Notes on counting

Peter J. Cameron
School of Mathematical Sciences
Queen Mary, University of London
Mile End Road
London E1 4NS
UK
p.j.cameron@qmul.ac.uk

Contents

Preface

Combinatorics is a subject which stands in an uneasy relation with the rest of mathematics, and has often been treated with scorn by traditional mathematicians. (Many people know Henry Whitehead's reported remark, "Combinatorics is the slums of topology".)

In defence of the subject, several eminent practitioners (notably Gian-Carlo Rota and André Joyal) have attempted to take at least part of combinatorics and re-formulate it as mathematics in the axiomatic, twentieth-century style. This has led to many important developments (matroid theory, the Möbius function, species) some of which are touched on here. In my view, though, this approach has not been completely successful, since combinatorics by its nature escapes any attempt to define it.

I find more congenial the view eloquently put by someone with impeccable credentials, Tim Gowers, in his paper "The two cultures of mathematics". He argues that, in combinatorics, it is *techniques* which play the role that big theorems do in more traditional mathematics.

Accordingly, these notes are not laden with theorems, big or small. If you need a particular binomial identity or the enumeration of a particular class of graphs, chances are you won't find it here. Instead, you may possibly find the technique which will help you to prove the identity or count the graphs yourself. (I have been asked by colleages such questions as "How many partially ordered sets can be obtained from the trivial poset by nesting and crossing?" or "How many orbits does a finite linear group have on n-tuples of vectors?" You won't find the answers here, but you will find the techniques needed to answer these questions.)

If you require a much more complete compendium, you are referred to the books by Goulden and Jackson and by Stanley listed in the bibliography. Stanley's book is particularly rich in exercises, which are the lifeblood of the subject.

These notes began as the course notes for the course MTHM C50, "Enumerative and asymptotic combinatorics", which I taught at Queen Mary, University of

London, in the spring of 2003. This is a second course in combinatorics for those who have already taken the equivalent of the undergraduate course MAS 219. The syllabus for the course reads:

1. Techniques: Inclusion-exclusion, recurrence relations and generating functions.
2. Subsets, partitions, permutations: binomial coefficients; partition, Bell, and Stirling numbers; derangements. q-analogues: Gaussian coefficients, q-binomial theorem.
3. Linear recurrence relations with constant coefficients.
4. Counting up to group action: Orbit-counting lemma, cycle index theorem.
5. Posets and Möbius inversion, Möbius function of projective space.
6. Asymptotic techniques: Order notation: O, o, $\sim$. Stirling's formula. Techniques from complex analysis including Hayman's Theorem.

I am grateful to the students on the course for their critical comments and for debugging the notes. (In particular, a solution by Pablo Spiga to one of the prize questions is included.) Any remaining errors are, of course, mine. Also, there are some topics included here which were not in the lecture course.

Peter J. Cameron
April 10, 2003

Chapter 1

Introduction

This course is about counting. Of course this doesn't mean just counting a single finite set. Usually, we have a family of finite sets indexed by a natural number n, and we want to find $F(n)$, the cardinality of the nth set in the family.

1.1 What is counting?

There are several kinds of answer to this question:

- An explicit formula (which may be more or less complicated, and in particular may involve a number of summations).

- A recurrence relation expressing $F(n)$ in terms of values of $F(m)$ for $m < n$.

- A closed form for a *generating function* for F. (The two types of generating function most often used are the *ordinary generating function* $\sum F(n)x^n$, and the *exponential generating function* $\sum F(n)x^n/n!$.) These are elements of the ring $\mathbb{Q}[[x]]$ of *formal power series*. They may or may not converge if a given non-zero complex number is substituted for x. (Formal power series are discussed further in the next section.)

  If a generating function converges, it is possible to find the coefficients by analytic methods (differentiation or contour integration).

- An asymptotic estimate for $F(n)$ is a function $G(n)$, typically expressed in terms of the standard functions of analysis, such that $F(n) - G(n)$ is of smaller order of magnitude than $G(n)$. (If $G(n)$ does not vanish, we can write this as $F(n)/G(n) \to 1$ as $n \to \infty$.) We write $F(n) \sim G(n)$ if this holds. This might be accompanied by an asymptotic estimate for $F(n) - G(n)$, and so on; we obtain an *asymptotic series* for F. (The basics of asymptotic analysis are described further in the third section of this chapter.)

- Related to counting combinatorial objects is the question of generating them. The first thing we might ask for is a system of sequential generation, where we can produce an ordered list of the objects. Again there are two possibilities.

  If the number of objects is $F(n)$, we might ask for a construction which, given i with $0 \le i \le F(n) - 1$, produces the ith object on the list directly.

  Alternatively, we may simply require a method of moving from each object to the next.

- We could also ask for a method for random generation of an object. If we have a technique for generating the ith object directly, we simply choose a random number in the range $\{0, \dots, F(n) - 1\}$ and generate the corresponding object. If not, we have to rely on other methods such as Markov chains.

Here are a few examples. These will be considered in more detail later in the course.

Example: subsets The number of subsets of $\{1, \dots, n\}$ is 2^n. Not only is this a simple formula to write down; it is easy to compute as well. At most $2\log_2 n$ integer multiplications are required.

To see this, write n in base 2: $n = 2^{a_1} + 2^{a_2} + \cdots + 2^{a_r}$, where $a_1 > \cdots > a_r$. Now we can compute 2^{2^i} for $1 \le i \le a_1$ by a_1 successive squarings (noting that $2^{2^{i+1}} = \left(2^{2^i}\right)^2$); then $2^n = (2^{2^{a_1}}) \cdots (2^{2^{a_r}})$ requires $r - 1$ further multiplications.

There is a simple recurrence relation for $F(n) = 2^n$, namely

$$F(0) = 1, \qquad F(n) = 2F(n-1) \text{ for } n \ge 1.$$

Using this, $F(n)$ can be found with just $n - 1$ integer doublings.

The ordinary generating function of the sequence (2^n) is $1/(1 - 2x)$, while the exponential generating function is $\exp(2x)$. (I will use $\exp(x)$ instead of e^x in these notes, except in some places involving calculus.)

No asymptotic estimate is needed, since we have a simple exact formula.

Choosing a random subset, or generating all subsets in order, are easily achieved by the following method. For each $i \in \{0, \ldots, 2^n - 1\}$, write i in base 2, producing a string of length n of zeros and ones. Now j belongs to the ith subset if and only if the jth symbol in the string is 1.

Example: permutations The number of permutations of $\{1, \ldots, n\}$ is $n!$, defined as usual as the product of the natural numbers from 1 to n. This formula is not so satisfactory, involving an n-fold product. It can be expressed in other ways, as a sum:

$$n! = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} (n-k)^n,$$

or as an integral:

$$n! = \int_0^\infty x^n \mathrm{e}^{-x}\, \mathrm{d}x.$$

Neither of these is easier to evaluate than the original definition.

The recurrence relation for $F(n) = n!$ is

$$F(0) = 1, \qquad F(n) = nF(n-1) \text{ for } n \geq 1.$$

This leads to the same method of evaluation as we saw earlier.

The ordinary generating function for $F(n) = n!$ fails to converge anywhere. The exponential generating function is $1/(1-x)$, convergent for $|x| < 1$.

As an example to show that convergence is not necessary for a power series to be useful, let

$$\left(1 + \sum_{n \geq 1} n! x^n\right)^{-1} = 1 - \sum_{n \geq 1} c(n) x^n.$$

Then $c(n)$ is the number of connected permutations on $\{1, \ldots, n\}$. (A permutation π is *connected* if there does not exist k with $1 \leq k \leq n-1$ such that π maps $\{1, \ldots, k\}$ to itself.)

An asymptotic estimate for $n!$ is given by *Stirling's formula*:

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{\mathrm{e}}\right)^n.$$

It is possible to generate permutations sequentially, or choose a random permutation, by a method similar to that for subsets.

Example: derangements A derangement is a permutation with no fixed points. Let $d(n)$ be the number of *derangements* of n.

There is a simple formula for $d(n)$: it is the nearest integer to $n!/\mathrm{e}$. This is also satisfactory as an asymptotic expression for $d(n)$; we can supplement it with the fact that $|d(n) - n!/\mathrm{e}| < 1/(n+1)$ for $n > 0$.

This formula is not very good for calculation, since it requires accurate knowledge of e and operations of real (rather than integer) arithmetic. There are, however, two recurrence relations for $d(n)$; the second, especially, leads to efficient calculation:

$$\begin{aligned} d(0) = 1, d(1) = 0, \qquad & d(n) = (n-1)(d(n-1) + d(n-2)) \text{ for } n \geq 2; \\ d(0) = 1, \qquad & d(n) = nd(n-1) + (-1)^n \text{ for } n \geq 1. \end{aligned}$$

The ordinary generating function for $d(n)$ fails to converge, but the exponential generating function is equal to $\exp(-x)/(1-x)$.

Since the probability that a random permutation is a derangement is about $1/\mathrm{e}$, we can choose a random derangement as follows: repeatedly choose a random permutation until a derangement is obtained. The expected number of choices necessary is about e.

Example: partitions The *partition number* $p(n)$ is the number of non-increasing sequences of positive integers with sum n. There is no simple formula for $p(n)$. However, quite a bit is known about it:

- The ordinary generating function is

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

- There is a recurrence relation:

$$p(n) = \sum (-1)^{k-1} p(n - k(3k-1)/2),$$

  where the sum is over all non-zero values of k, positive and negative, for which $n - k(3k-1)/2 \geq 0$. Thus,

$$p(n) = p(n-1) + p(n-2) - p(n-5) - p(n-7) + p(n-12) + \cdots,$$

  where there are about $\sqrt{8n/3}$ terms in the sum.

- The asymptotics of $p(n)$ are rather complicated, and were worked out by Hardy, Littlewood, and Rademacher:

  $$p(n) \sim \frac{1}{4n\sqrt{3}} \mathrm{e}^{\pi\sqrt{2n/3}}$$

  (more precise estimates, including a convergent series representation, exist).

Example: set partitions The *Bell number* $B(n)$ is the number of partitions of the set $\{1,\dots,n\}$. Again, no simple formula is known, and the asymptotics are very complicated. There is a recurrence relation,

$$B(n) = \sum_{k=1}^{n} \binom{n-1}{k-1} B(n-k),$$

and the exponential generating function is

$$\sum \frac{B(n)x^n}{n!} = \exp(\exp(x)-1).$$

Based on the recurrence one can derive a sequential generation algorithm.

1.2 Formal power series

Let R be a commutative ring with identity. A *formal power series* over R is just a function from the natural numbers to R; that is, an infinite sequence

$$r_0, r_1, r_2, \dots, r_n, \dots \tag{1.1}$$

of elements of R. We define addition and multiplication of such infinite series to make the set of formal power series into a ring. The definitions look more natural if we write the sequence (1.1) as

$$r_0 + r_1x + r_2x^2 + \cdots + r_nx^n + \cdots \tag{1.2}$$

The symbol x in this expression is just a dummy with no meaning; the “power” of x allows us to keep track of our place in the series. No infinite summation is actually involved! We denote the set of all formal power series by $R[[x]]$. If we had used a different symbol, say y, in the expression (1.2), we would write $R[[y]]$ instead. We often abbreviate (1.2) to

$$\sum_{n\geq 0} r_n x^n. \tag{1.3}$$

A *polynomial* is simply a formal power series in which all but finitely many of the terms are zero. The *degree* of a polynomial is the index of the last non-zero term. The set of polynomials is denoted by $R[x]$.

We define addition and multiplication of formal power series by

$$\begin{aligned}\left(\sum_{n\geq 0} r_n x^n\right) + \left(\sum_{n\geq 0} s_n x^n\right) &= \sum_{n\geq 0} (r_n + s_n) x^n,\\ \left(\sum_{n\geq 0} r_n x^n\right) \cdot \left(\sum_{n\geq 0} s_n x^n\right) &= \sum_{n\geq 0} t_n x^n,\end{aligned}$$

where

$$t_n = \sum_{k=0}^{n} r_k s_{n-k}.$$

Note that these operations involve only finite additions and multiplications of ring elements.

With these operations, $R[[x]]$ is a ring, and $R[x]$ a subring. We don't stop to prove this, as the verifications are routine.

Various other apparently "infinitary" operations can be defined which only involve finite sums and products. For example,

- Suppose that $f_0, f_1, \ldots \in R[[x]]$ have the property that the index of the smallest non-zero term in f_n tends to infinity with n. Then

  $$\sum_{n\geq 0} f_n$$

  is defined. In particular, if $f_n = r_n x^n$, the condition is satisfied, and this definition of the infinite sum agrees with our notation for the formal power series $\sum r_n x^n$.

- With the same conditions,

  $$\prod_{n\geq 0} (1 + f_n)$$

is defined: it is the sum of terms, each of which is the product of finitely many f_n (taking 1 from the remaining factors in the infinite product); and by assumption only finitely many such products contribute to the coefficient of x^n for any n.

- Let f and g be formal power series in which the constant term of g is zero. Then the result of substituting g into f is defined: if $f(x) = \sum r_n x^n$, then $f(g(x)) = \sum r_n g^n$.
- We can differentiate formal power series. The rule is, as you would expect,

$$\frac{\mathrm{d}}{\mathrm{d}x} \sum_{n \geq 0} a_n x^n = \sum_{n \geq 1} n a_n x^{n-1}.$$

  (No calculus needed, and no need to wonder if a function has a derivative!) The usual calculus rules for differentiating sums, products, and composite functions (the chain rule) are valid. Note that, if we differentiate n times and put $x = 0$ (that is, take the constant term), we obtain $n! a_n$.

A result which is important for enumeration is the following, though we are more concerned with the method of proof than the statement.

Proposition 1.1 *A formal power series is invertible if and only if its constant term is invertible.*

Proof Suppose that $f = \sum r_n x^n$ and $g = \sum s_n x^n$ satisfy $fg = 1$. Considering the term of degree zero, we see that $r_0 s_0 = 1$, so that r_0 is invertible.

Conversely, suppose that $r_0 s_0 = 1$, where $f = \sum r_n x^n$. The inverse $g = \sum s_n x^n$ must satisfy

$$\sum_{k=0}^{n} r_k s_{n-k} = 0$$

for $n > 0$; so

$$s_n = -s_0 \sum_{k=1}^{n} r_k s_{n-k}.$$

Thus the coefficients of g satisfy a linear recurrence relation, and can be determined recursively.

In general, knowledge of the inverse of a formal power series is equivalent to knowledge of a linear recurrence relation for its coefficients.

Example: Fibonacci numbers Let $f(x) = 1-x-x^2$. Then the coefficients of the inverse $(1-x-x^2)^{-1} = \sum s_n x^n$ satisfy the recurrence

$$s_0 = s_1 = 1, \qquad s_n = s_{n-1} + s_{n-2} \text{ for } n \geq 2;$$

in other words, they are the Fibonacci numbers.

For the purposes of enumeration, the coefficients of formal power series are usually integers or rational numbers. Often it is convenient to consider them as real numbers, and apply to them the processes of analysis.

For example, considering the Fibonacci numbers above, let α and β be the roots of the quadratic equation $z^2 - z - 1 = 0$: thus, $\alpha = (\sqrt{5}+1)/2$ and $\beta = (-\sqrt{5}+1)/2$. Then

$$\begin{aligned} \frac{1}{1-x-x^2} &= \frac{1}{\alpha-\beta}\left(\frac{\alpha}{1-\alpha x} - \frac{\beta}{1-\beta x}\right) \\ &= \frac{1}{\sqrt{5}}\left(\sum_{n\geq 0} \alpha^{n+1}x^n - \sum_{n\geq 0} \beta^{n+1}x^n\right); \end{aligned}$$

so the nth Fibonacci number is

$$F_n = \frac{1}{\sqrt{5}}(\alpha^{n+1} - \beta^{n+1}).$$

Since $|\beta| < 1$, we see that F_n is the nearest integer to $\alpha^{n+1}/\sqrt{5}$.

Particular formal power series of great importance include

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

1.3 Asymptotics

We introduce the notation for describing the asymptotic behaviour of functions here, though we will not do any serious asymptotic estimation for a while.

Let F and G be functions of the natural number n. For convenience we assume that G does not vanish. We write

- $F = O(G)$ if $F(n)/G(n)$ is bounded above as $n \to \infty$;
- $F = \Theta(G)$ if $F(n)/G(n)$ is bounded below as $n \to \infty$;
- $F = o(G)$ if $F(n)/G(n) \to 0$ as $n \to \infty$;
- $F \sim G$ if $F/G \to 1$ as $n \to \infty$.

Typically, F is a combinatorial enumeration function, and G a combination of standard functions of analysis. For example, Stirling's formula gives the asymptotics of the number of permutations of $\{1,\ldots,n\}$. We give the proof as an illustration.

Theorem 1.2

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$$

Proof Consider the graph of the function $y = \log x$ between $x = 1$ and $x = n$, together with the piecewise linear functions shown in Figure 1.1.

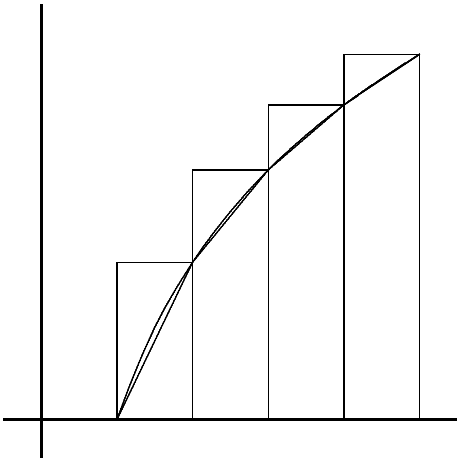

Figure 1.1: Stirling's formula

Let $f(x) = \log x$, let $g(x)$ be the function whose value is $\log m$ for $m \leq x < m+1$, and let $h(x)$ be the function defined by the polygon with vertices $(m, \log m)$, for $1 \leq m \leq n$. Clearly

$$\int_1^n g(x)\,\mathrm{d}x = \log 2 + \cdots + \log n = \log n!\,.$$

The difference between the integrals of g and h is the sum of the areas of triangles with base 1 and total height $\log n$; that is, $\frac{1}{2}\log n$.

Some calculus[1] shows that the difference between the integrals of f and g tends to a finite limit c as $n \to \infty$.

Finally, a simple integration shows that

$$\int_1^n f(x)\,\mathrm{d}x = n\log n - n + 1.$$

We conclude that

$$\log n! = n\log n - n + \tfrac{1}{2}\log n + (1-c) + o(1),$$

so that

$$n! \sim \frac{Cn^{n+1/2}}{\mathrm{e}^n}.$$

To identify the constant C, we can proceed as follows. Consider the integral

$$I_n = \int_0^{\pi/2} \sin^n x\,\mathrm{d}x.$$

Integration by parts shows that

$$I_n = \frac{n-1}{n} I_{n-2},$$

[1] Let $F(x) = f(x) - g(x)$. The convexity of $\log x$ shows that $F(x) \geq 0$ for all $x \in [m, m+1]$. For an upper bound we use the fact, a consequence of Taylor's Theorem, that

$$\log x \leq \log m + \frac{x-m}{m} \leq \log m + \frac{1}{m}$$

for $x \in [m, m+1]$. Then

$$F(x) = \log x - \log m - \log\left(1 + \frac{1}{m}\right)(x-m) \leq \frac{1}{m} - \log\left(1 + \frac{1}{m}\right) \leq \frac{1}{2m^2},$$

where the last inequality comes from another application of Taylor's Theorem which yields $\log(1+x) \geq x - x^2/2$ for $x \in [0,1]$. Now $\sum(1/m^2)$ converges, so the integral is bounded.

and hence

$$\begin{aligned} I_{2n} &= \frac{(2n)!\,\pi}{2^{2n+1}(n!)^2}, \\ I_{2n+1} &= \frac{2^{2n}(n!)^2}{(2n+1)!}. \end{aligned}$$

On the other hand,

$$I_{2n+2} \le I_{2n+1} \le I_{2n},$$

from which we get

$$\frac{(2n+1)\pi}{4(n+1)} \le \frac{2^{4n}(n!)^4}{(2n)!(2n+1)!} \le \frac{\pi}{2},$$

and so

$$\lim_{n\to\infty} \frac{2^{4n}(n!)^4}{(2n)!(2n+1)!} = \frac{\pi}{2}.$$

Puttiing $n! \sim Cn^{n+1/2}/\mathrm{e}^n$ in this result, we find that

$$\frac{C^2\mathrm{e}}{4} \lim_{n\to\infty} \left(1+\frac{1}{2n}\right)^{-(2n+3/2)} = \frac{\pi}{2},$$

so that $C = \sqrt{2\pi}$.

The last part of this proof is taken from Alan Slomson, *An Introduction to Combinatorics*, Chapman and Hall 1991. It is more-or-less the proof of Wallis' product formula for π.

The series $G_0(n) + G_1(n) + G_2(n) + \cdots$ is an *asymptotic series* for $F(n)$ if

$$F(n) - \sum_{j=0}^{i-1} G_j(n) \sim G_i(n)$$

for $i \ge 0$. (So in particular $F(n) \sim G_0(n)$, $F(n) - G_0(n) \sim G_1(n)$, and so on. Note that $G_i(n) = o(G_{i-1}(n))$ for all i.)

Warnings:

- an asymptotic series is not necessarily convergent;

- it is not necessarily the case that taking more terms in the series gives a better approximation to $F(n)$ for a fixed n.

For example, Stirling's formula can be extended to an asymptotic series for $n!$, namely

$$\sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n\left(1+\frac{1}{12n}+\frac{1}{288n^2}+\cdots\right).$$

Regarding a generating function for a sequence as a function of a real or complex variable is a powerful method for studying the asymptotic behaviour of the sequence. We will see examples of this later; here is a simple observation.

Suppose that $A(z)=\sum a_nz^n$ defines a function which is analytic in some neighbourhood of the origin in the complex plane. Suppose that the smallest modulus of a singularity of $A(z)$ is R. Then $\limsup a_n^{1/n}=1/R$, so a_n is bounded by $(c+\varepsilon)^n$ but not by $(c-\varepsilon)^n$ for large n, where $c=1/R$.

For example, we saw that the generating function of the Fibonacci numbers is $1/(1-z-z^2)$. So these numbers grow roughly like α^n, where α is the reciprocal of the smaller root of $1-z-z^2=0$, namely $\alpha=(1+\sqrt{5})/2$.

On the other hand, if $A(z)$ is analytic everywhere, then $a_n\le\varepsilon^n$ for $n>n_0(\varepsilon)$, for any positive ε. Indeed, $a_n=o(\varepsilon^n)$ for any positive ε.

For example, if $B(n)$ is the nth Bell number, then

$$\sum_{n\ge0}\frac{B(n)z^n}{n!}=\mathrm{e}^{\mathrm{e}^z-1},$$

which is analytic everywhere. So $B(n)=o(\varepsilon^n n!)$, for any positive ε.

1.4 Complexity

A formula like 2^n (the number of subsets of an n-set) can be evaluated quickly for a given value of n. A more complicated formula with multiple sums and products will take longer to calculate. We could regard a formula which takes more time to evaluate than it would take to generate all the objects and count them as being useless in practice, even if it has theoretical value.

Traditional computational complexity theory refers to decision problems, where the answer is just "yes" or "no" (for example, "Does this graph have a Hamiltonian circuit?"). The size of an instance of a problem is measured by the number of bits of data required to specify the problem (for example, $n(n-1)/2$ bits to specify a graph on n vertices). Then the time complexity of a problem is the function

f, where $f(n)$ is the maximum number of steps required by a Turing machine to compute the answer for an instance of size n. To allow for variations in the format of the input data and in the exact specification of a Turing machine, complexity classes are defined with a broad brush: for example, P (or “polynomial-time”) consists of all problems whose time complexity is at most n^c for some constant c. (For more details, see Garey and Johnson, *Computers and Intractability*.)

For counting problems, the answer is a number rather than a single Boolean value (for example, “How many Hamiltonian circuits does this graph have?”). Complexity theorists have defined the complexity class $\#\mathsf{P}$ (“number-P”) for this purpose.

Even this class is not really appropriate for counting problems of the type we mostly consider. Consider, for example, the question “How many partitions does an n-set have?” The input data is the integer n, which (if written in base 2) requires only $m = \lceil 1 + \log_2 n \rceil$ bits to specify. The question asks us to calculate the Bell number $B(n)$, which is greater than 2^{n-1} for $n > 2$, and so it takes time exponential in m simply to write down the answer! To get round this difficulty, it is usual to pretend that the size of the input data is actually n rather than $\log n$. (We can imagine that n is given by writing n consecutive 1s on the input tape of the Turing machine, that is, by writing n as a tally rather than in base 2.)

We have seen that computing 2^n (the number of subsets of an n-set) requires only $O(\log n)$ integer multiplications. But the integers may have as many as n digits, so each multiplication takes $O(n)$ Turing machine steps. Similarly, the solution to a recurrence relation can be computed in time polynomial in n, provided that each individual computation can be.

On the other hand, a method which involves generating and testing every subset or permutation will take exponentially long, even if the generation and testing can be done efficiently.

A notion of complexity relevant to this situation is the polynomial delay model, which asks that the time required to generate each object should be at most n^c for some fixed c, even if the number of objects to be generated is greater than polynomial.

Of course, it is easy to produce combinatorial problems whose solution grows faster than, say, the exponential of a polynomial. For example, how many intersecting families of subsets of an n-set are there? The total number, for n odd, lies between $2^{2^{n-1}}$ and 2^{2^n}, so that even writing down the answer takes time exponential in n.

We will not consider complexity questions further in this course.

Exercises

1.1. Prove directly that $(1-x)^{-1} = \sum_{n\geq 0} x^n$ (in the ring of formal power series).

1.2. Suppose that a collection of complex power series all define functions analytic in some neighbourhood of the origin, and satisfy some identity there. Are we allowed to conclude that this identity holds between the series regarded as formal power series?

1.3. Suppose that $A(x)$, $B(x)$ and $C(x)$ are the exponential generating functions of sequences (a_n), (b_n) and (c_n) respectively. Show that $A(x)B(x) = C(x)$ if and only if

$$c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k},$$

where

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

1.4. Show that the identity $\exp(\log(1+x)) = 1+x$ between formal power series is equivalent to the equation

$$\sum_{k=1}^{n} \frac{(-1)^k}{k!} T(n,k) = 0,$$

for $n > 1$, where $T(n,k)$ is computed as follows: write n as an ordered sum of k positive integers $a_1,\ldots,a_k$ in all possible ways; for each such expression compute the product $a_1\cdots a_k$; and sum the reciprocals of the resulting numbers.

What is the analogous interpretation of the identity $\log(1+(\exp(x)-1)) = x$?

1.5. Show that the identity $\exp(x+y) = \exp(x)\exp(y)$ is equivalent to the Binomial Theorem for all positive integer exponents.

1.6. Prove that $n^k = o(c^n)$ for any constants $k > 0$ and $c > 1$, and that $\log n = o(n^{\varepsilon})$ for any $\varepsilon > 0$.

1.7. Let $f(n)$ be the number of partitions of an n-set into parts of size 2.

(a) Prove that

$$f(n) = \begin{cases} 0 & \text{if } n \text{ is odd;} \\ 1\cdot 3\cdot 5\cdots(n-1) & \text{if } n \text{ is even.} \end{cases}$$

(b) Prove that the exponential generating function for the sequence $(f(n))$ is $\exp(x^2)$.

(c) Prove that

$$f(n) \sim \sqrt{2}\left(\frac{2m}{\mathrm{e}}\right)^m$$

for $n = 2m$.

1.8. Show that it is possible to generate all subsets of $\{1, \ldots, n\}$ successively in such a way that each subset differs from its predecessor by the addition or removal of precisely one element. (Such a sequence is known as a *Gray code*.)

Chapter 2

Subsets, partitions and permutations

The basic objects of combinatorics are subsets, partitions and permutations. In this chapter, we consider the problem of counting these. The counting functions have two parameters: n, the size of the underlying set; and k, a measure of the object in question (the number of elements of a subset, parts of a partition, or cycles of a permutation respectively).

2.1 Subsets

The number of k-element subsets of the set $\{1,\ldots,n\}$ is the *binomial coefficient*

$$\binom{n}{k} = \begin{cases} 0 & \text{if } k < 0 \text{ or } k > n; \\ \dfrac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} & \text{if } 0 \le k \le n. \end{cases}$$

For, if $0 \le k \le n$, there are $n(n-1)\cdots(n-k)$ ways to choose in order k distinct elements from $\{1,\ldots,n\}$; each k-element subset is obtained from $k!$ such ordered selections. The result for $k < 0$ or $k > n$ is clear.

Proposition 2.1 *The recurrence relation for the binomial coefficients is*

$$\binom{n}{0} = \binom{n}{n} = 1, \qquad \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \text{ for } 0 < k < n.$$

Proof Partition the k-element subsets into two classes: those containing n (which have the form $\{n\} \cup L$, where L is a $(k-1)$-element subset of $\{1,\ldots,n-1\}$, and so are $\binom{n-1}{k-1}$ in number); and those not containing n (which are k-element subsets of $\{1,\ldots,n-1\}$, and so are $\binom{n-1}{k}$ in number).

The *Binomial Theorem* for natural number exponents n asserts:

Proposition 2.2 $(x+y)^n = \displaystyle\sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$

Proof The proof is straightforward. On the left we have the product

$$(x+y)(x+y)\cdots(x+y) \qquad (n \text{ factors});$$

multiplying this out we get the sum of 2^n terms, each of which is obtained by choosing y from a subset of the factors and x from the remainder. There are $\binom{n}{k}$ subsets of size k, and each contributes a term $x^{n-k}y^k$ to the sum, for $k = 0,\ldots,n$.

The Binomial Theorem can be looked at in various ways. From one point of view, it gives the generating function for the binomial coefficients $\binom{n}{k}$ for fixed n:

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

Since the binomial coefficients have two indices, we could ask for a two-variable generating function:

$$\begin{aligned} \sum_{n\geq 0}\sum_{k\geq 0} \binom{n}{k} x^n y^k &= \sum_{n\geq 0} x^n (1+y)^n \\ &= \frac{1}{1-x(1+y)}. \end{aligned}$$

If we expand this in powers of y, we obtain

$$\begin{aligned} \frac{1}{(1-x)-xy} &= \frac{1}{1-x}\cdot\frac{1}{1-(x/(1-x))y} \\ &= \sum_{k\geq 0} \left(\frac{x^k}{(1-x)^{k+1}}\right) y^k, \end{aligned}$$

so that we have the following:

Proposition 2.3 $\displaystyle\sum_{n\geq k}\binom{n}{k}x^n = \frac{x^k}{(1-x)^{k+1}}.$

Our next observation on the Binomial Theorem concerns *Pascal's Triangle,* the triangle whose nth row contains the numbers $\binom{n}{k}$ for $0\leq k\leq n$. (Despite the name, this triangle was not invented by Pascal but occurs in earlier Chinese sources. Figure 2.1 shows the triangle as given in Chu Shi-Chieh's *Ssu Yuan Yü Chien*, dated 1303.) The recurrence relation shows that each entry of the triangle is the sum of the two above it.

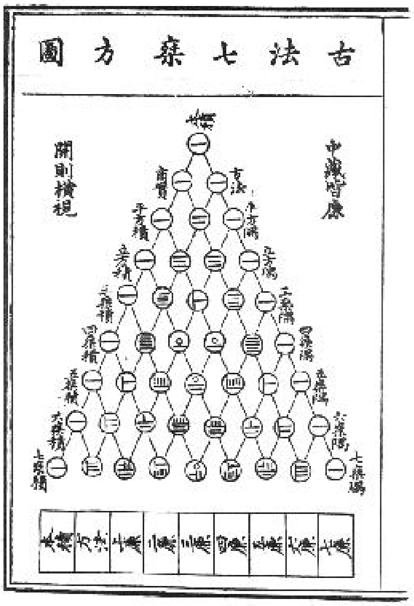

Figure 2.1: Chu Shi-Chieh's Triangle

At risk of making the triangle asymmetric, we turn it into a matrix $B=(b_{nk})$, where $b_{nk}=\binom{n}{k}$ for $n,k\geq 0$. This infinite matrix is lower triangular, with ones on the diagonal. Now when two lower triangular matrices are multiplied, each term of the product is only a finite sum: the (n,k) entry of BC is $\sum_m b_{nm}c_{mk}$, and this is non-zero only for $k\leq m\leq n$. In particular, we can ask "What is the inverse of B?"

The *signed matrix of binomial coefficients* is the matrix B^* with (n,k) entry $(-1)^{n-k}\binom{n}{k}$. That is, it is the same as B except that signs of alternate terms are changed in a chessboard pattern. Now:

Proposition 2.4 *The inverse of the matrix B of binomial coefficients is the matrix B^* of signed binomial coefficients.*

Proof We consider the vector space of polynomials (over $\mathbb{R}$). There is a natural basis consisting of the polynomials $1, x, x^2, \ldots$. Now, since

$$(1+x)^n = \sum_k \binom{n}{k} x^k,$$

we see that B represents the change of basis to $1, y, y^2, \ldots$, where $y = 1+x$. Hence the inverse of B represents the basis change in the other direction, given by $x = y-1$. Since

$$(y-1)^n = \sum_k (-1)^{n-k} \binom{n}{k} y^k,$$

the matrix of this basis change is B^*.

The other aspect of the Binomial Theorem is its generalisation to arbitrary real exponents (due to Isaac Newton). This depends on a revised definition of the binomial coefficients.

Let a be an arbitrary real (or complex) number, and k a non-negative integer. Define

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

Note that this agrees with the previous definition in the case when n is a non-negative integer, since if $k > n$ then one of the factors in the numerator is zero. We do not define this version of the binomial coefficients if k is not a natural number.

Now the *Binomial Theorem* asserts that, for any real number a, we have

$$(1+x)^a = \sum_{k \geq 0} \binom{a}{k} x^k. \tag{2.1}$$

Is this a theorem or a definition? If we regard it as an equation connecting real functions (where the left-hand side is defined by

$$(1+x)^a = \exp(a \log(1+x)), \tag{2.2}$$

and the series on the right-hand side is convergent for $|x| < 1$), it is a theorem, and was understood by Newton in this form. As an equation connecting formal power series, we may follow the same approach, or we may instead choose to regard

(2.1) as the definition and (2.2) as the theorem, according to taste. Whichever approach we take, we need to know that the laws of exponents hold:

$$\begin{aligned}(1+x)^a\cdot(1+y)^a &= (1+(x+y+xy))^a,\\ (1+x)^{a+b} &= (1+x)^a\cdot(1+x)^b,\\ (1+x)^{ab} &= ((1+x)^a)^b.\end{aligned}$$

If (2.1) is our definition, these verifications will reduce to identities between binomial coefficients; if (2.2) is the definition, they depend on properties of the power series for exp and log, defined as in the last chapter.

Binomial coefficients can be estimated by using Stirling's formula. (See Exercise 2.4, for example.)

The *Central Limit Theorem* from probability theory can also be used to get estimates for binomial coeffients. Suppose that a fair coin is tossed n times. Then the probability of obtaining k heads is equal to $\binom{n}{k}/2^n$. Now the number of heads is a binomial random variable X; so we have

$$\mathbb{P}(X=k)=\binom{n}{k}\Big/2^n. \tag{2.3}$$

According to the Central Limit Theorem, if n is large then X is approximated by a normal random variable Y with the same expected value $n/2$ and variance $n/4$. The probability density function of Y is given by

$$f_Y(y)=\frac{1}{\sqrt{\pi n/2}}\mathrm{e}^{2(k-n/2)^2/n}. \tag{2.4}$$

If $k=n/2+O(\sqrt{n})$ and $n\to\infty$, then a precise statement of the Central Limit Theorem shows that (2.4) gives an asymptotic formula for (2.3). In particular, when $k=n/2$, we obtain the result of Exercise 2.4.

2.2 Partitions

The *Bell number* $B(n)$ is the number of partitions of the set $\{1,\dots,n\}$. There is a related "unlabelled" counting number $p(n)$, the *partition number*, which is the number of partitions of the number n (that is, lists in non-increasing order of positive integers with sum n). Thus, given any set partition, the list of sizes of its parts is a number partition; and two set partitions are equivalent under relabelling the elements of the underlying set (that is, under permutations of $\{1,\ldots,n\}$) if and only if the corresponding number partitions are equal.

What would be the analogous "unlabelled" counting function for subsets? Two subsets of $\{1,\ldots,n\}$ are equivalent under permutations if and only if they have the same cardinality; so the unlabelled counting function f for subsets would be simply $f(n)=n+1$.

Set partitions

The *Stirling numbers of the second kind*, denoted by $S(n,k)$, are defined by the rule that $S(n,k)$ is the number of partitions of $\{1,\ldots,n\}$ into k parts if $1\le n\le k$, and zero otherwise. Clearly we have

$$\sum_{k=1}^{n} S(n,k)=B(n),$$

where the Bell number $B(n)$ is the total number of partitions of $\{1,\ldots,n\}$.

Proposition 2.5 *The recurrence relation for the Stirling numbers is*

$$S(n,1)=S(n,n)=1, \qquad S(n,k)=S(n-1,k-1)+kS(n-1,k) \textit{ for } 1<k<n.$$

Proof We split the partitions into two classes: those for which $\{n\}$ is a single part (obtained by adjoining this part to a partition of $\{1,\ldots,n-1\}$ into $k-1$ parts), and the remainder (obtained by taking a partition of $\{1,\ldots,n-1\}$ into k parts, selecting one part, and adding n to it).

Proposition 2.6 *(a) The Stirling numbers satisfy the recurrence*

$$S(n,k)=\sum_{i=1}^{n-1}\binom{n-1}{i-1}S(n-i,k-1).$$

(b) The Bell numbers saisfy the recurrence

$$B(n)=\sum_{i=1}^{n}\binom{n-1}{i-1}B(n-i).$$

Proof Consider the part containing n of an arbitrary partition with k parts; suppose that it has cardinality i. Then there are $\binom{n-1}{k-1}$ choices for the remaining $i-1$ elements in this part, and $S(n-i,k-1)$ partitions of the remaining $n-i$ elements into $k-1$ parts. This proves (a); the proof of (b) is almost identical.

The Stirling numbers also have the following property. Let $(x)_k$ denote the polynomial $x(x-1)\cdots(x-k+1)$.

Proposition 2.7 $x^n = \sum_{k=1}^{n} S(n,k)(x)_k.$

Proof We prove this first when x is a positive integer. We take a set X with x elements, and count the number of n-tuples of elements of x. The total number is of course x^n. We now count them another way. Given an n-tuple $(x_1,\ldots,x_n)$, we define an equivalence relation on $\{1,\ldots,n\}$ by $i \equiv j$ if and only if $x_i = x_j$. If this relation has k different classes, then there are k distinct elements among $x_1,\ldots,x_n$, say $y_1,\ldots,y_k$ (listed in order). The choice of the partition and the k-tuple $(y_1,\ldots,y_k)$ uniquely determines $(x_1,\ldots,x_n)$. So the number of n-tuples is given by the right-hand expression also.

Now this equation between two polynomials of degree n holds for any positive integer x, so it must be a polynomial identity.

Stirling numbers are involved in the substitution of $\exp(x)-1$ for x in formal power series. The result depends on the following lemma:

Lemma 2.8

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

Proof The proof is by induction on k, the result being true when $k=1$ since $S(n,1)=1$. Suppose that it holds when $k=l-1$. Then (setting $S(n,k)=0$ if $n<k$) we have

$$\begin{aligned}
\frac{(\exp(x)-1)^l}{l!} &= \frac{1}{l}\cdot(\exp(x)-1)\cdot\frac{(\exp(x)-1)^{l-1}}{(l-1)!} \\
&= \frac{1}{l}\left(\sum_{n\geq 1}\frac{x^n}{n!}\right)\cdot\left(\sum_{n\geq 1}\frac{S(n,l-1)x^n}{n!}\right).
\end{aligned}$$

The coefficient of $x^n/n!$ here is

$$\begin{aligned}\frac{n!}{l}\sum_{i=1}^{n-1}\frac{1}{i!}\cdot\frac{S(n-i,l-1)}{(n-i)!} &= \frac{1}{l}\sum_{i=1}^{n-1}\binom{n}{i}S(n-i,l-1)\\ &= \frac{1}{l}(S(n+1,l)-S(n,l-1)),\end{aligned}$$

using the recurrence relation of Proposition 2.6(a). Finally, the recurrence relation of Proposition 2.5 shows that this is $S(n,l)$, as required.

Proposition 2.9 *Let $(a_0,a_1,\ldots)$ and $(b_0,b_1,\ldots)$ be two sequences of numbers, with exponential generating functions $A(x)$ and $B(x)$ respectively. Then the following two conditions are equivalent:*

(a) $b_0 = a_0$ and $b_n = \sum_{k=1}^{n} S(n,k)a_k$ for $n \geq 1$;

(b) $B(x) = A(\exp(x)-1)$.

Proof Suppose that (a) holds. Without loss of generality we may assume that $a_0 = b_0 = 0$. Then

$$\begin{aligned}B(x) &= \sum_{n\geq 1}\frac{b_n x^n}{n!}\\ &= \sum_{n\geq 1}\frac{x^n}{n!}\sum_{k=1}^{n}S(n,k)a_k\\ &= \sum_{k\geq 1}a_k\sum_{n\geq k}\frac{S(n,k)x^n}{n!}\\ &= \sum_{k\geq 1}\frac{a_k(\exp(x)-1)^k}{k!}\\ &= A(\exp(x)-1),\end{aligned}$$

by Lemma 2.8.

The converse is proved by reversing the argument.

Corollary 2.10 *The exponential generating function for the Bell numbers is*

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

Proof Apply Proposition 2.9 to the sequence with $a_n = 1$ for all n; or sum the equation of Lemma 2.8 over k.

Number partitions

The partition number $p(n)$ is the number of partitions of an n-set, up to permutations of the set.

The key to evaluating $p(n)$ is its generating function:

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

For $(1-x^k)^{-1} = 1+x^k+x^{2k}+\cdots$. Thus a term in x^n in the product, with coefficient 1, arises from every expression $n=\sum c_k k$, where the c_k are non-negative integers, all but finitely many equal to zero. This number is $p(n)$, since we can regard $n=\sum c_k k$ as an alternative expression for a partition of n.

We will use this in the next chapter to give a recurrence relation for $p(n)$.

2.3 Permutations

A permutation of $\{1,\ldots,n\}$ is a bijective function from this set to itself.

In the nineteenth century, a more logical terminology was used. Such a function was called a substitution, while a permutation was a sequence $(a_1,a_2,\ldots,a_n)$ containing each element of the set precisely once. Since there is a natural ordering of $\{1,2,\ldots,n\}$, there is a one-to-one correspondence between "permutations" and "substitutions": the sequence $(a_1,a_2,\ldots,a_n)$ corresponds to the function $\pi : i \mapsto a_i$, for $i=1,\ldots,n$.

The correspondence between permutations and total orderings of an n-set has profound consequences for a number of enumeration problems. For now we return to the usage "permutation = bijective function". We refer to the sequence $(a_1,\ldots,a_n)$ as the *passive form* of the permutation π in the last paragraph; the function is the *active form* of the permutation.

Following the conventions of algebra, we write a permutation on the right of its argument, so that $i\pi$ is the image of i under the permutation π (that is, the ith term of the passive form of π).

The set of permutations of $\{1,\ldots,n\}$, with the operation of composition, is a group, called the *symmetric group S_n*. Products, identity, and inverses of permutations always refer to the operations in this group.

Unlabelled permutations

As for partitions, we can consider unlabelled or labelled permutations, that is, permutations of an n-set or equivalence classes of permutations. We dispose of unlabelled permutations first.

Two permutations π_1 and π_2 of $\{1,\ldots,n\}$ are equivalent if there is a bijection σ of $\{1,\ldots,n\}$ (that is, a permutation!) such that, for all $i \in \{1,\ldots,n\}$, we have

$$(i\sigma)\pi_2 = j\sigma \quad \text{if and only if} \quad i\pi_1 = j,$$

in other words, $i\sigma\pi_2 = i\pi_1\sigma$ for all i, so that $\pi_2 = \sigma^{-1}\pi_1\sigma$. Thus, this equivalence relation is the algebraic relation of *conjugacy* in the symmetric group; the unlabelled permutations are conjugacy classes of S_n.

Now recall the *cycle decomposition* of permutations:

> Any permutation of a finite set can be written as the disjoint union of cycles, uniquely up to the order of the factors and the choices of starting points of the cycles.

Moreover,

> Two permutations are equivalent if and only if the lists of cycle lengths of the two permutations (written in non-increasing order) are equal.

Thus equivalence classes of permutations correspond to partitions of the integer n. This means that the enumeration theory for "unlabelled permutations" is the same as that for "unlabelled partitions", discussed in the last section.

Labelled permutations

The *parity* of a permutation π of $\{1,\ldots,n\}$ is defined as the parity of $n-k$, where k is the number of cycles of π (in its decomposition as a product of distinct cycles). The *sign* of π is $(-1)^p$, where p is the parity of π.

Parity and sign have various important algebraic properties. For example,

- the parity of π is equal to the parity of the number of factors in any expression for π as a product of disjoint cycles;
- parity is a homomorphism from the symmetric group S_n to the group $\mathbb{Z}/(2)$ of integers mod 2, and hence sign is a homomorphism to the multiplicative group $\{\pm 1\}$.

- For $n > 1$, these homomorphisms are onto; their kernel (the set of permutations of even parity, or of sign $+1$) is a normal subgroup of index 2 in S_n, called the *alternating group* A_n.

The *Stirling numbers of the first kind* are defined by the rule that $s(n,k)$ is $(-1)^{n-k}$ times the number of permutations of $\{1,\ldots,n\}$ having k cycles. Sometimes the number of such permutations is referred to as the *unsigned Stirling number*.

Clearly we have

$$\sum_{k=1}^{n} |s(n,k)| = n!\,.$$

Slightly less obviously,

$$\sum_{k=1}^{n} s(n,k) = 0$$

for $n > 1$. The algebraic proof of this depends on the fact that sign is a homomorphism to $\{\pm 1\}$, so that the two values are taken equally often. We will see a combinatorial proof later.

Proposition 2.11 *The recurrence relation for the Stirling numbers is*

$$\begin{aligned} &s(n,1) = (-1)^{n-1}(n-1)!, \quad s(n,n) = 1, \\ &s(n,k) = s(n-1,k-1) - (n-1)s(n-1,k) \textit{ for } 1 < k < n. \end{aligned}$$

Proof We split the permutations into two classes: those for which (n) is a single part (obtained by adjoining this cycle to a permutation of $\{1,\ldots,n-1\}$ with $k-1$ cycles), and the remainder (obtained by taking a permutation of $\{1,\ldots,n-1\}$ with k cycles and interpolating n at some position in one of the cycles). The second construction, but not the first, changes the sign of the permutations.

To see that there are $(n-1)!$ permutations with a single cycle, note that if we choose to start the cycle with 1 then the remaining $n-1$ elements can be written into the cycle in any order.

Note that, if we instead define $s(n,0)$ and $s(n,n+1)$ to be equal to 0 for $n \geq 1$, then the recurrence holds also for $k = 1$ and $k = n$. We use this below.

The generating function is given by the following result:

Proposition 2.12 $\displaystyle\sum_{k=1}^{n} s(n,k)x^k = (x)_n.$

Proof The result is clear for $n = 1$, Suppose that it holds for $n = m-1$.

$$\begin{aligned}\sum_{k=1}^{m} s(m,k)x^k &= \sum_{k=1}^{m} s(m-1,k-1)x^k - \sum_{k=1}^{m}(m-1)s(m-1,k)x^k \\ &= (x-m+1)(x)_{m-1} \\ &= (x)_m.\end{aligned}$$

Note that substituting $x = 1$ into this equation shows that $\sum_k s(n,k) = 0$ for $n \geq 2$.

Corollary 2.13 *The triangular matrices S_1 and S_2 whose entries are the Stirling numbers of the first and second kinds are inverses of each other.*

Proof Propositions 2.7 and 2.12 show that S_1 and S_2 are the transition matrices between the bases $(x^n : n \geq 1)$ and $((x)_n\ n \geq 1)$ of the space of real polynomials with constant term zero.

Proposition 2.14 *Let $(a_0, a_1, \ldots)$ and $(b_0, b_1, \ldots)$ be two sequences of numbers, with exponential generating functions $A(x)$ and $B(x)$ respectively. Then the following two conditions are equivalent:*

(a) $b_0 = a_0$ and $b_n = \sum_{k=1}^{n} s(n,k)a_k$ for $n \geq 1$;

(b) $B(x) = A(\log(1+x))$.

Proof This is the “inverse” of Proposition 2.9.

We have counted permutations by number of cycles. A more refined count is by the list of cycle lengths.

Let $c_k(\pi)$ be the number of k-cycles in the cycle decomposition of π.

Proposition 2.15 *The size of the conjugacy class of π in S_n is*

$$\frac{n!}{\prod_k k^{c_k(\pi)} c_k(\pi)!}.$$

Proof Write out the pattern for the cycle structure of a permutation with $c_k(\pi)$ cycles of length k for all k, leaving blank the entries in the cycles. There are $n!$ ways of entering the numbers $1,\ldots,n$ in the pattern. However, each cycle of length k can be written in k different ways, since the cycle can start at any point; and the cycles of length k can be written in any of the $c_k(\pi)!$ possible orders. So the number of ways of entering the numbers $1,\ldots,n$ giving rise to each permutation in the conjugacy class is $\prod k^{c_k(\pi)} c_k(\pi)!$.

The *cycle index* of the symmetric group S_n is the generating function for the numbers $c_k(\pi)$, for $k = 1,\ldots,n$. By convention it is normalised by dividing by $n!$. Thus,

$$Z(S_n) = \sum_{\pi\in S_n} \prod_{k=1}^{n} s_k^{c_k(\pi)}.$$

Because of the normalisation, this can be thought of as the probability generating function for the cycle structure of a random permutation: that is, the coefficient of the monomial $\prod s_k^{a_k}$ (where $\sum kc_k = n$) is the probability that a random permutation π has $c_k(\pi) = a_k$ for $k = 1,\ldots,n$ — this is

$$\frac{1}{\prod_k k^{a_k} a_k!}.$$

One result which we will meet later is the following. We adopt the convention that $Z(S_0) = 1$.

Proposition 2.16 $\displaystyle\sum_{n\geq 0} Z(S_n) = \exp\left(\sum_{k\geq 1} \frac{s_k}{k}\right)$.

Proof The left-hand side is equal to

$$\begin{aligned}
\sum_{n\geq 0} \sum_{\sum a_k = n} \prod_{k\geq 1} \frac{s_k^{a_k}}{k^{a_k} a_k!} &= \sum_{a_1,a_2,\ldots} \prod_{k\geq 1} \frac{s_k^{a_k}}{k^{a_k} a_k!} \\
&= \prod_{k\geq 1} \sum_{a\geq 0} \frac{s_k^a}{k^a a!} \\
&= \prod_{k\geq 1} \exp\left(\frac{s_k}{k}\right) \\
&= \exp\left(\sum_{k\geq 1} \frac{s_k}{k}\right)
\end{aligned}$$

as required. (The sum on the right-hand side of the first line is over all infinite sequences of natural numbers $(a_1, a_2, \ldots)$ with only finitely many entries non-zero.)

We will see much more about cycle index in the chapter on orbit counting.

2.4 More on formal power series

The enumeration of subsets and partitions makes an unexpected appearance in the rules for differentiating products and composites of formal power series. In fact, the formulae below work as well for n-times differentiable functions in the usual sense of calculus, since the depend only on the standard rules for differentiating sums and products and the Chain Rule.

For brevity, we use $f^{(n)}(x)$ for the result of differentiating $f(x)$ n times, and write $f'(x)$ for $f^{(1)}(x)$.

Products The standard product rule

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x)$$

extends to *Leibniz's rule*:

Proposition 2.17

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(f(x)g(x)) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x)g^{(n-k)}(x).$$

Proof The proof is by induction. By the product rule, terms in $f^{(k)}(x)g^{(n-k)}(x)$ arise by differentiating terms in $f^{(k-1)}(x)g^{(n-k)}(x)$ or $f^{(k)}(x)g^{(n-k-1)}(x)$, so the coefficient of $f^{(k)}(x)g^{(n-k)}(x)$ is

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}.$$

Taking $f(x) = \mathrm{e}^{ax}$ and $g(x) = \mathrm{e}^{bx}$, we obtain

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

the Binomial Theorem for positive integer exponents. Similarly, taking $f(x) = x^a$ and $g(x) = x^b$, we obtain

$$(a+b)_{(n)} = \sum_{k=0}^{n} \binom{n}{k} a_{(k)} b_{(n-k)}.$$

Substitution The Chain Rule tells us that

$$\frac{\mathrm{d}}{\mathrm{d}x} f(g(x)) = f'(g(x))g'(x).$$

As we have seen, the substitution of g in f is valid provided that $g(0) = 0$.

The generalisation of this to repeated derivatives is *Faà di Bruno's rule*. If $a_1, \ldots, a_k$ are positive integers with sum n, let $P(n; a_1, \ldots, a_k)$ be the number of partitions of $\{1, \ldots, n\}$ into parts of size $a_1, \ldots, a_k$.

Proposition 2.18

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} f(g(x)) = \sum_{a_1+\cdots+a_k=n} P(n; a_1, \ldots, a_k) f^{(k)}(g(x)) g^{(a_1)}(x) \cdots g^{(a_k)}(x).$$

Proof Again by induction. Suppose that we have a bijection between partitions of $\{1, \ldots, n\}$ and terms in the nth derivative of $f(g(x))$. When we differentiate the term $f^{(k)}(g(x))g^{a_1}(x) \cdots g^{(a_k)}(x)$, corresponding to a partition of $\{1, \ldots, n\}$ into parts of sizes $a_1, \ldots, a_k$, we obtain $k+1$ terms:

- $f^{(k+1)}(g(x))g^{a_1}(x) \cdots g^{(a_k)}(x)g'(x)$, corresponding to the partition of $\{1, \ldots, n+1\}$ in which $n+1$ is a singleton part;
- $f^{(k)}(g(x))g^{a_1}(x) \cdots g^{(a_i+1)}(x) \cdots g^{(a_k)}(x)$, in which $n+1$ is adjoined to the ith part of the partition.

So each partition of $\{1, \ldots, n+1\}$ corresponds to a unique term in the sum, and we are done.

For example, we have

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} f(\exp(x) - 1) = \sum_{k=1}^{n} S(n,k) f^{(k)}(\exp(x) - 1) \exp(kx),$$

since the sum of $P(n; a_1, \ldots, a_k)$ over all $(a_1, \ldots, a_k)$ with fixed n and k is just the number $S(n,k)$ of partitions with k parts. Putting $x = 0$ we obtain the formula

$$b_n = \sum_{k=1}^{n} S(n,k) a_k$$

relating the coefficients of $f(x)$ and $f(\exp(x) - 1)$.

Exercises

2.1. Show that the number of ways of selecting k objects from a set of n distinguished objects, if we allow the same object to be chosen more than once and pay no attention to the order in which the choices are made, is $\binom{n+k-1}{k}$.

2.2. Prove that, if n is even, then

$$\frac{2^n}{n+1} \leq \binom{n}{n/2} \leq 2^n.$$

Use Stirling's formula to prove that

$$\binom{n}{n/2} \sim \frac{2^n}{\sqrt{\pi n/2}}.$$

How accurate is this estimate for small n?

2.3. Use the method of the preceding exercise, together with the Central Limit Theorem, to deduce the constant in Stirling's formula.

2.4. Prove directly that, if $0 \leq k < n$, then

$$\sum_m (-1)^{m-k} \binom{n}{m} \binom{m}{k} = \sum_m (-1)^{n-m} \binom{n}{m} \binom{m}{k} = 0.$$

2.5. Formulate and prove an analogue of Proposition 2.9 for binomial coefficients.

2.6. Let $B(n)$ be the number of partitions of $\{1,\ldots,n\}$. Prove that

$$\sqrt{n!} \leq B(n) \leq n!\,.$$

2.7. Prove that $\log n!$ is greater than $n\log n - n + 1$ and differs from it by at most $\frac{1}{2}\log n$. Deduce that

$$\frac{n^n}{\mathrm{e}^{n-1}} \leq n! \leq \frac{n^{n+1/2}}{\mathrm{e}^{n-1}}.$$

2.8. Let $c(n)$ be the number of connected permutations on $\{1,\ldots,n\}$. (A permutation π is *connected* if there does not exist k with $1 \leq k \leq n-1$ such that π maps $\{1,\ldots,k\}$ to itself.) Prove that

$$n! = \sum_{k=1}^{n} c(k)(n-k)!\ ,$$

and deduce that

$$\left(1+\sum_{n\geq 1} n!\,x^n\right)^{-1} = 1-\sum_{n\geq 1} c(n)x^n,$$

2.9. Prove that

$$(-1)^{n-k}\binom{n}{k} = \binom{-n+k-1}{k}$$

for $0 \leq k \leq n$. Use this and Proposition 2.3 to prove the Binomial Theorem for negative integer exponents.

2.10. Prove that

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

for $k \geq 1$. What happens when this equation is summed over k?

2.11. What is the relation between the numbers $T(n,k)$ defined in Exercise 1.4 and Stirling numbers?

2.12. A *total preorder* on a set X is a binary relation ρ on x which is symmetric and transitive and satisfies the condition that, for all $x,y \in X$, either $x\,\rho\,y$ or $y\,\rho\,x$ holds.

(a) Let ρ be a total preorder on X. Define a relation σ on X by the rule that $x\,\sigma\,y$ if and only if both $x\,\rho\,y$ and $y\,\rho\,x$ hold. Prove that σ is an equivalence relation whose equivalence classes are totally ordered by ρ. Show that ρ is determined by σ and the ordering of its equivalence classes. Show further that any equivalence relation and any total ordering of its equivalence classes aise in this way from a total preorder.

(b) Show that the number of total preorders of an n-set is

$$\sum_{k=1}^{n} S(n,k)k!\ .$$

(c) Show that the exponential generating function for the sequence in (b) is $1/(2-\exp(x))$.

(d) What can you deduce about the asymptotic behaviour of the sequence?

2.13. For $1 \le k \le n$, the *Lah number* $L(n,k)$ is defined by the formula

$$L(n,k) = \sum_{m=k}^{n} |s(n,m)|S(m,k).$$

(That is, the Lah numbers form a lower triangular matrix which is the product of the matrices of unsigned Stirling numbers of the first and second kinds. They are sometimes called Stirling numbers of the third kind.) Prove that

$$L(n,k) = \frac{n!}{k!}\binom{n-1}{k-1}.$$

2.14. Prove that

$$\begin{aligned}\sum_{k=0}^{n}\binom{n}{k}^2 &= \binom{2n}{n};\\ \sum_{k=0}^{n}(-1)^k\binom{n}{k}^2 &= \begin{cases} 0 & \text{if } n \text{ is odd;}\\ (-1)^{n/2}\binom{n}{n/2} & \text{if } n \text{ is even.}\end{cases}\end{aligned}$$

2.15. Prove that the generating function for the central binomial coefficients is

$$\sum_{n\ge 0}\binom{2n}{n}x^n = (1-4x)^{-1/2},$$

and deduce that

$$\sum_{k=1}^{n}\binom{2k}{k}\binom{2(n-k)}{n-k} = 4^n.$$

[Note: Finding a counting proof of this identity is quite challenging!]

2.16. Find a formula for the number $P(n;a_1,\ldots,a_k)$ appearing in Faà di Bruno's formula.

A prize question

In the course, a prize was offered for the first solution to this question.

> The following problem arises in the theory of clinical trials. A new drug is to be tested out. Of $2n$ subjects in the trial, n will receive the new drug and n will get a placebo. To avoid bias, it is important that the doctor administering the treatments does not know, and cannot reliably guess, which treatment each patient receives. The patients enter the trial one at a time, and are numbered from 1 to $2n$.
>
> If the treatments were allocated randomly with probability $1/2$, the doctor's guesses could be no better than random (so that the expected values for the numbers of correct and incorrect guesses are both n); but then the numbers of patients receiving drug and placebo would be unlikely to be equal. Given that they must balance, the doctor can certainly guess at least the last patient's treatment correctly.
>
> If we allocated the drug and the placebo randomly to patients $2i-1$ and $2i$ for $i = 1, \ldots, n$, then the doctor can correctly guess the treatment for each even-numbered patient.
>
> Suppose that instead we choose a random set of n patients to allocate the drug to, and the remaining n get the placebo; each of the $\binom{2n}{n}$ sets is equally likely. Suppose also that the doctor guesses according to the following rule. If the number of patients so far having the drug and the placebo are equal, he guesses at random about the next treatment. If the drug has occurred more often than the placebo, he guesses that the next treatment is the placebo, and *vice versa* if the placebo has occurred more often than the drug.
>
> Find a formula, and an asymptotic estimate, for the expected value of the difference between the number of correct guesses and the number of incorrect guesses that the doctor makes.

Solution We use the result of Problem 2.15 above, the identity

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

Following the hint, we first calculate the expected number of times during the trial when the numbers of patients receiving drug and placebo are equal. This is obtained by summing, over all n-element subsets A of $\{1,\ldots,2n\}$, the number of values of k for which $|A\cap\{1,\ldots,2k\}| = k$, and dividing by the number $\binom{2n}{n}$ of subsets. Now the sum can be calculated by counting, for each value of k, the number of n-subsets A for which $|A\cap\{1,\ldots,2k\}| = k$, and summing the result over k.

For a given k, the number of subsets is $\binom{2k}{k}\binom{2(n-k)}{n-k}$, since we must choose k of the numbers $1,\ldots,2k$, and $n-k$ of the numbers $2k+1,\ldots,2n$. Hence, by the stated result, the sum is 2^{2n}, and the average is $2^{2n}/\binom{2n}{n}$.

Now consider the doctor's guesses in any particular trial. At any stage where equally many patients have received drug and placebo, he guesses at random, and is equally likely to be right as wrong. Such points contribute zero to the expected number of correct minus incorrect guesses. In each interval between two consecutive such stages, say $2k$, and $2l$, the doctor will guess right one more time than he guesses wrong. (For example, if the $2k$th patient gets the drug, then between stages $2k+1$ and $2l$ the number of patients getting the drug is $l-k-1$ and the number getting the placebo is $l-k$, but the doctor will always guess the placebo.) So the expected number of correct minus incorrect guesses is the number of such intervals, which is one less than the number of times that the numbers are equal.

So the expected number is $2^{2n}/\binom{2n}{n}-1$, which is asymptotically $\sqrt{\pi n}$, by the result of Problem 2.2.

Chapter 3

Recurrence relations

A recurrence relation expresses the nth term of a sequence as a function of the preceding terms. The most general form of a recurrence relation takes the form

$$x_n = F_n(x_0, \ldots, x_{n-1}) \text{ for } n \geq 0.$$

Clearly such a recurrence has a unique solution. (Note that this allows the possibility of prescribing some initial values, by choosing the first few functions to be constant.)

Example: Ordered number partitions In how manny ways is it possible to write the positive integer n as a sum of positive integers, where the order of the summands is significant?

Let x_n be this number. One possible expression has a single summand n. In any other expression, if $n-i$ is the first summand, then it is followed by an expression for i as an ordered sum, of which there are x_i possibilities. Thus

$$x_n = 1 + x_1 + x_2 + \cdots + x_{n-1},$$

for $n \geq 1$. (When $n = 1$, this reduces to $x_1 = 1$.)

Since

$$x_{n-1} = 1 + x_1 + x_2 + \cdots + x_{n-2},$$

the recurrence reduces to the much simpler form

$$x_n = 2x_{n-1} \text{ for } n > 1,$$

with initial condition $x_1 = 1$. This obviously has the solution $x_n = 2^{n-1}$ for $n \geq 1$.

3.1 Linear recurrences with constant coefficients

Bounded recurrences

One type of linear recurrence which can be solved completely is of the form

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_k x_{n-k} \tag{3.1}$$

for $n \geq k$, where the k values $x_0, x_1, \ldots, x_{k-1}$ are prescribed.

If we consider the recurrence (3.1) without the initial values, we see that sums and scalar multiples of solutions are solutions. So, taking sequences over a field such as the rational numbers, we see that the set of solutions is a vector space over the field. Its dimension is k, since the k initial values can be prescribed abitrarily.

Thus, if we can write down k linearly independent solutions, the general solution is a linear combination of them.

The *characteristic equation* of the recurrence (3.1) is the equation

$$x^k - a_1 x^{k-1} - \cdots - a_k = 0.$$

This polynomial has k roots, some of which may be repeated. Suppose that its distinct roots are $\alpha_1, \ldots, \alpha^r$ with multiplicities $m_1, \ldots, m_r$, where $m_1 + \cdots + m_r = k$. Then a short calculation shows that the k functions

$$x_n = \alpha_1^n, \ldots, n^{m_1 - 1}\alpha_1^n, \ldots, \alpha_r^n, \ldots, n^{m_r - 1}\alpha_r^n$$

are solutions of (3.1); they are clearly linearly independent. So the general solution is a linear combination of them.

Example: Fibonacci numbers Consider the Fibonacci recurrence

$$F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 2.$$

The characteristic equation is

$$x^2 - x - 1 = 0$$

with roots $\alpha, \beta = (1 \pm \sqrt{5})/2$. So the general solution is

$$F_n = A\alpha^n + B\beta^n,$$

and A and B can be determined from the initial conditions.

For the usual Fibonacci numbers, we have $F_0 = F_1 = 1$, giving the two equations

$$\begin{aligned} A+B &= 1, \\ A\alpha+B\beta &= 1. \end{aligned}$$

Solving these equations gives the solution we found earlier.

Example: Sequences with forbidden subwords Let a be a binary sequence of length k. How many binary sequences of length n do not contain a as a consecutive subword?

Suppose, for example, that $a = 11$, so that we are counting binary strings with no two consecutive ones. Let $f(n)$ denote the number of such sequences of length n, and let $g(n)$ the number of sequences commencing with 11 but having no other occurrence of 11. Then

$$2f(n) = f(n+1)+g(n+1),$$

since if we take a string with no occurrence of 11 and precede it with a 1, then the only possible position of 11 is at the beginning. Also, if we take a string with no occurrence of 11 and precede it with 11, then the resulting sequence contains 11, but possibly two occurrences (if the original string began with a 1); so we have

$$f(n) = g(n+1)+g(n+2).$$

Then $f(n) = (2f(n)-f(n+1))+(2f(n+1)-f(n+2))$, so we have the Fibonacci recurrence

$$f(n+2) = f(n)+f(n+1).$$

Since $f(0) = 1 = F_1$ and $f(1) = 2 = F_2$, a simple induction proves that $f(n) = F_{n+1}$ for all $n \geq 0$.

Guibas and Odlyzko extended this approach to arbitrary forbidden substrings. They defined the *correlation polynomial* of a binary string a of length k to be

$$C_a(x) = \sum_{j=0}^{k-1} c_a(j)x^j,$$

where $c_a(0) = 1$ and, for $1 \leq j \leq k-1$,

$$c_a(j) = \begin{cases} 1 & \text{if } a_1a_2\cdots a_{k-j} = a_{j+1}a_{j+2}\cdots a_k, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, for $a = 11$, we have $C_a(x) = 1+x$.

Theorem 3.1 *Let $f_a(n)$ be the number of binary strings of length n excluding the substring a of length k. Then the generating function $F_a(x) = \sum_{n\geq 0} f_a(n)x^n$ is given by*

$$F_a(x) = \frac{C_a(x)}{x^k + (1-2x)C_a(x)},$$

where $C_a(x)$ is the correlation polynomial of a.

Proof We define $g_a(n)$ to be the number of binary sequences of length n which commence with a but have no other occurrence of a as a consecutive subsequence, and $G_a(x) = \sum_{n\geq 0} g_a(n)x^n$ the generating function of this sequence of numbers.

Let b be a sequence counted by $f_a(n)$. Then for $x \in \{0,1\}$, the sequence xb contains a at most once at the beginning. So

$$2f_a(n) = f_a(n+1) + g_a(n+1).$$

Multiplying by x^n and summing over $n \geq 0$ gives

$$2F_a(x) = x^{-1}(F_a(x) - 1 + G_a(x)). \tag{3.2}$$

Now let c be the concatenation ab. Then c starts with a, and may contain other occurrences of a, but only at positions overlapping the initial a, that is, where $a_{k-j+1}\cdots a_k b_1 \cdots b_j = a_1 \cdots a_k$. This can only occur when $c_a(k-j) = 1$, and the sequence $a_{k-j+1}\cdots a_k b$ then has length $n+j$ and has a unique occurrence of a at the beginning. So

$$f_a(n) = \sum g_a(n+j),$$

where the sum is over all j with $1 \leq j \leq k$ for which $c_a(k-j) = 1$. This can be rewritten

$$f_a(n) = \sum_{j=1}^{k} c_a(k-j)g_a(n+j),$$

or in terms of generating functions,

$$F_a(x) = x^{-k}C_a(x)G_a(x). \tag{3.3}$$

Combining equations (3.2) and (3.3) gives the result.

In the case where $a = 11$, we obtain

$$F_{11}(x) = \frac{1+x}{x^2 + (1-2x)(1+x)} = \frac{1+x}{1-x-x^2},$$

so that $f_{11}(n) = F_n + F_{n-1} = F_{n+1}$, as previously noted.

Unbounded recurrences

We will give here just one example. Recall from the last chapter that the generating function for the number $p(n)$ of partitions of the integer n is given by

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

Thus, to get a recurrence relation for $p(n)$, we have to understand the coefficients of its inverse:

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

Now a term on the right arises from each expression for n as a sum of distinct positive integers; its value is $(-1)^k$, where k is the number of terms in the sum. Thus, $a(n)$ is equal to the number of expressions for n as the sum of an even number of distinct parts, minus the number of expressions for n as the sum of an odd number of distinct parts.

This number is evaluated by *Euler's Pentagonal Numbers Theorem*:

Proposition 3.2

$$a(n) = \begin{cases} (-1)^k & \textit{if } n = k(3k-1)/2 \textit{ for some } k \in \mathbb{Z}, \\ 0 & \textit{otherwise.} \end{cases}$$

Putting all this together, the recurrence relation for $p(n)$ is

$$\begin{aligned} p(n) &= \sum_{k\neq 0}(-1)^{k-1}p(n-k(3k-1)/2) \\ &= p(n-1)+p(n-2)-p(n-5)-p(n-7)+p(n-12)+\cdots \end{aligned}$$

where the summation is over all values of k for which $n-k(3k-1)/2$ is non-negative.

The number of terms in the recurrence grows with n, but only as $O(\sqrt{n})$. So evaluating $p(n)$ for $n \leq N$ requires only $O(n^{3/2})$ additions and subtractions.

3.2 Other recurrence relations

There is no recipe for solving more general recurrence relations. We do a few examples for illustration.

Example: derangements Let $d(n)$ be the number of derangements of $\{1,\ldots,n\}$ (permutations which have no fixed points). We obtain a recurrence relation as follows. Each derangement maps n to some i with $1 \leq i \leq n-1$, and by symmetry each i occurs equally often. So we need only count the derangements mapping n to $n-1$, and multiply by $n-1$.

We divide these derangements into two classes. The first type map $n-1$ back to n. Such a permutation must be a derangement of $\{1,\ldots,n-2\}$ composed with the transposition $(n-1,n)$; so there are $d(n-2)$ such. The second type map i to n for some $i \neq n-1$. Replacing the sequence $i \mapsto n \mapsto n-1$ by the sequence $i \mapsto n-1$, we obtain a derangement of $n-1$; every such derangement arises. So there are $d(n-1)$ deraggements of this type.

Thus,

$$d(n) = (n-1)(d(n-1)+d(n-2)).$$

There is a simpler recurrence satisfied by $d(n)$, which can be deduced from this one, namely

$$d(n) = nd(n-1)+(-1)^n.$$

To prove this by induction, suppose that it is true for $n-1$. Then $(n-1)d(n-2) = d(n-1)-(-1)^{n-1}$; so $d(n) = (n-1)d(n-1)+d(n-1)+(-1)^n$, and the inductive step is proved. (Starting the induction is an exercise.)

Now this is a special case of a general recursion which can be solved, namely

$$x_0 = c, \qquad x_n = p_n x_{n-1} + q_n \text{ for } n \geq 1.$$

We can include the initial condition in the recursion by setting $q_0 = c$ and adoopting the convention that $x_{-1} = 0$.

If $q_n = 0$ for $n \geq 1$, then the solution is simply $x_n = P_n$ for all n, where

$$P_n = c\prod_{i=1}^{n} p_i.$$

So we compare x_n to p_n. Putting $y_n = x_n/P_n$, the recurrence becomes

$$y_0 = 1, \qquad y_n = y_{n-1} + \frac{q_n}{P_n} \text{ for } n \geq 1,$$

with solution

$$y_n = \sum_{i=0}^{n} \frac{q_i}{P_i}.$$

(Remember that $q_0 = P_0 = c$.) Finally,

$$x_n = P_n \sum_{i=0}^{n} \frac{q_i}{P_i}.$$

For derangements, we have $p_n = n$, $c = 1$ (so that $P_n = n!$), and $q_n = (-1)^n$. Thus

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

It follows that $d(n)$ is the nearest integer to $n!/\mathrm{e}$, since

$$n!/\mathrm{e} - d(n) = n! \sum_{i \geq n+1} \frac{(-1)^i}{i!},$$

and the modulus of the alternating sum of decreasing terms on the right is smaller than that of the first term, which is $n!/(n+1)! = 1/(n+1)$.

Example: Catalan numbers It is sometimes possible to use a recurrence relation to derive an algebraic or differential equation for a generating function for the sequence. If we are lucky, this equation can be solved, and the resulting function used to find the terms in the sequence.

The nth *Catalan number* C_n is the number of ways of bracketing a product of n terms, where we are not allowed to assume that the operation is associatuve or commutative. For example, for $n = 4$, there are five bracketings

$$(a(b(cd))), (a((bc)d)), ((ab)(cd)), ((a(bc))d), (((ab)c)d),$$

so $C_4 = 5$.

Any bracketed product of n terms is of the form (AB), where A and B are bracketed products of i and $n - i$ terms respectively. So

$$C_n = \sum_{i=1}^{n-1} C_i C_{n-i} \text{ for } n \geq 2.$$

Putting $F(x) = \sum_{n \geq 1} C_n x^n$, the recurrence relation shows that F and F^2 agree in all coefficients except $n = 1$. Since $C_1 = 1$ we have $F = F^2 + x$, or $F^2 - F + x = 0$. Solving this equation gives

$$F(x) = \tfrac{1}{2}(1 \pm \sqrt{1-4x}).$$

Since $C_0 = 0$ by definition, we must take the negative sign here.

This expression gives us a rough estimate for C_n: the nearest singularity to the origin is a branchpoint at $1/4$, so C_n grows "like" 4^n. However, we can get the solution explicitly.

From the binomial theorem, we have

$$F(x) = \tfrac{1}{2}\left(1 - \sum_{n\geq 0}\binom{1/2}{n}(-4)^n\right).$$

Hence

$$\begin{aligned}
C_n &= -\frac{1}{2}\binom{1/2}{n}(-4)^n \\
&= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{2}\cdots\frac{2n-3}{2}\cdot\frac{2^{2n}}{n!} \\
&= \frac{1}{2^{n+1}}\cdot\frac{(2n-2)!}{2^{n-1}(n-1)!}\cdot\frac{2^{2n}}{n\cdot(n-1)!} \\
&= \frac{1}{n}\binom{2n-2}{n-1}.
\end{aligned}$$

Sometimes we cannot get an explicit solution, but can obtain some information about the growth rate of the sequence.

Example: Wedderburn–Etherington numbers Another interpretation of the Catalan number C_n is the number of rooted binary trees with n leaves, where "left" and "right" are distinguished. If we do not distinguish left and right, we obtain the *Wedderburn–Etherington numbers* W_n.

Such a tree is determined by the choice of trees with i and $n-i$ leaves, but the order of the choice is unimportant. Thus, if $i = n/2$, the number of trees is only $W_i(W_i+1)/2$, rather than W_i^2. For $i \neq n/2$, we simply halve the number. This gives the recurrence

$$W_n = \begin{cases} \tfrac{1}{2}\displaystyle\sum_{i=1}^{n-1} W_i W_{n-i} & \text{if } n \text{ is odd,} \\ \tfrac{1}{2}\left(\displaystyle\sum_{i=1}^{n-1} W_i W_{n-i} + W_{n/2}\right) & \text{if } n \text{ is even.} \end{cases}$$

Thus, $F(x) = \sum W_n x^n$ satisfies

$$F(x) = x + \tfrac{1}{2}(F(x)^2 + F(x^2)).$$

This cannot be solved explicitly. We will obtain a rough estimate for the rate of growth. Later, we find more precise asymptotics.

We seek the nearest singularity to the origin. Since all coefficients are real and positive, this will be on the positive real axis. (If a power series with positive real coefficiets converges at $z = r$, then it converges absolutely at any z with $|z| = r$.) Let s be the required point. Then $s < 1$, so $s^2 < s$; so $F(z^2)$ is analytic at $z = s$. Now write the equation as

$$F(z)^2 - 2F(z) + (F(z^2) + 2z) = 0,$$

with "solution"

$$F(z) = 1 - \sqrt{1 - 2z - F(z^2)}$$

(taking the negative sign as before). Thus, s is the real positive solution of

$$F(s^2) = 1 - 2s.$$

Solving this equation numerically (using the fact that $F(s^2)$ is the sum of a convergent Taylor series and can be estimated from knowledge of a finite number of terms), we find that $s \approx 0.403\ldots$, so that W_n grows "like" $(2.483\ldots)^n$.

We will find more precise asymptotics for W_n later in the course.

Example: Bell numbers We already calculated the exponential generating function for the Bell numbers. Here is how to do it using the recurrence relation

$$B(n) = \sum_{k=1}^{n} \binom{n-1}{k-1} B(n-k).$$

Multiply by $x^n/n!$ and sum over n: the e.g.f $F(x)$ is given by

$$F(x) = \sum_{n\ge 0} \frac{x^n}{n!} \sum_{k=1}^{n} \binom{n-1}{k-1} B(n-k).$$

Differentiating with respect to x we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x} F(x) &= \sum_{n\ge 1} \frac{x^{n-1}}{(n-1)!} \sum_{k=1}^{n} \binom{n-1}{k-1} B(n-k) \\ &= \sum_{l\ge 0} \frac{x^l}{l!} \sum_{m\ge 0} \frac{B(m)x^m}{m!}. \end{aligned}$$

Here we use new variables $l = k-1$ and $m = n-k$; the constraints of the original sum mean that l and m independently take all natural number values. Hence

$$\frac{\mathrm{d}}{\mathrm{d}x}F(x) = \exp(x)F(x).$$

This first-order differential equation can be solved in the usual way with the initial condition $F(0) = 1$ to give

$$F(x) = \exp(\exp(x)-1),$$

in agreement with our earlier result.

Exercises

3.1. Some questions on Fibonacci numbers.

(a) Show that the number of expressions for n as an ordered sum of ones and twos is F_n.

(b) Verify the following formula for the sloping diagonals of Pascal's triangle:

$$\sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n-i}{i} = F_n.$$

(c) Let n be a positive integer. Write down all expressions for n as an ordered sum of positive integers. For each such expression, multiply the summands together; then add the resulting products. Prove that the answer is F_{2n-1}.

(d) In (c), if instead of multiplying the summands, we multiply 2^{d-2} for each summand $d > 2$, then the answer is F_{2n-2}.

(e) Prove that, for $n \geq 0$,

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{n+2} = \begin{pmatrix} F_n & F_{n+1} \\ F_{n+1} & F_{n+2} \end{pmatrix}.$$

(f) Use (e) to show that F_n can be computed with $O(\log n)$ arithmetic operations on integers.

3.2. Let A be a finite set of positive integers. Suppose that the currency of a certain country has A as the set of denominations. Prove that the number $f(n)$ of ways of paying a bill of n units, where coins are paid in order, has generating function $1/(1-\sum_{a\in A}x^a)$.

Suppose that $A = \{1,2,5,10\}$. Prove that $f(n) \sim c\,\alpha^n$ for some constants c and α, and estimate α.

What is the generating function for the number in the case when the order of the coins is not significant?

3.3. Let a be a binary string of length k with correlation polynomial $C_a(x)$. A random binary sequence is obtained by tossing a fair coin, recording 1 for heads and 0 for tails. Let E_a be the expected number of coin tosses until the first occurrence of a as a consecutive substring. Prove that E_a is the sum, over n, of the probability that a doesn't occur in the first n terms of the sequence. Deduce that $E_a = 2^k\,C_a(1/2)$.

3.4. This exercise is due to Wilf, and illustrates his "snake oil" method.

(a) Prove that

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

(b) Let

$$a_n = \sum_{k=0}^{n}\binom{n+k}{2k}2^{n-k}$$

for $n \geq 0$. Prove that the ordinary generating function for (a_n) is

$$\sum_{n\geq 0}a_n x^n = \frac{1-2x}{(1-x)(1-4x)},$$

and deduce that $a_n = (2^{2n+1}+1)/3$ for $n \geq 0$.

(c) Write down a linear recurrence relation with constant coefficients satisfied by the numbers a_n.

3.5. Let s_n be the number of partitions of an n-set into parts of size 1 or 2 (equivalently, the number of permutations of an n-set whose square is the identity). Show that

$$s_n = s_{n-1} + (n-1)s_{n-2} \text{ for } n \geq 2,$$

and hence find the exponential generating function for (s_n) in closed form.

3.6. Let a_n be the number of strings that can be formed from n distinct letters (using each letter at most once, and including the empty string). Prove that

$$a_0 = 1, \qquad a_n = na_{n-1} + 1 \text{ for } n \geq 1,$$

and deduce that $a_n = \lfloor \mathrm{e}\, n! \rfloor$. What is the exponential generating function for this sequence?

3.7. Prove that

$$\frac{x^{m+1} - y^{m+1}}{x - y} = \sum_{k=0}^{\lfloor m/2 \rfloor} \binom{m-k}{k} (-xy)^k (x+y)^{m-2k}.$$

By taking x and y to be the roots of the equation $z^2 - z - 1 = 0$, deduce the equality of two well-known expressions for the Fibonacci numbers.

(I am grateful to Marcio Soares for this exercise.)

Chapter 4

q-analogues

Much of the enumerative combinatorics of sets and functions can be generalised in a manner which, at first sight, seems a bit unmotivated. In this chapter, we develop a small amount of this large body of theory.

4.1 Motivation

We can look at q-analogues in several ways:

- The q-analogues are, typically, formulae which tend to the classical ones as $q \to 1$. Most basic is the fact that

  $$\lim_{q\to 1} \frac{q^a - 1}{q - 1} = a$$

  for any real number a (this is immediate from l'Hôpital's rule).

- There is a formal similarity between statements about subsets of a set and subspaces of a vector space, with cardinality replaced by dimension. For example, the inclusion-exclusion rule

  $$|U \cup V| + |U \cap V| = |U| + |V|$$

  for sets becomes

  $$\dim(U+V) + \dim(U \cap V) = \dim(U) + \dim(V)$$

  for vector spaces. Now, if the underlying field has q elements, then the number of 1-dimensional subspaces of an n-dimensional vector space is $(q^n - 1)/(q - 1)$, which is exactly the q-analogue of n.

- The analogy can be interpreted at a much higher level, in the language of *braided categories*. I will not pursue this here. You can read more in various papers of Shahn Majid, for example Braided Groups, *J. Pure Appl. Algebra* **86** (1993), 187–221; Free braided differential calculus, braided binomial theorem and the braided exponential map, *J. Math. Phys.* **34** (1993), 4843–4856.

In connection with the second interpretation, note the theorem of Galois:

Theorem 4.1 *The cardinality of any finite field is a prime power. Moreover, for any prime power q, there is a unique field with q elements, up to isomorphism.*

To commemorate Galois, finite fields are called *Galois fields*, and the field with q elements is denoted by $\mathrm{GF}(q)$.

Definition The *Gaussian coefficient*, or *q-binomial coefficient*, $\begin{bmatrix} n \\ k \end{bmatrix}_q$, where n and k are natural numbers and q a real number different from 1, is defined by

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n-1)(q^{n-1}-1)\cdots(q^{n-k+1}-1)}{(q^k-1)(q^{k-1}-1)\cdots(q-1)}.$$

It can be shown that this expression is a polynomial in q, if we regard q as an indeterminate. If instead we regard q as a complex number, it has a well-defined value as long as q is not a dth root of unity for some d dividing k. (In the excluded cases, the denominator is zero, but the limit still exists.)

Proposition 4.2 *(a)* $\lim_{q\to 1} \begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{pmatrix} n \\ k \end{pmatrix}.$

(b) If q is a prime power, then the number of k-dimensional subspaces of an n-dimensional vector space over $\mathrm{GF}(q)$ *is equal to* $\begin{bmatrix} n \\ k \end{bmatrix}_q$.

Proof The first assertion is almost immediate from $\lim_{q\to 1}(q^n-1)/(q-1) = n$.

For the second, note that the number of choices of k linearly independent vectors in $\mathrm{GF}(q)^n$ is

$$(q^n-1)(q^n-q)\cdots(q^n-q^{k-1}),$$

since the ith vector must be chosen outside the span of its predecessors. Any such choice is the basis of a unique k-dimensional subspace. Putting $n = k$, we see that the number of bases of a k-dimensional space is

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

Dividing and cancelling powers of q gives the result.

4.2 The q-binomial theorem

The q-binomial coefficients satisfy an analogue of the recurrence relation for binomial coefficients.

Proposition 4.3 $\begin{bmatrix} n \\ 0 \end{bmatrix}_q = \begin{bmatrix} n \\ n \end{bmatrix}_q = 1, \qquad \begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q + q^k \begin{bmatrix} n-1 \\ k \end{bmatrix}_q$ *for* $0 < k < n$.

Proof This comes straight from the definition. Suppose that $0 < k < n$. Then

$$\begin{aligned} \begin{bmatrix} n \\ k \end{bmatrix}_q - \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q &= \left(\frac{q^n - 1}{q^k - 1} - 1\right) \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \\ &= q^k \left(\frac{q^{n-k} - 1}{q^k - 1}\right) \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \\ &= q^k \begin{bmatrix} n \\ k-1 \end{bmatrix}_q. \end{aligned}$$

The array of Gaussian coefficients has the same symmetry as that of binomial coefficients. From this we can deduce another recurrence relation.

Proposition 4.4 *(a) For* $0 \le k \le n$,

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n \\ n-k \end{bmatrix}_q.$$

(b) For $0 < k < n$,

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k \end{bmatrix}_q.$$

Proof (a) is immediate from the definition. For (b),

$$\begin{aligned}
\begin{bmatrix} n \\ k \end{bmatrix}_q &= \begin{bmatrix} n \\ n-k \end{bmatrix}_q \\
&= \begin{bmatrix} n-1 \\ n-k-1 \end{bmatrix}_q + q^{n-k} \begin{bmatrix} n-1 \\ n-k \end{bmatrix}_q \\
&= \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q .
\end{aligned}$$

We come now to the q-analogue of the binomial theorem, which states the following.

Theorem 4.5 *For a positive integer n, a real number $q \neq 1$, and an indeterminate z, we have*

$$\prod_{i=1}^{n} (1 + q^{i-1} z) = \sum_{k=0}^{n} q^{k(k-1)/2} z^k \begin{bmatrix} n \\ k \end{bmatrix}_q .$$

Proof The proof is by induction on n; starting the induction at $n = 1$ is trivial. Suppose that the result is true for $n - 1$. For the inductive step, we must compute

$$\left(\sum_{k=0}^{n-1} q^{k(k-1)/2} z^k \begin{bmatrix} n-1 \\ k \end{bmatrix}_q \right) \left(1 + q^{n-1} z\right) .$$

The coefficient of z^k in this expression is

$$\begin{aligned}
& q^{k(k-1)/2} \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{(k-1)(k-2)/2+n-1} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \\
= \;& q^{k(k-1)/2} \left(\begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \right) \\
= \;& q^{k(k-1)/2} \begin{bmatrix} n \\ k \end{bmatrix}_q
\end{aligned}$$

by Proposition 4.4(b).

4.3 Elementary symmetric functions

In this section we touch briefly on the theory of elementary symmetric functions.

Let $x_1,\ldots,x_n$ be n indeterminates. For $1 \le k \le n$, the kth *elementary symmetric function* $e_k(x_1,\ldots,x_n)$ is the sum of all monomials which can be formed by multiplying together k *distinct* indeterminates. Thus, e_k has $\binom{n}{k}$ terms, and

$$e_k(1,1,\ldots,1) = \binom{n}{k}.$$

For example, if $n = 3$, the elementary symmetric functions are

$$e_1 = x_1 + x_2 + x_3, \quad e_2 = x_1x_2 + x_2x_3 + x_3x_1, \quad e_3 = x_1x_2x_3.$$

We adopt the convention that $e_0 = 1$.

Newton observed that the coefficients of a polynomial of degree n are the elementary symmetric functions of its roots, with appropriate signs:

Proposition 4.6 $\displaystyle\prod_{i=1}^{n}(z - x_i) = \sum_{k=0}^{n}(-1)^k e_k(x_1,\ldots,x_n)z^{n-k}.$

Consider the generating function for the e_k:

$$E(z) = \sum_{k=0}^{n} e_k(x_1,\ldots,x_n)z^k.$$

A slight rewriting of Newton's Theorem shows that

$$E(z) = \prod_{i=1}^{n}(1 + x_iz).$$

Hence the binomial theorem and its q-analogue give the following specialisations:

Proposition 4.7 *(a) If $x_1 = \ldots = x_n = 1$, then*

$$E(z) = (1+z)^n = \sum_{k=0}^{n}\binom{n}{k}z^k,$$

so

$$e_k(1,1,\ldots,1) = \binom{n}{k}.$$

(b) If $x_i = q^{i-1}$ for $i = 1,\ldots,n$, then

$$E(z) = \prod_{i=1}^{n}(1+q^{i-1}z) = \sum_{k=0}^{n} q^{k(k-1)/2} z^k \begin{bmatrix} n \\ k \end{bmatrix}_q,$$

so

$$e_k(1,q,\ldots,q^{n-1}) = q^{k(k-1)/2} \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

4.4 Partitions and permutations

The number of permutations of an n-set is $n!$. The linear analogue of this is the number of linear isomorphisms from an n-dimensional vector space to itself; this is equal to the number of choices of basis for the n-dimensional space, which is

$$(q^n-1)(q^n-q)\cdots(q^n-q^{n-1}).$$

These linear maps form a group, the *general linear group* $\mathrm{GL}(n,q)$.

Using the q-binomial theorem, we can transform this multiplicative formula into an additive formula:

Proposition 4.8

$$|\,\mathrm{GL}(n,q)| = (-1)^n q^{n(n-1)/2} \sum_{i=0}^{n} (-1)^k q^{k(k+1)/2} \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

Proof We have

$$|\,\mathrm{GL}(n,q)| = (-1)^n q^{n(n-1)/2} \prod_{i=1}^{n}(1-q^i),$$

and the right-hand side is obtained by substituting $z = -q$ in the q-binomial theorem.

The total number of $n \times n$ matrices is q^{n^2}, so the probability that a random matrix is invertible is

$$p_n(q) = \prod_{i=1}^{n}(1-q^{-i}).$$

As $n \to \infty$, we have

$$p_n(q) \to p(q) = \prod_{i\geq 1}(1-q^{-i}).$$

According to Euler's Pentagonal Numbers Theorem, we have

$$p(q) = \sum_{k\in\mathbb{Z}} (-1)^k q^{-k(3k-1)/2} = 1 - q^{-1} - q^{-2} + q^{-5} + q^{-7} - q^{-12} - \cdots$$

So, for example, $p(2) = 0.2887\ldots$ is the limiting probability that a large random matrix over $\mathrm{GF}(2)$ is invertible.

What is the q-analogue of the Stirling number $S(n,k)$, the number of partitions of an n-set into k parts? This is a philosophical, not a mathematical question; I argue that the q-analogue is the Gaussian coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_q$.

The number of surjective maps from an n-set to a k-set is $k!S(n,k)$, since the preimages of the points in the k-set form a partition of the n-set whose k parts can be mapped to the k-set in any order. The q-analogue is the number of surjective linear maps from an n-space V to a k-space W. Such a map is determined by its kernel U, an $(n-k)$-dimensional subspace of V, and a linear isomorphism from V/U to W. So the analogue of $S(n,k)$ is the number of choices of U, which is

$$\begin{bmatrix} n \\ n-k \end{bmatrix}_q = \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

4.5 Irreducible polynomials

Though it is not really a q-analogue of a classical result, the following theorem comes up in various places. Recall that a polynomial of degree n is *monic* if the coefficient of x^n is equal to 1.

Theorem 4.9 *The number $f_q(n)$ of monic irreducible polynomials of degree n over* $\mathrm{GF}(q)$ *satisfies*

$$\sum_{k\mid n} k f_q(k) = q^n.$$

Proof We give two proofs, one depending on some algebra, and the other a rather nice exercise in manipulating formal power series.

First proof: We use the fact that the roots of an irreducible polynomial of degree k over $\mathrm{GF}(q)$ lie in the unique field $\mathrm{GF}(q^k)$ of degree k over $\mathrm{GF}(q)$. Moreover, $\mathrm{GF}(q^k) \subseteq \mathrm{GF}(q^n)$ if and only if $k \mid n$; and every element of $\mathrm{GF}(q^n)$ generates some subfield over $\mathrm{GF}(q)$, which has the form $\mathrm{GF}(q^k)$ for some k dividing n.

Now each of the q^n elements of GF$(q^n$ satisfies a unique minimal polynomial. of degree k for some k; and every irreducible polynomial arises in this way, and has k distinct roots. So the result holds.

Second proof: All the algebra we use in this proof is that each monic polynomial of degree n can be factorised uniquely into monic irreducible factors. If the number of monic irreducibles of degree k is m_k, then we obtain all monic polynomials of degree n by the following procedure:

- Express $n = \sum a_k k$, where a_k are non-negative integers;
- Choose a_k monic irreducibles of degree k from the set of all m_k such, with repetitions allowed and order not important;
- Multiply the chosen polynomials together.

Altogether there are q^n monic polynomials $x^n + c_1 x^{n-1} + \cdots + c_n$ of degree n, since there are q choices for each of the n coefficients. Hence

$$q^n = \sum \prod_k \binom{m_k + a_k - 1}{a_k}, \tag{4.1}$$

where the sum is over all sequences $a_1, a_2, \ldots$ of natural numbers which satisfy $\sum k a_k = n$.

Multiplying by x^n and summing over n, we get

$$\begin{aligned}
\frac{1}{1-qx} &= \sum_{n\geq 0} q^n x^n \\
&= \sum_{a_1,a_2,\ldots} \prod_{k\geq 1} \binom{m_k + a_k - 1}{a_k} x^{k a_k} \\
&= \prod_{k\geq 1} \sum_{a\geq 0} \binom{m_k + a - 1}{a} (x^k)^a \\
&= \prod_{k\geq 1} (1 - x^k)^{-m_k}.
\end{aligned}$$

Here the manipulations are similar to those for the sum of cycle indices in Chapter 2; we use the fact that the number of choices of a things from a set of m, with repetition allowed and order unimportant, is $\binom{m+a-1}{a}$, and in the fourth line we invoke the Binomial Theorem with negative exponent.

Taking logarithms of both sides, we obtain

$$\begin{aligned}\sum_{n\geq 1}\frac{q^nx^n}{n} &= -\log(1-qx)\\ &= \sum_{k\geq 1} -m_k\log(1-x^k)\\ &= \sum_{k\geq 1} m_k \sum_{r\geq 1}\frac{x^{kr}}{r}.\end{aligned}$$

The coefficient of x^n in the last expression is the sum, over all divisors k of n, of $m_k/r = km_k/n$. This must be equal to the coefficient on the left, which is q^n/n. We conclude that

$$q^n = \sum_{k|n} km_k, \tag{4.2}$$

as required.

Note how the very complicated recurrence relation (4.1) for the numbers m_k changes into the much simpler recurrence relation (4.2) after taking logarithms!

We will see how to solve such a recurrence in the chapter on Möbius inversion.

Exercises

4.1. Prove that $\begin{bmatrix} n \\ k \end{bmatrix}_q$ is a polynomial of degree $k(n-k)$ in the indeterminate q.

4.2.

(a) Prove that, for $0 < k < n$,

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + (q^{n-1}-1)\begin{bmatrix} n-2 \\ k-1 \end{bmatrix}_q.$$

(b) Let

$$F_q(n) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q,$$

so that, if q is a prime power, then $F_q(n)$ is the total number of subspaces of an n-dimensional vector space over $\mathrm{GF}(q)$. Prove that

$F_q(0) = 1, F_q(1) = 2, \quad F_q(n) = 2F_q(n-1) + (q^{n-1}-1)F_q(n-2)$ for $n \geq 2$.

(c) Deduce that, if $q > 1$, then $F_q(n) \geq c\, q^{n^2/4}$ for some constant c (depending on q).

4.3. This exercise shows that the Gaussian coefficients have a counting interpretation for all positive integer values of q (not just prime powers).

Suppose that q is an integer greater than 1. Let Q be a finite set of cardinality q containing two distinguished elements 0 and 1. We say that a $k \times n$ matrix with entries from Q is in *reduced echelon form* if the following conditions hold:

- If a row has any non-zero entries, then the first such entry is 1 (such entries are called “leading 1”);
- if $i < j$ and row j is non-zero, then row i is also non-zero, and its leading 1 occurs to the left of the leading 1 in row j;
- if a column contains the leading 1 of some row, then all other entries in that column are 0.

Prove that $\begin{bmatrix} n \\ k \end{bmatrix}_q$ is the number of $k \times n$ matrices in reduced echelon form with no rows of zeros.

4.4. A matrix is said to be in *echelon form* if it satisfies the first two conditions in the definition of reduced echelon form. Show that, if q is an integer greater than 2, the right-hand side of the q-binomial theorem with $x = 1$ counts the number of $n \times n$ matrices in echelon form.

How many $n \times n$ matrices in reduced echelon form are there?

4.5. Let $h_k(x_1, \dots, x_n)$ be the *complete symmetric function* of degree k in the indeterminates $x_1, \dots, x_n$ (the sum of *all* monomials of degree k that can be formed using these indeterminates). For example,

$$h_2(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 + x_3x_1.$$

Prove that

$$\sum_{k=0}^{\infty} h_k(x_1, \dots, x_n) z^k = \prod_{i=1}^{n} (1 - x_i z)^{-1}.$$

Deduce that

(a) $h_k(1, 1, \dots, 1) = \binom{n+k-1}{k}$;

(b) $h_k(1,q,\ldots,q^{n-1}) = \begin{bmatrix} n+k-1 \\ k \end{bmatrix}_q$ for $q \neq 1$.

Hint for (b): show that

$$\sum_{i=0}^{k} q^i \begin{bmatrix} n+i-2 \\ i \end{bmatrix}_q = \begin{bmatrix} n+k-1 \\ k \end{bmatrix}_q.$$

4.6. The second proof of Theorem 4.9 shows that the number of irreducible polynomials over $\mathrm{GF}(q)$ is exactly what is required if every element of $\mathrm{GF}(q^n)$ is the root of a unique irreducible of degree dividing n over $\mathrm{GF}(q)$. Turn the argument around to gove a counting proof of the existence and uniqueness of $\mathrm{GF}(q^n)$, given that of $\mathrm{GF}(q)$.

4.7. Let ω be a primitive dth root of unity. Express $\begin{bmatrix} n \\ k \end{bmatrix}_\omega$ in terms of binomial coefficients (whenever you can).

Solution by Pablo Spiga Let d be a natural number, and let ω be a primitive dth root of unity in $\mathbb{C}$, i.e. $\omega^d = 1$. Then, if $0 \leq a,b \leq d-1$, we have

$$\begin{bmatrix} nd+a \\ kd+b \end{bmatrix}_\omega = \binom{n}{k} \begin{bmatrix} a \\ b \end{bmatrix}_\omega.$$

Note that we are assuming that $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]_\omega = 0$ whenever $a < b$.

Solution By induction on a. We have

$$\begin{aligned} 1-\xi^d &= \prod_{i=1}^{d} (\omega^{i-1} - \xi) \\ &= \prod_{i=1}^{d} (\omega^{i-1} \cdot (1-\omega^{-i+1}\xi)) \\ &= \prod_{i=1}^{d} \omega^{i-1} \cdot \prod_{i=1}^{d} (1-\omega^{i-1}\xi). \end{aligned}$$

Thus, we get

$$\prod_{i=1}^{nd} (1+\omega^{i-1}(-\xi)) = \sum_{j=0}^{nd} \omega^{j(j-1)/2} (-1)^j \begin{bmatrix} nd \\ j \end{bmatrix}_\omega \xi^j, \tag{4.3}$$

but

$$\prod_{i=1}^{nd}(1+\omega^{i-1}(-\xi)) = (1-\xi^d)^n = \sum_{k=0}^{n}\binom{n}{k}(-1)^k\xi^{kd}. \tag{4.4}$$

We have proved that $\begin{bmatrix} nd \\ j \end{bmatrix}_\omega = 0$ if d does not divide j. Assume $j = dk$. By (4.3) and (4.4), as

$$\omega^{dk(dk-1)/2}(-1)^{k(d+1)} = 1, \tag{4.5}$$

we get

$$\begin{bmatrix} nd \\ kd \end{bmatrix}_\omega = \binom{n}{k}.$$

(For (4.5), note that if d is odd then $-1^{d+1} = 1$, while if d is even than we can write -1 as $\omega^{d/2}$, and we find $\omega^{dk(dk+d)/2} = \omega^{d^2k(k+1)/2}$.) This proves the result for $a = 0$.

Assume $a \geq 1$. If $b \neq 0$ then, by induction hypothesis and by the usual recurrence relation, we get

$$\begin{aligned}
\begin{bmatrix} nd+a \\ kd+b \end{bmatrix}_\omega &= \begin{bmatrix} nd+a-1 \\ kd+b-1 \end{bmatrix}_\omega + \omega^{kd+b}\begin{bmatrix} nd+a-1 \\ kd+b \end{bmatrix}_\omega \\
&= \binom{n}{k}\begin{bmatrix} a-1 \\ b-1 \end{bmatrix}_\omega + \omega^b\binom{n}{k}\begin{bmatrix} a-1 \\ b \end{bmatrix}_\omega \\
&= \binom{n}{k}\begin{bmatrix} a \\ b \end{bmatrix}_\omega.
\end{aligned}$$

Finally, if $b = 0$, then, as $a-1 < d-1$,

$$\begin{aligned}
\begin{bmatrix} nd+a \\ kd \end{bmatrix}_\omega &= \begin{bmatrix} nd+a-1 \\ (k-1)d+d-1 \end{bmatrix}_\omega + \omega^{kd}\begin{bmatrix} nd+a-1 \\ kd \end{bmatrix}_\omega \\
&= \binom{n}{k}\omega^0\begin{bmatrix} a-1 \\ 0 \end{bmatrix}_\omega \\
&= \binom{n}{k}\begin{bmatrix} a \\ b \end{bmatrix}_\omega.
\end{aligned}$$

Remark Compare Lucas' formula

$$\binom{np+a}{kp+b} \equiv \binom{n}{k}\binom{a}{b} \pmod{p}$$

if p is prime and $0 \leq a, b < p$.

Chapter 5

Group actions and cycle index

A cube has six faces, so if we paint each face red, white or blue, the total numbers of ways that we can apply the colours is $3^6 = 729$. However, if we can pick up the cube and move it around, it is natural to count in a different way, where two coloured cubes differing only by a rotation are counted as “the same”. There are 24 rotations of the cube into itself, but the answer to our question is not obtained just by dividing 729 by 24. The purpose of this section is to develop tools for answering such questions.

5.1 Group actions

Let X be a set, and G a set of permutations of X. We write the image of $x \in X$ under the permutation g as x^g. We denote the identity permutation (leaving every element of X where it is) by 1, and the inverse of a permutation g (the permutation h with $x^g = y \Leftrightarrow x^h = x$) by g^{-1}. The composition of two permutations g and h, denoted by gh, is defined by the rule that

$$x^{gh} = (x^g)^h$$

(in other words, apply first g, then h).

We say that G is a *permutation group* if the following conditions hold:

- G contains the identity permutation;
- G contains the inverse of each of its elements;
- G contains the composition of any two of its elements.

For example, the 24 rotational symmetries of a cube form a permutation group on the set of points of the cube.

Until the middle of the nineteenth century, what we have just defined would have simply been called a *group*.Now the definition of a group is more abstract. We don't go into abstract group theory here, but note some terminology arising from this. If G is an abstract group in the modern sense, an *action* of G on the set X is a function associating a permutation of X with each group element, in such a way that the identity, inverse, and composition of permutations correspond to the same concepts in the abstract group.

In particular, if G is a permutation group on a set X, then we can construct actions of G on various auxiliary sets built from X: for example, the set of ordered pairs of elements of X, the set of subsets of X, the set of functions from X to another set (or from another set to X).

For example, G acts on the set $X \times X$ of ordered pairs of elements of X by the rule

$$(x,y)^g = (x^g, y^g)$$

for $x,y \in X$, $g \in G$; that is, the permutation g acts coordinate-wise on ordered pairs, mapping (x,y) to (x^g, y^g).

Thus, the phrases "G is a permutation group on X" and "G acts on X" are almost synonymous; the difference is of less interest to a combinatorialist than to an algebraist.

Suppose that G acts on X. We define a relation $\sim$ on X by the rule that $x \sim y$ if $y = x^g$ for some $g \in G$.

Proposition 5.1 *$\sim$ is an equivalence relation.*

Proof We check the three conditions.

- $x = x^1$, so $x \sim x$: $\sim$ is reflexive.
- Let $x \sim y$. Then $y = x^g$, so $x = y^{g^{-1}}$, so $y \sim x$: $\sim$ is symmetric.
- Let $x \sim y$ and $y \sim z$. Then $x = x^g$ and $z = y^h$, for some $g,h \in G$. Thus, $z = (x^g)^h = x^{gh}$, so $x \sim z$: $\sim$ is transitive.

Note that the three conditions in the definition of a permutation group translate precisely into the three conditions of an equivalence relation.

The equivalence classes of this relation are the *orbits* of G on X.

In our coloured cube example, the group of 24 rotations of the cube acts on the set of 729 colourings of the faces of the cube. Two colourings count "the same" if and only if they are in the same orbit. So our task is to count orbits.

5.2 The Orbit-Counting Lemma

For any permutation g of X, we let $\mathrm{fix}(g)$ denote the number of *fixed points* of g (elements $x \in X$ such that $x^g = x$).

Theorem 5.2 ***(Orbit-Counting Lemma)*** *Let G be a permutation group on the finite set X. Then the number of orbits of G on X is given by the formula*

$$\frac{1}{|G|}\sum_{g\in G}\mathrm{fix}(g).$$

Proof We count in two different ways the number N of pairs (x,g), with $x \in X$, $g \in G$, and $x^g = x$.

On the one hand, clearly

$$N = \sum_{g\in G}\mathrm{fix}(g).$$

On the other hand, we claim that if the point x lies in an orbit $\{x = x_1,\dots,x_n\}$, then the number of permutations $g \in G$ with $x^g = x$ is $|G|/n$. More generally, for any i with $1 \le i \le n$, the number of permutations $g \in G$ with $x^g = x_i$ is independent of i (the proof is an exercise), and so is $|G|/n$.

Hence the number of pairs (y,g) with $y^g = y$ for which y lies in a fixed orbit of size n is $n \cdot |G|/n = |G|$. So each orbit contributes $|G|$ to the sum, and so $N = |G|k$, where k is the number of orbits.

Equating the two values gives the result.

Using this, we can count our coloured cubes. We have to examine the 24 rotations and find the number of colourings fixed by each.

- The identity fixes all $3^6 = 729$ colourings.

- There are three axes of rotation through the mid-points of opposite faces. A rotation through a half-turn about such an axis fixes $3^4 = 81$ colourings: we can choose arbitrarily the colour for the top face, the bottom face, the east and west faces, and the north and south faces (assuming that the axis is vertical). A rotation about a quarter turn fixes $3^3 = 27$ colourings, since all four faces except top and bottom must have the same colour. There are three half-turns and six quarter-turns.

- A half-turn about the axis joining the midpoints of opposite edges fixes $3^3 = 27$ colourings. There are six such rotations.

- A third-turn about the axis joining opposite vertices fixes $3^2 = 9$ colourings. There are eight such rotations.

By Theorem 5.2, the number of orbits is

$$\frac{1}{24}(1\cdot 729+3\cdot 81+6\cdot 27+6\cdot 27+8\cdot 9) = 57,$$

so there are 57 different colourings up to rotation.

At this point, we can give a more combinatorial proof of the formula

$$x(x-1)\cdots(x-n+1) = \sum_{k=1}^{n} s(n,k)x^k$$

from chapter 2. We prove the equivalent form

$$x(x+1)\cdots(x+n-1) = \sum_{k=1}^{n} |s(n,k)|x^k$$

from which the required equation is obtained by substituting $-x$ for x and multiplying by $(-1)^n$. Suppose first that x is a positive integer. Consider the set of functions from $\{1,\ldots,n\}$ to a set X of cardinality x. There are x^n such functions. Now the symmetric group S_n acts on these functions: the permutation g maps the function f to f^g, where

$$f^g(i) = f(ig^{-1}).$$

The orbits are simply the selections of n things from X, where repetitions are allowed and order is not important. So the number of orbits is

$$\binom{x+n-1}{n} = x(x+1)\cdots(x+n-1)/n!$$

(see Chapter 2, Exercise 1).

We can also count the orbits using the Orbit-Counting Lemma. Let g be a permutation in S_n having k cycles. How many functions are fixed by g? Clearly a function f is fixed if and only if it is constant on each cycle of g; its values on the cycles can be chosen arbitrarily. So there are x^k fixed functions. Since the number of permutations with k cycles is $|s(n,k)|$, the Orbit-Counting Lemma shows that the number of orbits is

$$\frac{1}{n!}\sum_{k=1}^{n}|s(n,k)|x^k.$$

Equating the two expressions and multiplying by $n!$ gives the result.

Now the required equation holds for all positive integer values of x, and so it is a polynomial identity.

5.3 Cycle index

It is possible to develop a method for solving the coloured cubes problem which doesn't require extensive recalculation when small changes are made (such as changing the number of colours).

Suppose that we have a set F of objects called "figures", each of which (say f) has a non-negative integer "weight" $w(f)$ associated with it. The number of figures may be infinite, but we assume that there are only a finite number of any given weight: say a_n figures of weight n. The *figure-counting series* is the (ordinary) generating function for these numbers:

$$A(x) = \sum_{n\geq 0} a_n x^n.$$

We attach a figure to each point of a finite set X. (Equivalently, we take a function ϕ from X to the set F of figures.) The *weight* of the function ϕ is just

$$w(\phi) = \sum_{x\in X} w(\phi(x)).$$

Finally, we have a group G of permutations of X. Then G acts on the set of functions by the rule that

$$\phi^g(x) = \phi(xg^{-1}).$$

Clearly $w(\phi^g) = w(\phi)$ for any function ϕ.

We want to find the generating function for the number of functions of each possible weight, but counting two functions as "the same" if they lie in the same orbit of G with the above action. In other words, we want to calculate the *function-counting series*

$$B(x) = \sum_{n\geq 0} b_n x^n,$$

where b_n is the number of orbits consisting of functions of weight n.

In the coloured cubes example, if we take three figures Red, White and Blue, each of weight 0, the figure-counting series is simply 3, and the function-counting series is 57. We could, say, change the weight of Red to 1, so that the figure-counting series is $2+x$; then the function-counting series is the generating function for the numbers of colourings with $0,1,2,\ldots,6$ red faces (up to rotations).

The gadget that does this job is the *cycle index* of G. Each element $g\in G$ can be decomposed into disjoint cycles; let $c_i(g)$ be the number of cycles of length i, for $i=1,\ldots,n=|X|$. Now put

$$z(g) = s_1^{c_1(g)} s_2^{c_2(g)} \cdots s_n^{c_n(g)},$$

where $s_1,\ldots,s_n$ are indeterminates. Then the *cycle index* of G is defined to be

$$Z(G) = \frac{1}{|G|}\sum_{g\in G} z(g).$$

For example, our analysis of the rotations of the cube shows that the cycle index of this group (acting on faces) is

$$\frac{1}{24}(s_1^6 + 3s_1^2s_2^2 + 6s_1^2s_4 + 6s_2^3 + 8s_3^2).$$

We use the notation

$$Z(G; s_i \leftarrow f_i \text{ for } i=1,\ldots,n)$$

for the result of substituting the expression f_i for the indeterminate s_i for $i=1,\ldots,n$.

Theorem 5.3 *If G acts on X, and we attach figures to the points of X with figure-counting series $A(x)$, then the function-counting series is given by*

$$B(x) = Z(G; s_i \leftarrow A(x^i) \textit{ for } i=1,\ldots,n).$$

For example, in the coloured cubes, let Red have weight 1 and the other colours weight 0. Then $A(x) = 2+x$, and the function-counting series is

$$\begin{aligned} B(x) &= \frac{1}{24}((2+x)^6+3(2+x)^2(2+x^2)^2+6(2+x)^2(2+x^4) \\ &\qquad +6(2+x^2)^3+8(2+x^3)^2) \\ &= 10+12x+16x^2+10x^3+6x^4+2x^5+x^6. \end{aligned}$$

Note that putting $x = 1$ recovers the value 57.

Proof The first step is to note that, if we ignore the group action and simply count all the functions, the function-counting series is $B(x) = A(x)^n$, where $n = |X|$. For the term in x^m in $A(x)^n$ is obtained by taking all expressions $m = m_1 + \cdots + m_n$ for m as a sum of n non-negative integers, multiplying the corresponding terms $a_{m_i}^{m_i}$ in $A(x)$, and summing the result. The indicated product counts the number of choices of functions of weights $m_1, \ldots, m_n$ to attach at the points $1, \ldots, n$ of X, so the result is indeed the function-counting series.

Note that this proves the theorem in the case where G is the trivial group.

Next, we have to count the functions of given weight fixed by a permutation $g \in G$. As we have seen, a function is fixed by g if and only if it is constant on the cycles of g. Now if we choose a function of weight r to attach to the points of a particular i-cycle of g, the number of choices is a_r but the contribution to the weight is ir. Arguing as above, the generating function for the number of fixed functions is

$$A(x)^{c_1(g)}A(x^2)^{c_2(g)}\cdots A(x^n)^{c_n(g)} = z(g; s_i \leftarrow A(x^i) \text{ for } i = 1, \ldots, n).$$

Finally, by the Orbit-Counting Lemma, if we sum over $g \in G$ and divide by $|G|$, we find that the function-counting series is

$$B(x) = Z(G; s_i \leftarrow A(x^i) \text{ for } i = 1, \ldots, n).$$

5.4 Labelled and unlabelled

Group actions can be used to clarify the difference between two types of counting of combinatorial objects, namely counting labelled and unlabelled objects.

Typically, we are counting structures “based on” a set of n points: these may be partitions or permutations, or more elaborate relational structures such as graphs, trees, partially ordered sets, etc. An *isomorphism* between two such objects is a bijection between their base sets which preserves the structure.

A *labelled object* is simply an object whose base set is $\{1,2,\ldots,n\}$. Two objects count as different unless they are identical. On the other hand, for unlabelled objects, we wish to count them as the same obtain one from the other by re-labelling the points of the base set. In other words, an *unlabelled object* is an isomorphism class of objects.

For example, for graphs on three vertices, there are eight labelled objects, but four unlabelled ones.

Now the symmetric group S_n acts on the set of all labelled objects on the set $\{1,\ldots,n\}$; its orbits are the unlabelled objects. So counting unlabelled objects is equivalent to counting orbits of S_n in an appropriate action.

A given object A has an automorphism group $\mathrm{Aut}(A)$, consisting of all permutations of the set of points which map the object to itself. The number of different labellings of A is $n!/|\mathrm{Aut}(A)|$, since of the $n!$ labellings, two are the same if and only if they are related by an automorphism of A. (More formally, labellings correspond bijectively to cosets of $\mathrm{Aut}(A)$ in the symmetric group S_n.) So the number of labelled objects is

$$\sum_A \frac{n!}{|\mathrm{Aut}(A)|},$$

where the sum is over the unlabelled objects on n points.

The cycle index method can be applied to give more sophisticated counts. For example, let us count graphs on 4 vertices. The number of pairs of vertices is 6, and each pair is either an edge or a non-edge. So the number of labelled graphs is $2^6 = 64$, and the number of labelled graphs with k edges is $\binom{6}{k}$ for $k = 0,\ldots,6$.

In order to count orbits, we must let S_4 act on the set of 64 graphs. But we can think of a graph as the set of $\binom{4}{2} = 6$ pairs of vertices with a figure (either an edge or a non-edge) attached to each. So we must compute the cycle index of S_4 acting on pairs of vertices. Table 5.1 gives details. The notation $1^2 2^1$, for example, means “two fixed points and one 2-cycle”. Such an element, say the transposition $(1,2)$, fixes the two pairs $\{1,2\}$ and $\{3,4\}$, and permutes the other four pairs in two 2-cycles; so its cycle structure on pairs is $1^2 2^2$.

So the cycle index of the permutation group G induced on pairs by S_4 is

$$Z(G) = \frac{1}{24}(s_1^6 + 9s_1^2 s_2^2 + 8s_3^2 + 6s_2 s_4).$$

Now if we take edges to have weight 1 and non-edges to have weight 0 (that

Cycles on vertices	Cycles on pairs	Number
1^4	1^6	1
$1^2 2^1$	$1^2 2^2$	6
2^2	$1^2 2^2$	3
13	3^2	8
4	24	6

Table 5.1: Cycle index of S_4

is, figure-counting series $A(x) = 1 + x$), the function-counting series is

$$B(x) = 1 + x + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6,$$

the generating function for unlabelled graphs on four vertices by number of edges.

We conclude by summarising some of our earlier results on counting labelled and unlabelled structures. Table 5.2 gives the numbers of labelled and unlabelled structures on n points; $B(n)$ and $p(n)$ are the Bell and partition numbers.

Structure	Labelled	Unlabelled
Subsets	2^n	$n+1$
Partitions	$B(n)$	$p(n)$
Permutations	$n!$	$p(n)$
Total orders	$n!$	1

Table 5.2: Labelled and unlabelled

We see from the table that it is possible, even in very natural cases, to have the same number of labelled objects but different numbers of unlabelled ones, or *vice versa*.

Exercises

5.1. Let G be a permutation group on a finite set X, where $|X| = n > 1$. Suppose that G has only one orbit. Prove that there is an element of G which is a derangement of X (that is, which has no fixed point). Show further that at least a fraction $1/n$ of the elements of G are derangements.

5.2. Use the Cycle Index Theorem to write down a polynomial in two variables x and y in which the coefficient of x^iy^j is the number of cubes in which the faces are coloured red, white and blue, having i red and j blue faces, up to rotations of the cube.

5.3. Find a formula for the number of ways of colouring the faces of the cube with r colours, up to rotations of the cube. Repeat this exercise for the other four Platonic solids.

5.4. A necklace has ten beads, each of which is either black or white, arranged on a loop of string. A cyclic permutation of the beads counts as the same necklace. How many necklaces are there? How many are there if the necklace obtained by turning over the given one is regarded as the same?

Chapter 6

Möbius inversion

Often we are in the situation where we have a number of conditions of varying strength, and we have information about the number of objects which satisfy various combinations of conditions (inclusion); we want to count the objects satisfying none of the conditions (exclusion), or perhaps satisfying some but not others. Of course, the conditions may not all be independent!

6.1 The Principle of Inclusion and Exclusion

Let $A_1,\ldots,A_n$ be subsets of a finite set X. For any non-empty subset J of the index set $\{1,\ldots,n\}$, we put

$$A_J = \bigcap_{j\in J} A_j;$$

by convention, we take $A_\emptyset = X$. The *Principle of Inclusion and Exclusion* (PIE, for short) asserts the following.

Theorem 6.1 *The number of elements of X lying in none of the sets A_i is equal to*

$$\sum_{J\subseteq\{1,\ldots,n\}} (-1)^{|J|}|A_J|.$$

Proof The expression in the theorem is a linear combination of the cardinalities of the sets A_J, and so we can calculate it by working out, for each $x\in X$, the contribution of x to the sum. If K is the set of all indices j for which $x\in A_j$, then x contributes to the terms involving sets $J\subseteq K$, and the contribution is

$$\sum_{J\subseteq K} (-1)^{|J|}.$$

If $|K| = k > 0$, then there are $\binom{k}{j}$ sets of size j in the sum, which is

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

whereas if $K = \emptyset$ then the sum is 1. So the points with $K = \emptyset$ (those lying in no set A_i) each contribute 1 to the sum, and the remaining points contribute nothing. So the theorem is proved.

If there are numbers $m_0, \ldots, m_n$ such that $|A_J| = m_j$ whenever $|J| = j$, then PIE can be written in the simpler form

$$\sum_{j=0}^{n} (-1)^j m_j.$$

Here are a couple of applications.

Example: Surjections The number of functions from an m-set *onto* an n-set is given by the formula

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} (n-j)^m.$$

For let M and N be the sets, with $N = \{1, \ldots, n\}$. Let X be the set of all functions $f : M \longrightarrow N$, and A_i the set of functions whose range does not include the point i. Then A_J is the set of functions whose range includes none of the points of J (that is, functions from M to $N \setminus J$); so $|A_J| = (n-j)^m$ when $|J| = j$. A function is a surjection if and only if it lies in none of the sets A_i. The result follows.

In particular, if $m = n$, then surjections are permutations, and we have

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} (n-j)^n = n!.$$

Example: Derangements This time, let X be the set of all permutations of $\{1, \ldots, n\}$, and A_i the set of permutations fixing i. Then A_j is the set of permutations fixing every point in J; so $|A_J| = (n-j)!$ when $|J| = j$. The permutations lying in none of the sets A_i are the derangements, and so we have

$$\begin{aligned} d(n) &= \sum_{j=0}^{n} (-1)j \binom{n}{j} (n-j)! \\ &= n! \sum_{j=0}^{n} \frac{(-1)^j}{j!}, \end{aligned}$$

in agreement with our earlier result.

The statement of PIE can be generalised to give a formula for the number of elements of X which lie in a given collection of sets A_i and not in the remaining ones (see Exercise 6.5). Indeed, the same formula applies if the numbers concerned are arbitrary real numbers rather than cardinalities of sets:

Theorem 6.2 *Let real numbers a_J and b_J be given for each subset J of $N = \{1,\dots,n\}$. Then the following are equivalent:*

(a) $a_J = \displaystyle\sum_{J\subseteq I\subseteq N} b_I$ *for all* $J \subseteq N$*;*

(b) $b_J = \displaystyle\sum_{J\subseteq I\subseteq N} (-1)^{|I|} a_I$ *for all* $J \subseteq N$*.*

Proof The theorem asserts the form of the solution to a system of linear equations; in other words, the inverse of a certain matrix. However, the same matrix occurs in the original form of PIE.

The theorem as stated involves sums over supersets of the given index set. However, it is easily transformed to involve sums over subsets (see Exercise 6.5. In this form, it is a generalisation of the inverse relationship between the triangular matrix of binomial coefficients and the signed version (see Exercise 6.5).

6.2 Partially ordered sets

In this section, we formalise the kind of lower-triangular matrices which occurred in the last.

A *partial order* on a set X is a binary relation $\leq$ on X which satisfies the following conditions:

- $x \leq x$ (*reflexivity*);
- if $x \leq y$ and $y \leq x$ then $x = y$ (*antisymmetry*);
- if $x \leq y$ and $y \leq z$ then $x \leq z$ (*transitivity*).

It is a *total order* if it satisfies the further condition

- for any x, y, exactly one of $x < y$, $x = y$, $y < x$ holds (*trichotomy*), where $x < y$ is short for $x \leq y$ and $x \neq y$. (Note that antisymmetry implies that at most one of these three conditions holds.)

The usual order relations on the natural numbers, integers, and real numbers are total orders. An important example of a partial order is the relation of *inclusion* on the set of all subsets of a given set. Other important examples of partially ordered sets include

- the positive integers ordered by divisibility (that is, $x \leq y$ if and only if $x \mid y$);
- the subspaces of a finite vector space, ordered by inclusion. (This is known as a *projective space*.)

Any finite totally ordered set can be written as $\{x_1, x_2, \ldots, x_n\}$, where $x_i \leq x_j$ if and only if $i \leq j$.

A set carrying a partial order relation is called a *partially ordered set*, or *poset* for short.

We need to use the following result. A relation σ is an *extension* of a relation ρ if *x rho y* $\Rightarrow$ *x sigma y*; that is, regarding a relation in the usual way as a set of ordered pairs, ρ is a subset of σ.

Theorem 6.3 *Any partial order on a set X can be extended to a total order on X.*

This theorem is easily proved for finite sets: take any pair of elements x,y which are incomparable in the given relation; set $x \leq y$, and include all consequences of transitivity (show that no conflicts arise from this); and repeat until all pairs are comparable. It is more problematic for infinite sets; it cannot be proved from the Zermelo–Fraenkel axioms, but requires an additional principle such as the Axiom of Choice.

The upshot of the theorem for finite sets is that any finite partially ordered set can be written as $X = \{x_1, \ldots, x_n\}$ so that, if $x_i \leq x_j$, then $i \leq j$ (but not necessarily conversely). This is often possible in many ways. For example, the subsets of $\{a,b,c\}$, ordered by inclusion, can be written as

$$\begin{array}{llll} X_1 = \emptyset, & X_2 = \{a\}, & X_3 = \{b\}, & X_4 = \{c\}, \\ X_5 = \{a,b\}, & X_6 = \{a,c\}, & X_7 = \{b,c\}, & X_8 = \{a,b,c\}. \end{array}$$

Now any function f from $X \times X$ to the real numbers can be written as an $n \times n$ matrix A_f, whose (i,j) entry is $f(x_i, x_j)$.

Our results extend to some infinite partially ordered sets, namely, those which are *locally finite*. (A partially ordered set X is locally finite if, for any $x,y \in X$, the *interval*

$$[x,y] = \{z \in X : x \leq z \leq y\}$$

is finite.)

Examples of infinite, locally finite posets include:

- The natural numbers; the integers (with the usual order).
- All finite subsets of an infinite set (ordered by inclusion).
- All finite-dimensional subspaces of an infinite-dimemsional vector space over a finite field (ordered by inclusion).
- The positive integers (ordered by divisibility).

6.3 The incidence algebra of a poset

The *incidence algebra* of the partially ordered set X is defined to be the set of all functions $\alpha : X \times X \to \mathbb{R}$ which have the property that $\alpha(x,y) = 0$ unless $x \leq y$. Note that, for such a function α, the matrix A_α is lower triangular. The algebra operations of addition and multiplication are defined to be the usual matrix operations on the corresponding matrices; that is,

$$\begin{aligned}(\alpha+\beta)(x,y) &= \alpha(x,y)+\beta(x,y),\\ (\alpha\beta)(x,y) &= \sum_{x\leq z\leq y} \alpha(x,z)\beta(z,y).\end{aligned}$$

(These equations shows that the way in which we extend the partial order to a total order does not affect the definitions.)

The definitions of addition and multiplication work equally well for an infinite locally finite poset (since the sum in the formula for multiplication is finite). So the incidence algebra of a locally finite poset is defined.

The incidence algebra has an identity, the function ι given by

$$\iota(x,y) = \begin{cases} 1 & \text{if } x = y,\\ 0 & \text{otherwise.}\end{cases}$$

(The matrix A_ι is the usual identity matrix.) Another important algebra element is the *zeta function* ζ, defined by

$$\zeta(x,y) = \begin{cases} 1 & \text{if } x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

Thus ζ is the characteristic function of the partial order, and an arbitrary function α belongs to the incidence algebra if and only if

$$\zeta(x,y) = 0 \Rightarrow \alpha(x,y) = 0.$$

A lower triangular matrix with ones on the diagonal has an inverse. The *Möbius function* μ of a poset is the inverse of the zeta function. In other words, it satisfies

$$\sum_{x \leq z \leq y} \mu(x,y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $\mu(x,x) = 1$ for all x. Moreover, if we know $\mu(x,z)$ for $x \leq z < y$, then we can calculate

$$\mu(x,y) = -\sum_{x \leq z < y} \mu(x,z).$$

In particular, we see that the values of the Möbius function are all integers.

6.4 Some Möbius functions

By definition, the Möbius function of a poset satisfies the following:

Proposition 6.4 *Let f and g be elements of the incidence algebra of a poset X (that is, functions on $X \times X$ satisfying $f(x,y) = g(x,y) = 0$ unless $x \leq y$. Then the following conditions are equivalent:*

(a) $g(x,y) = \sum_{x \leq z \leq y} f(x,z)$;

(b) $f(x,y) = \sum_{x \leq z \leq y} g(x,z)\mu(z,y)$.

This result is referred to as *Möbius inversion*. In order to use it, we have to compute the Möbius functions of various posets. Note that the Möbius function is local, in the sense that the value of $\mu(x,y)$ is determined by the structure of the interval $[x,y] = \{z : x \leq z \leq y\}$.

One important result is the following. Let $X_1,\dots,X_r$ be posets. The *direct product* $X_1\times\cdots\times X_r$ is the poset whose elements are all r-tuples $(x_1,\dots,x_r)$ with $x_i\in X_i$ for $1\le i\le r$; the order is given by

$$(x_1,\dots,x_r)\le(y_1,\dots,y_r)\Leftrightarrow x_i\le i_i \text{ for } 1\le i\le r,$$

where the order $x_i\le y_i$ is that in the poset X_i.

Proposition 6.5 *The Möbius function of the direct product $X_1\times\cdots\times X_r$ is given by*

$$\mu((x_1,\dots,x_r),(y_1,\dots,y_r))=\prod_{i=1}^{r}\mu_i(x_i,y_i),$$

where μ_i is the Möbius function of X_i.

Proof It is enough to show that

$$\sum_{\substack{x_i\le z_i\le y_i\\ 1\le i\le r}}\prod_{i=1}^{r}\mu_i(x_i,z_i)=0.$$

Now the left-hand side of this expression factorises as

$$\prod_{i=1}^{r}\sum_{x_i\le z_i\le y_i}\mu_i(x_i,z_i),$$

and the inner sum is zero by definition of the Möbius function μ_i.

Example: the integers In the poset of integers, with the usual order, the Möbius function is given by

$$\mu(x,y)=\begin{cases}1 & \text{if } y=x;\\ -1 & \text{if } y=x+1;\\ 0 & \text{otherwise.}\end{cases}$$

Example: Finite subsets of a set In this case, the Möbius function is

$$\mu(X,Y)=(-1)^{|Y-|X|}\text{ for } X\subseteq Y,$$

and of course $\mu(X,Y)=0$ otherwise. For let $X\subseteq Y$, and let $Y\setminus X=\{z_1,\dots,z_n\}$. We claim that the interval $[X,Y]$ is isomorphic to $\{0,1\}^n$, the direct product of n copies of $\{0,1\}\subseteq\mathbb{Z}$. The isomorphism takes a set Z with $X\le Z\le Y$ to the n-tuple $(e_1,\dots,e_n)$, where

$$e_i=\begin{cases}1 & \text{if } z_i\in Z,\\ 0 & \text{otherwise.}\end{cases}$$

So $\mu(X,Y)$ is equal to $\mu((0,\dots,0),(1,\dots,1))$ calculated in $\{0,1\}^n$; by Proposition 6.5 this is $\mu(0,1)^n$, and $\mu(0,1)=-1$ by the preceding example.

Example: Positive integers ordered by divisibility Suppose that m divides n. Let $n/m = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$, where $p_1, \ldots, p_r$ are distinct primes and $a_1, \ldots, a_r$ positive integers. Then the interval $[m,n]$ is isomorphic to the direct product

$$[0,a_1] \times \cdots \times [0,a_r]$$

of intervals $[0,a_i]$ in $\mathbb{Z}$. The correspondence is given by

$$(b_1, \ldots, b_r) \leftrightarrow m p_1^{b_1} \cdots p_r^{b_r}.$$

By the first example, we see that $\mu(m,n) = 0$ if any $a_i > 1$, that is, if n/m is divisible by the square of a prime. If n/m is the product of s distinct primes, then $\mu(m,n) = (-1)^s$. To summarise:

$$\mu(m,n) = \begin{cases} (-1)^s & \text{if } n/m \text{ is the product of } s \text{ distinct primes;} \\ 0 & \text{if } m \text{ doesn't divide } n \text{ or if } n/m \text{ is not squarefree.} \end{cases}$$

Example: Subspaces of a finite vector space By the Second Isomorphism Theorem, if U and W are subspaces of V with $U \subseteq W$, then the interval $[U,W]$ is isomorphic to the poset of subspaces of W/U, and in particular depends only on $\dim(W) - \dim(U)$. It suffices to calculate $\mu(\{0\},V)$, where V is an n-dimensional vector space over $\mathrm{GF}(q)$.

Now putting $x = -1$ in the q-binomial theorem, we obtain

$$\sum_{k=0}^{n-1} (-1)^k q^{k(k-1)/2} \begin{bmatrix} n \\ k \end{bmatrix}_q$$

for $n > 0$. This is exactly the inductive step in the proof that $\mu(\{0\},V) = (-1)^n q^{n(n-1)/2}$ for $n > 0$. For there are $\left[{n \atop k}\right]_q$ k-dimensional subspaces of V, and the induction hypothesis asserts that $\mu(\{0\},W) = (-1)^k q^{k(k-1)/2}$ for each such subspace; then the identity shows that $\mu(\{0\},V)$ must have the claimed value.

So, in general, $\mu(U,W) = (-1)^n q^{n(n-1)/2}$ if $U \subseteq W$ and $\dim(W/U) = n$; and of course, $\mu(U,W) = 0$ if $U \not\subseteq W$.

6.5 Classical Möbius inversion

All our examples in the preceding section have the special property that each interval $[x,y]$ is isomorphic to $[e,z]$, where e is a fixed element of the poset, and z

depends on x and y. Thus, for the integers, $e = 0$ and $z = y - x$; for subsets of a set, $e = \emptyset$ and $z = y \setminus x$; for positive integers ordered by divisibility, $e = 1$ and $z = y/x$; and for subspaces of a vector space, $e = \{0\}$ and $z = y/x$ (the quotient space).

Thus, in these cases, the Möbius function satisfies $\mu(x,y) = \mu(e,z)$, so it can be written as a function of one variable z. Abusing notation, we use the same symbol μ. In the four cases, we have:

- $\mu(0) = 1$, $\mu(1) = -1$, $\mu(z) = 0$ for $z \geq 2$;
- $\mu(Z) = (-1)^{|Z|}$;
- $\mu(z) = (-1)^s$ if z is the product of s distinct primes, $\mu(z) = 0$ if z is not squarefree;
- $\mu(Z) = (-1)^k q^{k(k-1)/2}$, where $k = \dim(Z)$.

The third of these is the “classical” Möbius function, and plays an important role in number theory. If you see $\mu(z)$ without any further explanation, it probably means the classical Möbius function. In this case, Möbius inversion can be stated as follows:

Proposition 6.6 *Let f and g be functions on the positive integers. Then the following are equivalent:*

(a) $g(n) = \sum_{m|n} f(m)$;

(b) $f(n) = \sum_{m|n} g(m)\mu(n/m)$.

Here are two applications of this result.

Example: Euler’s function Euler’s ϕ-function (sometimes called the *totient function* is the function ϕ defined on the positive integers by the rule that $\phi(n)$ is the number of integers x with $1 \leq x < n$ coprime to n.

If $\gcd(x,n) = d$, then $\gcd(x/d, n/d) = 1$. So the number of x in this range with $\gcd(x,n) = d$ is $\phi(n/d)$, and we have

$$\sum_{d|n} \phi(n/d) = n,$$

or, putting $m = n/d$,

$$\sum_{m|n} \phi(m) = n.$$

Now Möbius inversion gives

$$\phi(n) = \sum_{m|n} m\mu(n/m).$$

From this it is easy to deduce that, if $n = p_1^{a_1} \cdots p_r^{a_r}$, where p_i are distinct primes and $a_i > 0$, then

$$\phi(n) = p_1^{a_1-1}(p_1 - 1) \cdots p_r^{a_r-1}(p_r - 1).$$

Example: Irreducible polynomials Let $f_q(n)$ be the number of monic irreducible polynomials of degree n over $\mathrm{GF}(q)$. By Theorem 4.9,

$$\sum_{m|n} m f_q(m) = q^n.$$

So, by Möbius inversion, we have a formula for $f_q(n)$:

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

For example, the number of irreducible polynomials of degree 6 over $\mathrm{GF}(2)$ is

$$\frac{1}{6}(2^6 - 2^3 - 2^2 + 2^1) = 9.$$

(Why is the word "monic" not needed here?)

Exercises

6.1. Let $A_1, \ldots, A_n$ be subsets of X. For $J \subseteq N = \{1, \ldots, n\}$, let A_j consist of the points of X lying in A_i for all $i \in J$, and B_j the points lying in A_i if $i \in J$ and not if $i \notin J$. Show that

$$|B_J| = \sum_{J \subseteq K \subseteq N} (-1)^{|K \setminus J|} |A_K|.$$

6.2. Prove that, with the hypotheses of Theorem 6.2, the following conditions are equivalent:

(a) $a_J = \sum_{I \subseteq J} b_I$ for all $J \subseteq N$;

(b) $b_J = \sum_{I \subseteq J} (-1)^{|J \setminus I|} a_I$ for all $J \subseteq N$.

6.3. By taking the numbers a_J and b_J of the preceding exercise to depend only on the cardinality j of J, show that the following statements are equivalent for two sequences (x_i) and (y_i):

(a) $x_j = \sum_{i=0}^{j} y_i$;

(b) $y_j = \sum_{i=0}^{j} (-1)^{j-i} y_i$.

6.4. Prove that

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

6.5. Let x and y be elements of a poset X, with $x \leq y$. A *chain* from x to y is a sequence $x = x_0, x_1, \ldots, x_l = y$ with $x_{i-1} < x_i$ for $i = 1, \ldots, l$; its *length* is l. Show that

$$\mu(x,y) = \sum_{c \in C} (-1)^{l(c)},$$

where C is the set of all chains from x to y, and $l(c)$ is the length of c.

6.6. Let $d(n)$ be the number of divisors of the positive integer n. Prove that

$$\sum_{m|n} d(m)\mu(n/m) = 1$$

for $n > 1$.

6.7. Let $\mathcal{P}(X)$ denote the poset whose elements are the partitions of the set X, with $P \leq Q$ if P refines Q (that is, every part of P is contained in a part of Q). Let E be the partition into sets of size 1.

(a) Show that, if the parts of P have sizes $a_1,\dots,a_r$, then

$$\mu(E,P) = (a_1-1)!\cdots(a_r-1)!\,.$$

(b) Show that any interval $[P,Q]$ is isomorphic to a product of posets of the form $\mathcal{P}(X_j)$, and hence calculate $\mu(P,Q)$.

6.8. Let G be the cyclic group consisting of all powers of the permutation

$$g : 1 \to 2 \to \cdots \to n \to 1.$$

Show that the cycle index of G is

$$Z(G) = \frac{1}{n}\sum_{m|n}\phi(n/m)s_{n/m}^{m},$$

where ϕ is Euler's function.

6.9. A necklace is made of n beads of q different colours. Necklaces which differ only by a rotation are regarded as the same. Show that the number of different necklaces is

$$\frac{1}{n}\sum_{m|n}q^m\phi(n/m),$$

while the number which have no rotational symmetry is

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

(Notice that, if q is a prime power, the second expression is equal to the number of monic irreducible polynomials of degree n over $\mathrm{GF}(q)$. Finding a bijective proof of this fact is much harder!)

6.10. A function F on the natural numbers is said to be *multiplicative* if

$$\gcd(m,n) = 1 \Rightarrow F(mn) = F(m)F(n).$$

(a) Suppose that F and G are multiplicative. Show that the function H defined by

$$H(n) = \sum_{k|n}F(k)G(n/k)$$

is multiplicative.

(b) Show that the Möbius and Euler functions are multiplicative.

(c) Let $d(n)$ be the number of divisors of n, and $\sigma(n)$ the sum of the divisors of n. Show that d and σ are multiplicative.

6.11. Prove the following "approximate version" of PIE:

Let $A_1,\ldots,A_n,A'_1,\ldots,A'_n$ be subsets of a set X. For $I \subseteq N = \{1,\ldots,n\}$, let

$$a_I = \left|\bigcap_{i\in I} A_i\right|, \qquad a'_I = \left|\bigcap_{i\in I} A'_i\right|.$$

If $a_I = a'_I$ for all *proper* subsets I of N, then $|a_N - a'_N| \le |X|/2^{n-1}$.

Remark: For more general approximate versions of PIE, see N. Linial and N. Nisan, Approximate inclusion-exclusion, *Combinatorica* **10** (1990), 349–365.

Chapter 7

Species

Species, invented by André Joyal in 1980, provide an attempt to unify some of the many structures and techniques which appear in combinatorial enumeration. I don't attempt to be too precise about what a species is. Think of it as a set of "points" carrying some structure (a graph, a poset, a permutation, etc.) We can ask for the number of labelled or unlabelled structures on n points in a given species.

7.1 Cayley's Theorem

We begin with a particular species where there is a simple but unexpected formula for the labelled counting problem. A *tree* is a connected graph with no cycles. It is straightforward to show that a tree on n vertices contains $n-1$ edges, and that any connected graph has a spanning tree (that is, some set of $n-1$ of its edges forms a tree). Moreover, any tree has a vertex lying on only one edge (since the average number of edges per vertex is $2(n-1)/n < 2$). Such a vertex is called a *leaf*. If we remove from a tree a leaf and its incident edge, the result is still a tree.

Cayley's Theorem states:

Theorem 7.1 *The number of labelled trees on n vertices is* n^{n-2}.

There are many different proofs of this theorem. Below, we will see two proofs which are made clearer by means of the concept of species. But first, one of the classics:

Prüfer's proof of Cayley's Theorem We construct a bijection between the set of all trees on the vertex set $\{1,\ldots,n\}$ and the set of all $(n-2)$-tuples of elements from this set. The tuple associated with a tree is called its *Prüfer code*.

First we describe the map from trees to Prüfer codes. Start with the empty code. Repeat the following procedure until only two vertices remain: select the leaf with smallest label; append the label of its unique neighbour to the code; and then remove the leaf and its incident edge.

Next, the construction of a tree from a Prüfer code P. We use an auxiliary list L of vertices added as leaves, which is initially empty. Now, while P is not empty, we join the first element of P to the smallest-numbered vertex v which is not in either P or L, and then add v to L and remove the first element of P. When P is empty, two vertices have not been put into L; the final edge of the tree joins these two vertices.

I leave to the reader the task of showing that these two constructions define inverse bijections. The method actually gives much more information:

Proposition 7.2 *In the tree with Prüfer code P, the valency of the vertex i is one more than the number of occurrences of i in P.*

For, at the conclusion of the second algorithm, if we add in the last two vertices to L, then L contains each vertex precisely once; and edges join each of the first $n-2$ vertices of L to the corresponding vertex in P, together with an edge joining the last two vertices of L.

Using this, one can count labelled trees with any prescribed degree sequence.

7.2 Species and counting

Almost the only thing we assume about a species $\mathcal{G}$ is that, for each n, there are only a finite number of $\mathcal{G}$-objects on n points (so that we can count them). The only property we use of the objects in a species is that we "know" whether a bijective map between the point sets of two objects is an isomorphism between them (and hence we know the automorphism group of each object).

We make one further (inessential but convenient) assumption, namely that there is a unique object on the empty set of points.

We say that two species are *equivalent* (written $\mathcal{G} \sim \mathcal{H}$) if there is a bijection between the objects of the two species on a given point set such that the automorphism groups of corresponding objects are equal.

The most important formal power series associated with a species is its *cycle index,* which is defined by the rule

$$\tilde{Z}(\mathcal{G}) = \sum_{A \in \mathcal{G}} Z(\mathrm{Aut}(A)),$$

where $\mathrm{Aut}(A)$ is the automorphism group of A. Clearly, equivalent objects have the same cycle index.

The cycle index is well-defined since a monomial $s_1^{a_1} \cdots s_r^{a_r}$ arises only from cycle indices involving $n = \sum_{i=1}^{r} ia_i$ points, and by assumption there are only finitely many of these.

There are two important specialisations of the cycle index of a species $\mathcal{G}$; these are the exponential generating function

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

for the number G_n of labelled n-element $\mathcal{G}$-objects (that is, objects on the point set $\{1, \ldots, n\}$); and the ordinary generating function

$$g(x) = \sum_{n \geq 0} g_n x^n$$

for the number g_n of unlabelled n-element $\mathcal{G}$-objects (that is, isomorphism classes).

Theorem 7.3 *Let $\mathcal{G}$ be a species. Then*

(a) $G(x) = \tilde{Z}(\mathcal{G}; s_1 \leftarrow x, s_i \leftarrow 0$ for $i > 1)$;

(b) $g(x) = \tilde{Z}(\mathcal{G}; s_i \leftarrow x^i)$.

Proof The number of different labellings of an object A on n points is clearly $n!/|\mathrm{Aut}(A)|$. So it is enough to show that, for any permutation group G, we have

$$\begin{aligned} Z(G; s_1 \leftarrow x, s_i \leftarrow 0 \text{ for } i > 1) &= x^n/|G|, \\ Z(G; s_i \leftarrow x^i) &= x^n. \end{aligned}$$

The first equation holds because putting $s_i = 0$ for all $i > 1$ kills all permutations except the identity. The second holds because, with this substitution, each group element contributes x^n, and the result is $1/|G| \sum_{g \in G} x^n = x^n$.

7.3 Examples of species

There are a few simple species for which we can do all the sums explicitly.

Example: Sets The species $\mathcal{S}$ has as its objects the finite sets, with one set of each cardinality up to isomorphism. Its cycle index was calculated in Chapter 5:

$$\tilde{Z}(\mathcal{S}) = \sum_{n\geq 0}(S_n) = \exp\left(\sum_{i\geq 1}\left(\frac{s_i}{i}\right)\right).$$

Hence we find that

$$\begin{aligned} S(x) &= \exp(x),\\ s(x) &= \exp\left(\sum_{i\geq 1}\frac{x^i}{i}\right)\\ &= \exp(-\log(1-x))\\ &= \frac{1}{1-x}, \end{aligned}$$

in agreement with the fact that $S_n = s_n = 1$ for all $n \geq 0$.

Example: Total orders Let $\mathcal{L}$ be the species of total (or linear) orders. Each n-set can be totally ordered in $n!$ ways, all of which are isomorhic, and each of which is rigid (that is, has the trivial automorphism group).

We have

$$\tilde{Z}(\mathcal{L}) = \sum_{n\geq 0} s_1^n = \frac{1}{1-s_1},$$

so that

$$L(x) = l(x) = \frac{1}{1-x}.$$

Example: Circular orders The species $\mathcal{C}$ consists of *circular orders*. An element of this species corresponds to placing the points of the object around a circle, where only the relative positions are considered, and there is no distinguished starting point. Thus, there is just one unlabelled n-element object in $\mathcal{C}$ for all n, and the number of labelled objects is equal to the number $(n-1)!$ of cyclic permutations for $n \geq 1$. The unique n-element structure has $\phi(m)$ automorphisms each with n/m cycles of length m for all m dividing n, where ϕ is Euler's function. Hence (see Exercise 7.2),

$$\begin{aligned}
\tilde{Z}(\mathcal{C}) &= 1-\sum_{m\geq 1}\frac{\phi(m)}{m}\log(1-s_m),\\
C(x) &= 1+\sum_{n\geq 1}\frac{x^n}{n}=1-\log(1-x),\\
c(x) &= \sum x^n=\frac{1}{1-x}.
\end{aligned}$$

Example: Permutations An object of the species $\mathcal{P}$ consists of a set carrying a permutation. We will see later how $\mathcal{P}$ can be expressed as a composition, from which its cycle index can be deduced (Exercise 7.2). We have

$$\begin{aligned}
\tilde{Z}(\mathcal{P}) &= \prod_{n\geq 1}(1-s_n)^{-1},\\
P(x) &= \frac{1}{1-x},\\
p(x) &= \prod_{n\geq 1}(1-x^n)^{-1}.
\end{aligned}$$

The function $p(x)$ is the generating function for number partitions. For, as we saw earlier, an unlabelled permutation is the same as a conjugacy class of permutations; and conjugacy classes are determined by their cycle structure.

7.4 Operations on species

There are several ways of building new species from old; only a few important ones are discussed here.

Products Let $\mathcal{G}$ and $\mathcal{H}$ be species. We define the *product* $\mathcal{K}=\mathcal{G}\times\mathcal{H}$ as follows: an object of $\mathcal{K}$ on a set X consists of a distinguished subset Y of X, a $\mathcal{G}$-object on Y, and a $\mathcal{H}$-object on $X\setminus Y$.

Since these objects are chosen independently, it is easy to check that

$$\tilde{Z}(\mathcal{G}\times\mathcal{H})=\tilde{Z}(\mathcal{G})\tilde{Z}(\mathcal{H}).$$

Since the generating functions for labelled and unlabelled structures are specialisations of the cycle index, we have similar multiplicative formulae for them.

For example, if $\mathcal{S}$, $\mathcal{G}$ and $\mathcal{G}^\circ$ are the species of sets, graphs, and graphs with no isolated vertices respectively, then

$$\mathcal{G} \sim \mathcal{S} \times \mathcal{G}^\circ.$$

Substitution Let $\mathcal{G}$ and $\mathcal{H}$ be species. We define the *substitution* $\mathcal{K} = \mathcal{G}[\mathcal{H}]$ as follows: an object of $\mathcal{K}$ on a set X consists of a partition of X, an $\mathcal{H}$-object on each part of the partition, and a $\mathcal{G}$-object on the set of parts of the partition.

Alternatively, we may regard it as a $\mathcal{G}$-object in which every point is replaced by a *non-empty* $\mathcal{H}$-object.

The cycle index is obtained from that of $\mathcal{G}$ by the substitution

$$s_i \leftarrow \tilde{Z}(\mathcal{H}; s_j \leftarrow s_{ij}) - 1$$

for all i. (The -1 in the formula corresponds to removing the empty $\mathcal{H}$-structure before substituting.)

From this, we see that the exponential generating functions for labelled structures obey the simple substitution law:

$$K(x) = G(H(x) - 1).$$

The situaation for unlabelled structures is more complicated, and $k(x)$ cannot be obtained from $g(x)$ and $h(x)$ alone. Instead, we have

$$k(x) = \tilde{Z}(\mathcal{G}; s_i \leftarrow h(x^i) - 1).$$

This equation also follows from the Cycle Index Theorem, since we are counting functions on $\mathcal{G}$-structures where the figures are non-empty $\mathcal{H}$-structures with weight equal to cardinality.

For example, if $\mathcal{S}$, $\mathcal{P}$ and $\mathcal{C}$ are the species of sets permutations, and circular orders, then the standard decomposition of a permutation into disjoint cycles can be written

$$\mathcal{P} \sim \mathcal{S}[\mathcal{C}].$$

The counting series for labelled structures are given by

$$S(x) \quad = \quad \sum_{n\geq 0} \frac{x^n}{n!} = \exp(x),$$

$$\begin{aligned} P(x) &= \sum_{n\geq 0} \frac{n!x^n}{n!} = \frac{1}{1-x}, \\ C(x) &= 1+\sum_{n\geq 0} \frac{(n-1)!x^n}{n!} = 1-\log(1-x); \end{aligned}$$

so the equation above becomes

$$\frac{1}{1-x} = \exp(-\log(1-x)),$$

So the decomposition of a permutation into cycles is the combinatorial equivalent of the fact that exp and log are inverse functions!

Rooted (or pointed) structures Given a species $\mathcal{G}$, let $\mathcal{G}^*$ be the species of *rooted* $\mathcal{G}$*-structures*: such a structure consists of a $\mathcal{G}$-structure with a distinguished point.

We have

$$\tilde{Z}(\mathcal{G}^*) = s_1 \frac{\partial}{\partial s_1} \tilde{Z}(\mathcal{G}),$$

and so

$$G^*(x) = x\frac{\mathrm{d}}{\mathrm{d}x}G(x).$$

Sometimes it is convenient to remove the distinguished point. This just removes the factors s_1 and t in the above formulae, so that this operation corresponds to differentiation. As a result, we denote the result by $\mathcal{G}'$.

For example, if $\mathcal{C}$ is the class of cycles, then $\mathcal{C}'$ corresponds to the class $\mathcal{L}$ of total (linear) orders. We have

$$L(x) = \frac{\mathrm{d}}{\mathrm{d}x}C(x) = \frac{\mathrm{d}}{\mathrm{d}x}(1-\log(1-x)) = \frac{1}{1-x},$$

in agreement with the preceding example.

7.5 Cayley's Theorem revisited

The notion of species can be used to give two further proofs of Cayley's Theorem.

First proof Let $\mathcal{L}$ and $\mathcal{P}$ be the species of total (or linear) orders and permutations, respectively. These species are quite different, but have the property that the numbers of labelled objects on n points are the same (namely $n!$).

Hence the numbers of labelled objects in the two species $\mathcal{L}[\mathcal{T}^*]$ and $\mathcal{P}[\mathcal{T}^*]$ are equal. (Here $\mathcal{T}^*$ is the species of rooted trees.)

Consider an object in $\mathcal{L}[\mathcal{T}^*]$. This consists of a linear order $(x_1,\ldots,x_r)$, with a rooted tree T_i at x_i for all i. I claim that this is equivalent to a tree with two distinguished vertices. Take edges $\{x_i,x_{i+1}\}$ for $i=1,\ldots,r-1$, and identify x_i with the root of T_i for all i. The resulting graph is a tree. Conversely, given a tree with two distinguished vertices x and y, there is a unique path from x to y in the tree, and the remainder of the tree consists of rooted trees attached to the vertices of the path.

Now consider an object in $\mathcal{P}[\mathcal{T}^*]$. Identify the root of each tree with a point of the set on which the permutation acts, and orient each edge of this tree towards the root. The resulting structure defines a function f on the point set, where

- if v is a root, then $f(v)$ is the image of v under the permutation;
- if v is not a root, then $f(v)$ is the unique vertex for which $(v,f(v))$ is a directed edge of one of the trees.

Conversely, given a function $f:X\to X$, the set Y of periodic points of f has the property that f induces a permutation on it; the pairs $(v,f(v))$ for which v is not a periodic point have the structure of a family of rooted trees, attached to Y at the point for which the iterated images of v under f first enter Y.

So the numbers of trees with two distinguished points is equal to the number of functions from the vertex set to itself. Thus, if there are $F(n)$ labelled trees, we see that

$$n^2F(n)=n^n,$$

from which Cayley's Theorem follows.

Second proof As in the preceding proof, let $\mathcal{T}^*$ denote the species of rooted trees. If we remove the root from a rooted tree, the result consists of an unordered collection of trees, each of which has a natural root (at the neighbour of the root of the original tree). Conversely, given a collection of rooted trees, add a new root, joined to the roots of all the trees in the collection, to obtain a single rooted tree. So, if $\mathcal{E}$ denotes the species consisting of a single 1-vertex structure, and $\mathcal{S}$ the species of sets, we have

$$\mathcal{T}^*\sim\mathcal{E}\times\mathcal{S}[\mathcal{T}^*].$$

Hence, for the exponential generating functions for labelled structures, we have

$$T^*(x) = x\exp(T^*(x)).$$

This is, formally, a recurrence relation for the coefficients of $T^*(x)$, and we need to show that the nth coefficient is n^{n-1}. This can be done most easily with the technique of *Lagrange inversion*, which is discussed in the next chapter.

7.6 What is a species?

We have proceeded this far without ever giving a precise definition of a species. The informal idea is that an object of a species is constructed from a finite set, and bijections between finite sets induce isomorphisms of the objects built on them.

It turns out that mathematics does provide a language to describe this, namely *category theory*. It would take us too far afield to give all the definitions here. In essence, a category consists of a collection of *objects* with a collection of *morphisms* between them. In the only case with which we deal, objects are sets and morphisms are set mappings. In particular, the class $\mathfrak{S}$ whose objects are all finite sets and whose morphisms are all bijections between them satisfies the axioms for a category.

Now a species is simply a *functor F* from $\mathfrak{S}$ to itself. This means that F associates to each finite set S a set $F(S)$, and to each bijection $f : S \to S'$ a bijection $F(f) : F(S) \to F(S')$, such that F respects composition and identity (that is, $F(f_1 f_2) = F(f_1)F(f_2)$ and $F(1_S) = 1_{F(S)}$, where 1_S is the identity map on S).

The standard reference on species (apart from Joyal's original paper) is the book by Bergeron, Labelle and Leroux.

Exercises

7.1. Count the labelled trees in which the vertex i has valency a_i for $1 \le i \le n$, where $a_1, \ldots, a_n$ are positive integers with sum $2n-2$.

7.2. Show that the cycle index for the species $\mathcal{C}$ of circular structures is

$$\tilde{Z}(\mathcal{C}) = 1 - \sum_{m\ge1} \frac{\phi(m)}{m} \log(1 - s_m).$$

Use the fact that

$$\mathcal{P} \sim \mathcal{S}[\mathcal{C}]$$

to show that

$$\tilde{Z}(\mathcal{P}) = \prod_{n\geq 1}(1-s_n)^{-1}.$$

Can you give a direct proof of this?

7.3. Use the result of the preceding exercise, and the fact that $c_n = 1$ for all n (where c_n is the number of unlabelled n-element structures in $\mathcal{C}$) to prove the identity

$$\prod_{m\geq 1}(1-x^m)^{-\phi(m)/m} = \exp(x/(1-x)).$$

7.4. Suppose that g_n is the number of unlabelled n-element objects in the species $\mathcal{G}$. Show that the generating function for unlabelled structures in $\mathcal{S}[\mathcal{G}]$ is

$$\prod_{n\geq 1}(1-x^n)^{-g_n}.$$

Verify this combinatorially in the case $\mathcal{G} = \mathcal{S}$. How would you describe the objects of $\mathcal{S}[\mathcal{S}]$?

7.5. Let $\mathcal{G}$ be a species. The *Stirling numbers* of $\mathcal{G}$ are the numbers $S(\mathcal{G})(n,k)$, defined to be the number of partitions of an n-set into k parts with a $\mathcal{G}$-object on each part.

(a) Prove that, for $\mathcal{G} = \mathcal{S}$, $\mathcal{C}$ and $\mathcal{L}$ respectively, the Stirling numbers are respectively the Stirling numbers $S(n,k)$ of the second kind, the unsigned Stirling numbers $|s(n,k)|$ of the first kind, and the Lah numbers $L(n,k)$ respectively.

(b) Let $S(\mathcal{G})$ be the lower triangular matrix of Stirling numbers of $\mathcal{G}$. Prove that

$$S(\mathcal{G})S(\mathcal{H}) = S(\mathcal{H}[\mathcal{G}]).$$

(c) Let (a_n) and (b_n) be sequences of positive integers with exponential generating functions $A(x)$ and $B(x)$ respectively. Prove that the following two conditions are equivalent:

- $a_0 = b_0$ and $b_n = \sum_{k=1}^{n} S(\mathcal{G})(n,k)a_k$ for $n \geq 1$;
- $B(x) = A(G(x)-1)$.

7.6. A *forest* is a graph whose connected components are trees. Show that there is a bijection between labelled forests of rooted trees on n vertices, and labelled rooted trees on $n+1$ vertices with root $n+1$.

Hence show that, if a forest of rooted trees on n vertices is chosen at random, then the probability that it is connected tends to the limit $1/\mathrm{e}$ as $n \to \infty$.

Remark It is true but harder to prove that the analogous limit for unrooted trees is $1/\sqrt{\mathrm{e}}$.

7.7. Let $\mathcal{U}$ be the *subset* species: a $\mathcal{U}$-object consists of a distinguished subset of its ground set. Calculate the cycle index of $\mathcal{U}$. Hence or otherwise prove that the enumeration functions of $\mathcal{U}$ are

$$\begin{aligned} U(x) &= \exp(2x), \\ u(x) &= (1-x)^{-2}. \end{aligned}$$

Chapter 8

Lagrange inversion

A formal power series over a field, with zero constant term and non-zero term in x, has an inverse with respect to composition. Indeed, the set of all such formal power series is a group, which has recently become known as the *Nottingham group*. However, the basic facts are much older. The associative, closure, and identity laws are obvious, and the rule for finding the inverse in characteristic zero is known as *Lagrange inversion*.

8.1 The theorem

The basic fact can be stated as follows.

Proposition 8.1 *Let f be a formal power series over $\mathbb{R}$, with $f(0)=0$ and $f'(0)\neq 0$. Then there is a unique formal power series g such that $g(f(x))=x$; the coefficient of y^n in $g(y)$ is*

$$\left[\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}\left(\frac{x}{f(x)}\right)^n\right]_{x=0}\Big/ n!.$$

This can be expressed in a more convenient way for our purpose. Let

$$\phi(x)=\frac{x}{f(x)}.$$

Then the inverse function g is given by the functional equation

$$g(y)=y\phi(g(y)).$$

Then Lagrange inversion has the form

$$g(y) = \sum_{n\geq 1} \frac{b_n y^n}{n!},$$

where

$$b_n = \left[\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}\phi(x)^n\right]_{x=0}.$$

Example: Cayley's Theorem The exponential generating function for rooted trees satisfies the equation

$$T^*(x) = x\exp(T^*(x)).$$

With $\phi(x) = \exp(x)$, we find that the coefficient of $y^n/n!$ in $T^*(y)$ is

$$\left[\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}\exp(nx)\right]_{x=0} = n^{n-1},$$

proving Cayley's Theorem once again.

8.2 Proof of the theorem

The proof of Lagrange's inversion formula involves a considerable detour. The treatment here follows the book by Goulden and Jackson. Throughout this section, we assume that the coefficients form a field of characteristic zero; for convenience, we assume that the coefficient ring is $\mathbb{R}$.

First, we extend the notion of formal power series to *formal Laurent series*, defined to be a series of the form

$$f(x) = \sum_{n\geq m} a_n x^n,$$

where m may be positive or negative. If the series is not identically zero, we may assume without loss of generality that $a_m \neq 0$, in which case m is the *valuation* of f, written

$$m = \mathrm{val}(f).$$

We define addition, multiplication, composition, differentiation, etc., for formal Laurent series as for formal power series. In particular, $f(g(x))$ is defined for any formal Laurent series f,g with $\mathrm{val}(g) > 0$. (This is less trivial than the analogous result for formal power series. In particular, we need to know that $g(x)^{-m}$ exists as a formal Laurent series for $m > 0$. It is enough to deal with the case $m = 1$, since certainly $g(x)^m$ exists. If $\mathrm{val}(g) = r$, then $g(x) = x^r g_1(x)$, and so $g(x)^{-1} = x^{-r} g_1(x)^{-1}$, and we have seen that $g_1(x)_{-1}$ exists as a formal power series, since $g_1(0)$ is invertible.

We denote the derivative of the formal Laurent series $f(x)$ by $f'(x)$.

We also introduce the following notation: $[x^n]f(x)$ denotes the coefficient of x^n in the formal power series (or formal Laurent series) $f(x)$. The case $n = -1$ is especially important, as we learn from complex analysis. The value of $[x^{-1}]f(x)$ is called the *residue* of $f(x)$, and is also written as $\mathrm{Res}\, f(x)$.

Everything below hinges on the following simple observation, which is too trivial to need a proof.

Proposition 8.2 *For any formal Laurent series $f(x)$, we have* $\mathrm{Res}\, f'(x) = 0$.

Now the following result describes the residue of the composition of two formal Laurent series.

Theorem 8.3 (Residue Composition Theorem) *Let $f(x)$, $g(x)$ be formal Laurent series with* $\mathrm{val}(g) = r > 0$. *Then*

$$\mathrm{Res}(f(g(x))g'(x)) = r\,\mathrm{Res}(f(x)).$$

Proof It is enough to consider the case where $f(x) = x^n$, since Res is a linear function.

Suppose that $n \neq -1$, so that the right-hand side is zero. Then

$$\mathrm{Res}(g^n(x)g'(x)) = \frac{1}{n+1}\mathrm{Res}\left(\frac{\mathrm{d}}{\mathrm{d}x}g^{n+1}(x)\right) = 0.$$

So consider the case where $n = -1$. Let $g(x) = ax^r h(x)$, where $a \neq 0$ and $h(0) = 1$. Then

$$\begin{aligned} g'(x) &= rax^{r-1}h(x) + ax^r h'(x), \\ \frac{g'(x)}{g(x)} &= \frac{r}{x} + \frac{h'(x)}{h(x)}, \end{aligned}$$

so

$$\operatorname{Res} g'(x)g(x)^{-1} = r = r \operatorname{Res} x^{-1},$$

since $h'(x)/h(x) = (\mathrm{d}/\mathrm{d}x)\log h(x)$, and $\log h(x) = \log(1+k(x))$ is a formal power series since $k(x)$ is a f.p.s. with constant term zero.

It is tempting to say

$$\begin{aligned} g'(x)g(x)^{-1} &= \frac{\mathrm{d}}{\mathrm{d}x}\log g(x) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}(\log a + r\log x + \log h(x)) \\ &= \frac{r}{x} + \frac{\mathrm{d}}{\mathrm{d}x}\log h(x), \end{aligned}$$

but this is not valid; $\log g(x)$ may not exist as a formal Laurent series. Consider this point carefully; an error here would lead to the incorrect conclusion that $\operatorname{Res}(g'(x)/g(x)) = 0$.

From the Residue Composition Theorem, we can prove a more general version of Lagrange Inversion.

Theorem 8.4 (Lagrange Inversion) *Let ϕ be a formal power series with* $\operatorname{val}(\phi) = 0$*. Then the equation*

$$g(x) = x\phi(g(x))$$

has a unique formal power solution $g(x)$. Moreover, for any Laurent series f, we have

$$[x^n]f(g(x)) = \begin{cases} \frac{1}{n}[x^{n-1}](f'(x)\phi(x)^n) & \text{if } n \geq \operatorname{val}(f) \text{ and } n \neq 0, \\ f(0) + \operatorname{Res}(f'(x)\log(\phi(0)^{-1}\phi(x)) & \text{if } n = 0. \end{cases}$$

Proof Let $\Phi(x) = x/\phi(x)$, so that $\Phi(g(x)) = x$ and $\operatorname{val}(\Phi(x)) = 1$. Then g is the inverse function of Φ.

We have

$$\begin{aligned} [x^n]f(g(x)) &= \operatorname{Res} x^{-n-1} f(g(x)) \\ &= \operatorname{Res} \Phi(y)^{-n-1}\Phi'(y) f(y), \end{aligned}$$

where we have made the substitution $x = \Phi(y)$ (so that $y = g(x)$) and used the Residue Composition Theorem.

For $n \neq 0$, we have

$$\begin{aligned}[x^n] f(g(x)) &= -\frac{1}{n}[y^{-1}] f(y) \left(\Phi(y)^{-n}\right)' \\ &= \frac{1}{n}[y^{-1}] f'(y)\Phi(y)^{-n} \\ &= \frac{1}{n}[y^{n-1}] f'(y)\phi^n(y).\end{aligned}$$

Here, in the second line, we have used the fact that

$$\operatorname{Res}(f'(x)g(x)) = -\operatorname{Res}(f(x)g'(x)),$$

a consequence of the fact that $\operatorname{Res}(f(x)g(x))' = 0$); in the third line we use the fact that $\Phi(x) = x/\phi(x)$.

For $n = 0$, we have

$$\begin{aligned}[x^0] f(g(x)) &= [y^0] f(y) - [y^{-1}] f(y)\phi'(y)\phi(y)^{-1} \\ &= f(0) + \operatorname{Res}(f'(y)\log(\phi(y)\phi^{-1}(0)),\end{aligned}$$

using the same principle as before and the fact that

$$(\log(\phi(y)\phi^{-1}(0)))' = \phi'(y)\phi(y)^{-1}.$$

Taking $f(x) = x$ in this result gives the form of Lagrange Inversion quoted earlier.

We proceed to an application, also taken from Goulden and Jackson, of the Residue Composition Theorem.

Example: a binomial identity We use the Residue Composition Theorem to prove that

$$\sum_{k=0}^{n} \binom{2n+1}{2k+1}\binom{j+k}{2n} = \binom{2j}{2n}.$$

We begin with the sum of the odd terms in $(1+x)^{2n+1}$:

$$\sum_{k=0}^{n} \binom{2n+1}{2k+1} x^{2k} = \frac{1}{2x}\left((1+x)^{2n+1} - (1-x)^{2n+1}\right).$$

Call the right-hand side of this equation $f(x)$. Now, if S is the sum that we want to evaluate, then

$$\begin{aligned} S &= [y^{2n}](1+y)^j \sum_{k=0}^{n} \binom{2n+1}{2k+1}(1+y)^k \\ &= \operatorname{Res} y^{-(2n+1)}(1+y)^j f((1+y)^{1/2}). \end{aligned}$$

Now we do the following rather strange thing: make the substitution $y = z^2(z^2-2)$. Then $\mathrm{val}(y(z)) = 2$, and $(1+y)^{1/2} = 1-z^2$. So the Residue Composition Theorem gives

$$\begin{aligned} S &= \operatorname{Res}(z^2-1)^{2j}\left(\frac{1}{(z^2-2)^{2n+1}} - \frac{1}{z^{4n+2}}\right)z \\ &= \operatorname{Res}(z^2-1)^{2j}z^{-(4n+1)} \\ &= [z^{4n}](z^2-1)^{2j} \\ &= \binom{2j}{2n}, \end{aligned}$$

as required. (In the second line we have used the fact that $(z^2-2)^{-(2n+1)}$ is a formal power series and so its residue is zero.)

Chapter 9

Bernoulli, Euler, Maclaurin

We saw in Chapter 1 an asymptotic estimate for $n!$ which began by comparing $\log n! = \sum_{i=1}^{n} \log i$ to $\int_1^n \log x \, \mathrm{d}x$. Obviously the comparison is not exact, but the approximation can often be improved by the Euler–Maclaurin sum formula. This formula involves the somewhat mysterious Bernoulli numbers, which crop up in a wide variety of other situations too.

9.1 Bernoulli numbers

The Bernoulli numbers B_n can be defined by the recurrence relation

$$B_0 = 1, \quad \sum_{k=0}^{n} \binom{n+1}{k} B_k = 0 \text{ for } n \geq 1.$$

Note that we can write the recurrence as

$$\sum_{k=0}^{n+1} \binom{n+1}{k} B_k = B_{n+1},$$

since the term B_{n+1} cancels from this equation (which expresses B_n in terms of earlier terms).

Conway and Guy, in *The Book of Numbers*, have a typically elegant presentation of the Bernoulli numbers. They write this relation as

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

for $n \geq 1$, where B^k is to be interpreted as B_k *after* the left-hand expression has been evaluated using the Binomial Theorem.

Thus,

$$\begin{aligned} B_2+2B_1+1=B_2, &\quad \text{whence} \quad B_1=-\frac{1}{2},\\ B_3+3B_2+3B_1+1=B_3, &\quad \text{whence} \quad B_2=\frac{1}{6}, \end{aligned}$$

and so on. Note that, unlike most of the sequences we have considered before, the Bernoulli numbers are not integers.

Theorem 9.1 *The exponential generating function for the Bernoulli numbers is*

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

Proof Let $F(x)$ be the e.g.f., and consider $F(x)(\exp(x)-1)$. The coefficient of $x^{n+1}/(n+1)!$ is

$$(n+1)!\sum_{k=0}^{n}\left(\frac{B_k}{k!}\right)\left(\frac{1}{(n+1-k)!}\right)=\sum_{k=0}^{n}\binom{n+1}{k}B_k=0$$

for $n\geq 1$. (Note that the sum runs from 0 to n rather than $n+1$ because we subtracted the constant term from the exponential.) The coefficient of x, however, is clearly 1. So the product is x.

Corollary 9.2 $B_n=0$ *for all odd* $n>1$.

Proof

$$F(x)+\frac{x}{2}=\frac{x}{2}\cdot\frac{\exp(x/2)+\exp(-x/2)}{\exp(x/2)-\exp(-x/2)}=\frac{x}{2}\coth\left(\frac{x}{2}\right)$$

which is an even function of x; so the coefficients of the odd powers of x are zero.

Corollary 9.3

$$B_n=\sum_{k=1}^{n}\frac{(-1)^kk!S(n,k)}{k+1}.$$

Proof Let $f(x) = \log(1+x)/x = \sum a_n x^n/n!$, where

$$a_n = \frac{(-1)^n n!}{(n+1)}.$$

By Theorem 2.9, $f(\exp(x)-1) = x/(\exp(x)-1) = \sum B_n x^n/n!$, where

$$B_n = \sum_{k=1}^{n} S(n,k) a_k.$$

One application of the Bernoulli numbers is in *Faulhaber's formula* for the sum of the kth powers of the first n natural numbers. Everyone knows that

$$\begin{aligned} \sum_{i=1}^{n} i &= n(n+1)/2, \\ \sum_{i=1}^{n} i^2 &= n(n+1)(2n+1)/6, \\ \sum_{i=1}^{n} i^3 &= n^2(n+1)^2/4, \end{aligned}$$

but how does the sequence continue?

Theorem 9.4

$$\sum_{i=1}^{n} i^k = \frac{1}{k+1} \sum_{j=0}^{k} \binom{k+1}{j} B_j (n+1)^{k+1-j}.$$

So, for example,

$$\begin{aligned} \sum_{i=1}^{n} i^4 &= \frac{1}{5}\left((n+1)^5 - \frac{5}{2}(n+1)^4 + \frac{5}{3}(n+1)^3 - \frac{1}{6}(n+1)\right) \\ &= n(n+1)(6n^3+9n^2+n-1)/30. \end{aligned}$$

Proof This argument is written out in the shorthand notation of Conway and Guy. Check that you can turn it into a more conventional proof!

We calculate

$$(n+1+B)^{k+1} - (n+B)^{k+1} = \sum_{j=1}^{k+1} \binom{k+1}{j} n^{k-j} ((B+1)^j - B^j).$$

Now $(B+1)^j = B^j$ for all $j \geq 2$, so the only surviving term in this expression is

$$(k+1)n^k((B+1)^1 - B^1) = (k+1)n^k.$$

Thus we have

$$\frac{1}{k+1}((n+1+B)^{k+1} - (n+B)^{k+1}) = n^k,$$

from which by induction we obtain

$$\frac{1}{k+1}((n+1+B)^{k+1} - B^{k+1}) = \sum_{i=1}^{n} i^k.$$

The left-hand side of this expression is

$$\frac{1}{k+1}\sum_{j=0}^{k}\binom{k+1}{j}B_j(n+1)^{k+1-j},$$

as required.

Warning Conway and Guy use a non-standard definition of the Bernoulli numbers, as a result of which they have $B_1 = 1/2$ rather than $-1/2$. As a result, their formulae look a bit different.

How large are the Bernoulli numbers? The generating function $x/(\exp(x)-1)$ has a removable singularity at the origin; apart from this, the nearest singularities are at $\pm 2\pi\mathrm{i}$, and so B_n is about $n!(2\pi)^{-n}$; in fact, it can be shown that

$$|B_n| = \frac{2\,n!\,\zeta(n)}{(2\pi)^n}$$

for n even, where $\zeta(n) = \sum_{k\geq 1} k^{-n}$. Of course, $B_n = 0$ if n is odd and $n > 1$.

Another curious formula for B_n is due to von Staudt and Clausen:

$$B_{2n} = N - \sum_{p-1|2n}\frac{1}{p}$$

for some integer N, where the sum is over the primes p for which $p-1$ divides $2n$.

9.2 Bernoulli polynomials

The *Bernoulli polynomials* $B_n(t)$ are defined by the formula

$$\frac{x\exp(tx)}{\exp(x)-1} = \sum_{n\geq 0} \frac{B_n(t)x^n}{n!}.$$

Proposition 9.5 *The Bernoulli polynomials satisfy the following conditions:*

(a) $B_n(0) = B_n(1) = B_n$ *for* $n \neq 1$, *and* $B_1(0) = -1/2$, $B_1(1) = 1/2$..

(b) $B_n(t+1) - B_n(t) = nt^{n-1}$.

(c) $B_n'(t) = nB_{n-1}(t)$.

(d) $B_n(t) = \displaystyle\sum_{k=0}^{n} \binom{n}{k} B_{n-k}t^k$

Proof All parts are easy exercises. Let $F(t) = x\exp(tx)/(\exp(x)-1)$.

(a) $F(0)$ is the e.g.f. for the regular Bernoulli numbers, and $F(1) = x + F(0)$.
(b) $F(t+1) - F(t) = x\exp(tx)$.
(c) $F'(t) = xF(t)$.
(d) $F(t) = F(0)\exp(xt)$: use the rule for multiplying e.g.f.s.

The first few Bernoulli polynomials are

$$B_0(t) = 1, \qquad B_1(t) = t - \tfrac{1}{2}, \qquad B_2(t) = t^2 - t + \tfrac{1}{6},$$
$$B_3(t) = t^3 - \tfrac{3}{2}t^2 + \tfrac{1}{2}t, \qquad B_4(t) = t^4 - 2t^3 + t^2 - \tfrac{1}{30}.$$

9.3 The Euler–Maclaurin sum formula

Faulhaber's formula gives us an exact value for the sum of the values of a polynomial over the first n natural numbers. The Euler–Maclaurin formula generalises this to arbitrary well-behaved functions; instead of an exact value, we must be content with error estimates, which in some cases enable us to show that we have an asymptotic series.

The Euler–Maclaurin sum formula connects the sum

$$\sum_{i=1}^{n} f(i)$$

with the series

$$\int_1^n f(t)\,\mathrm{d}t + \frac{1}{2}(f(1)+f(n)) + \sum \frac{B_{2i}}{(2i)!}\left(f^{(2i-1)}(n) - f^{(2i-1)}(1)\right),$$

where f is a "sufficiently nice" function.

Here is a precise formulation due to de Bruijn.

Theorem 9.6 *Let f be a real function with continuous $(2k)$th derivative. Let*

$$S_k = \int_1^n f(t)\,\mathrm{d}t + \frac{1}{2}(f(1)+f(n)) + \sum_{i=1}^{k} \frac{B_{2i}}{(2i)!}\left(f^{(2i-1)}(n) - f^{(2i-1)}(1)\right).$$

Then

$$\sum_{i=1}^{n} f(i) = S_k - R_k,$$

where the error term is

$$R_k = \int_1^n f^{(2k)}(t)\frac{B_{2k}(\{t\})}{(2k)!}\,\mathrm{d}t,$$

with $B_{2k}(t)$ the Bernoulli polynomial and $\{t\} = t - \lfloor t \rfloor$ the fractional part of t.

Proof First let g be any function with continuous $(2k)$th derivative on $[0,1]$. We claim that

$$\begin{aligned}\frac{1}{2}(g(0)+g(1)) - \int_0^1 g(t)\,\mathrm{d}t \\ = \sum_{i=1}^{k} \frac{B_{2i}}{(2i)!}\left(g^{(2i-1)}(1) - g^{(2i-1)}(0)\right) - \int_0^1 g^{(2k)}(t)\frac{B_{2k}(t)}{(2k)!}\,\mathrm{d}t.\end{aligned}$$

The proof is by induction: both the start of the induction (at $k=1$) and the inductive step are done by integrating the last term by parts twice, using the fact that $B_n'(t) = nB_{n-1}(t)$ (see Proposition 9.5).

Now the result is obtained by applying this claim successively to the functions $g(x) = f(x+1)$, $g(x) = f(x+2)$, $\ldots$, $g(x) = f(x+n)$, and adding.

If f is a polynomial, then $f^{(2k)}(x) = 0$ for sufficiently large k, and the remainder term vanishes, giving Faulhaber's formula. For other applications, we must estimate the size of the remainder term.

There are various analytic conditions which guarantee a bound on the size of R_k, so that it can be shown that we have an asymptotic series for the sum. I will not give precise conditions here.

Example: Stirling's formula Let $f(x) = \log x$. Then $f^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{x^k}$. We obtain the asymptotic series

$$c + n\log n - n + \frac{1}{2}\log n + \sum \frac{B_{2k}}{2k(2k-1)n^{2k-1}}$$

for

$$\sum_{i=1}^{n} \log i = \log n!\,.$$

The series begins $1/(12n) - 1/(360n^3) + 1/(1260n^5) + \cdots$. Exponentiating term-by-term (using the fact that, if $\log X = \log Y + o(n^{-k})$ then $X = Y(1 + o(n^{-k}))$), we obtain

$$n! \sim \sqrt{2\pi}\,\frac{n^{n+1/2}}{\mathrm{e}^n}\left(1 + \frac{1}{12n} + \frac{1}{288n^2} + \cdots\right).$$

Note in passing that, for fixed n, this asymptotic series is divergent (see our earlier estimate for B_k).

Example: The harmonic series Applying Euler–Maclaurin to $f(x) = 1/x$, we get

$$\sum_{i=1}^{n} \frac{1}{i} \sim \log n + \gamma - \sum \frac{B_k}{kn^k},$$

where the sum begins $1/(2n) - 1/(12n^2) + 1/(120n^4) + \cdots$. Here γ is *Euler's constant*, a somewhat mysterious constant with value approximately $0.5772157\ldots$. Again the series is divergent for fixed n.

Chapter 10

Hayman's Theorem and other tools

A number of non-trivial analytic results have been proved for the purpose of obtaining asymptotic formulae for combinatorially defined numbers. These include theorems of Hayman, Meir and Moon, and Bender. I will not give proofs of these theorems, but treat them as black boxes and give examples to illustrate their use.

10.1 Hayman's Theorem

Hayman's Theorem is an important result on the asymptotic behaviour of the coefficients of certain *entire* functions (i.e., functions which are analytic in the entire complex plane).

The theorem applies only to a special class of such functions, the so-called *H-admissible* or *Hayman-admissible* functions. Rather than attempt to give a general definition of this class, I will state a theorem of Hayman showing that it is closed under certain operations, which suffice to show that any function in which we are interested is H-admissible. See Hayman's paper in the bibliography, or Odlyzko's survey.

Theorem 10.1 *(a) If f is H-admissible and p is a polynomial with real coefficients, then $f+p$ is H-admissible.*

(b) If p is a non-constant polynomial with real coefficients such that $\exp(p(x)) = \sum q_n x^n$ *with $q_n > 0$ for $n \geq n_0$, then* $\exp(p(x))$ *is H-admissible.*

(c) If p is a non-constant real polynomial with leading term positive, and f is H-admissible, then $p(f(x))$ is H-admissible.

(d) If f and g are H-admissible, then $\exp(f(x)$ *and* $f(x)g(x)$ *are H-admissible.*

Corollary 10.2 *The exponential function is H-admissible.*

Now *Hayman’s Theorem* is the following.

Theorem 10.3 *Let $f(x) = \sum_{n\geq 0} f_n x^n$ be H-admissible. Let $a(x) = xf'(x)/f(x)$ and $b(x) = xa'(x)$, and let r_n be the smallest positive root of the equation $a(x) = n$. Then*

$$f_n \sim \frac{1}{\sqrt{2\pi b_n}} f(r_n) r_n^{-n}.$$

Example: Stirling’s formula Take $f(x) = \exp(x)$ (we have noted that this function is admissible), so that $f_n = 1/n!$. Now $a(x) = x = b(x)$, and $r_n = n$. Thus

$$\frac{1}{n!} = \frac{1}{\sqrt{2\pi n}} \mathrm{e}^n n^{-n},$$

which is just Stirling’s formula the other way up!

Example: Bell numbers Let $f(x) = \exp(\exp(x) - 1)$, so that $f_n = B(n)/n!$, where $B(n)$ is the number of partitions of an n-set. This function is H-admissible. Now $a(x) = x\mathrm{e}^x$ and $b(x) = (x + x^2)\mathrm{e}^x$.

The number r_n is the smallest positive solution of $x\mathrm{e}^x = n$. In terms of this, we have

$$\frac{B(n)}{n!} \sim \frac{1}{\sqrt{2\pi n(1+r_n)}} \mathrm{e}^{n/r_n - 1} r_n^{-n},$$

and so by Stirling’s formula,

$$B(n) \sim \frac{1}{\sqrt{1+r_n}} \left(\frac{n}{\mathrm{e}r_n}\right)^n \mathrm{e}^{n/r_n - 1}.$$

Of course, this is not much use without a good estimate for r_n. However, for $n = 100$, the right-hand side is within 0.4% of $B(100)$.

In fact, it can be shown that

$$r_n = \log n - \log\log n + O\left(\frac{\log\log n}{\log n}\right),$$

from which it can be deduced that

$$\log B(n) \sim n\log n - n\log\log n - n.$$

10.2 The theorem of Meir and Moon

The theorem of Meir and Moon (which has been generalised by Bender) gives the asymptotics of the coefficients of a power series defined by Lagrange inversion (compare Chapter 6). Typically we have to find the inverse function of f. Setting $\phi(x) = x/f(x)$, the inverse function g is given by the functional equation $g(y) = y\phi(g(y))$. Replacing y by x and g by f, the theorem is as follows.

Theorem 10.4 *Let $y = f(x) = \sum f_n x^n$ satisfy the equation*

$$y = x\Phi(y),$$

where Φ is analytic in some neighbourhood of the origin, with $\Phi(x) = \sum a_n x^n$. Suppose that the following conditions hold:

(a) $a_0 = 1$ and $a_n \geq 0$ for $n \geq 0$.

(b) $\gcd\{n : a_n > 0\} = 1$.

(c) There is a positive real number α, inside the circle of convergence of Φ, satisfying

$$\alpha\Phi'(\alpha) = \Phi(\alpha).$$

Then

$$f_n \sim Cn^{-3/2}\beta^n,$$

where $C = \sqrt{\Phi(\alpha)/2\pi\Phi''(\alpha)}$ and $\beta = \Phi(\alpha)/\alpha = \Phi'(\alpha)$.

Example: Rooted trees The generating function $y = T^*(x)$ for labelled rooted trees satisfies

$$y = x\exp(y).$$

The exponential function converges everywhere, and the solution of $\alpha\exp(\alpha) = \exp(\alpha)$ is clearly $\alpha = 1$, so that $\beta = \mathrm{e}$ and $C = \sqrt{1/2\pi}$. Hence the number T_n^* of labelled rooted trees on n vertices satisfies

$$\frac{T_n^*}{n!} = \frac{1}{\sqrt{2\pi}} n^{-3/2}\mathrm{e}^n.$$

Since $T_n^* = n^{n-1}$ by Cayley's Theorem, we obtain

$$n! \sim \sqrt{2\pi}\frac{n^{n+1/2}}{\mathrm{e}^n},$$

in other words, Stirling's formula.

10.3 Bender's Theorem

Bender's Theorem generalises the theorem of Meir and Moon by treating a very much more general class of implicitly defined functions. Thus, y will be defined as a function of x by the equation $F(x,y)=0$. In the case of Meir and Moon, we have $F(x,y)=y-x\Phi(y)$.

Theorem 10.5 *Suppose that* $y=f(x)$ *is defined implicitly by the equation* $F(x,y)=0$*, and let* $f(x)=\sum_{n\geq 0}f_nx^n$*. Suppose that there exist real numbers* ξ *and* η *such that*

(a) F *is analytic in a neighbourhood of* (ξ,η)*;*

(b) $F(\xi,\eta)=0$ *and* $F_y(\xi,\eta)=0$*, but* $F_x(\xi,\eta)\neq 0$ *and* $F_{yy}(\xi,\eta)\neq 0$ *(subscripts denote partial derivatives);*

(c) the only solution of $F(x,y)=F_y(x,y)=0$ *with* $|x|\leq\xi$ *and* $|y|\leq\eta$ *is* $(x,y)=(\xi,\eta)$*.*

Then

$$f_n\sim Cn^{-3/2}\xi^{-n},$$

where

$$C=\sqrt{\frac{\xi F_x(\xi,\eta)}{2\pi F_{yy}(\xi,\eta)}}.$$

Example: Wedderburn–Etherington numbers Recall from Chapter 3 that the generating function for these numbers satisfies

$$f(x)=x+\frac{1}{2}(f(x)^2+f(x^2)).$$

Here we have $F(x,y)=y-x-(y^2+g(x))/2$, where $g(x)=f(x^2)$, which we regard as a "known" function (using a truncation of its Taylor series to approximate it).

The equation $F_y(\xi,\eta)=0$ gives us that $\eta=1$; the the equation $F(\xi,\eta)=0$ then gives $g(\xi)=1-2\xi$. This equation can be solved numerically (it is the same one we solved in Chapter 3 to find the radius of convergence of $f(x)$). The remaining conditions of the theorem can then be verified.

We obtain $\xi^{-1} = 2.483\ldots$, and hence

$$f_n \sim Cn^{-3/2}\xi^{-n},$$

where C can also be found numerically if desired.

Exercises

10.1. Let s_n be the number of permutations of $\{1,\ldots,n\}$ which are equal to their inverses. Prove that

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

and use Hayman's Theorem to show that

$$s_n \sim \frac{1}{\sqrt{2}}\left(\frac{n}{\mathrm{e}}\right)^{n/2} \mathrm{e}^{\sqrt{n}-1/4}.$$

Bibliography

Here are the details of the set books for the course and other sources referred to in the text.

E. A. Bender, Asymptotic methods in enumeration, *SIAM Review* **16** (1974), 485–515.

F. Bergeron, G. Labelle and P. Leroux, *Combinatorial Species and Tree-like Structures*, Encyclopedia of Mathematics and its Applications **67**, Cambridge University Press, Cambridge, 1998.

J. H. Conway and R. K. Guy, *The Book of Numbers*, Springer-Verlag, New York, 1996.

M. R. Garey and D. S. Johnson, *Computers and Intractability: An Introduction to the Theory of* NP*-completeness*, W. H. Freeman, San Francisco, 1979.

I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, Wiley-Interscience, New York, 1983.

W. T. Gowers, The two cultures of mathematics, pp. 65–78 in *Mathematics: Frontiers and Perspectives* (ed. V. Arnold, M. Atiyah P. Lax and B. Mazur), American Math. Soc., Providence, 1999.

L. J. Guibas and A. M. Odlyzko, String overlaps, pattern matching, and nontransitive games, *J. Combinatorial Theory* (A) **30** (1981), 183–208.

W. K. Hayman, A generalization of Stirling's formula, *J. Reine Angew. Math.* **196** (1956), 67–95.

A. Joyal, Une theorie combinatoire des séries formelles, *Advances in Math.* **42** (1981), 1–82.

N. Linial and N. Nisan, Approximate inclusion-exclusion, *Combinatorica* **10** (1990), 349–365.

S. Majid, Braided groups, *J. Pure Appl. Algebra* **86** (1993), 187–221; Free braided differential calculus, braided binomial theorem and the braided exponential map, *J. Math. Phys.* **34** (1993), 4843–4856.

A. M. Odlyzko, Asymptotic enumeration methods, pp. 1063–1230 in *Handbook of Combinatorics* (ed. R. L. Graham, M. Grötschel and L. Lovász), North Holland, Amsterdam, 1995.

A. Slomson, *An Introduction to Combinatorics*, Chapman and Hall, London, 1991.

R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, Cambridge University Press, Cambridge, 2000.

H. S. Wilf, The 'Snake Oil' method for proving combinatorial identities, pp. 208–217 in *Surveys in combinatorics* (ed. J. Siemons), London Math. Soc. Lecture Note Series **141**, Cambridge University Press, Cambridge, 1989.

An indispensible website for anyone with an interest in enumeration is Neil Sloane's *Encyclopedia of Integer Sequences*, at
`http://www.research.att.com:80/~njas/sequences/`

Index