Notes on Combinatorics

Peter J. Cameron

Preface: What is Combinatorics?

> Combinatorics, the mathematics of patterns, ..., helps us design computer networks, crack security codes, or solve sudokus
>
> Ursula Martin, Vice-Principal (Science and Engineering),
> Queen Mary, University of London

These notes accompanied the course MAS219, Combinatorics, at Queen Mary, University of London, in the Autumn semester 2007.

It is impossible to define combinatorics, but an approximate description would go like this. We are given the job of arranging certain objects or items according to a specified pattern. Some of the questions that arise include:

- Is the arrangement possible?
- In how many ways can the arrangement be made?
- How do we go about finding such an arrangement?

This is best illustrated by examples.

Example 1: Sudoku You are given a 9×9 grid, divided into nine 3×3 squares. Your job is to put the numbers $1, 2, \ldots, 9$ into the cells of the grid in such a way that each number occurs just once in each row, once in each column, and once in each 3×3 subsquare.

It is not hard to see that the arrangement is indeed possible. A heroic calculation by Bertram Felgenhauer and Frazer Jarvis in 2005 showed that there are

$$6,670,903,752,021,072,936,960$$

different ways of filling the grid.

Now suppose that someone has complicated the problem by writing some numbers into the grid already. In general it may or may not be possible to complete the grid; and even if it is, it may be very difficult to find a solution. Nevertheless, many people around the world enjoy engaging with this combinatorial problem every day.

Example 2: Euler's officers The great mathematician Leonhard Euler asked in 1782:

> *Six different regiments have six officers, each one holding a different rank (of six different ranks altogether). Can these* 36 *officers be arranged in a square formation so that each row and column contains one officer of each rank and one from each regiment?*

Euler conjectured that the answer is "no", and this guess was eventually proved correct in 1900. However Euler also conjectured that the answer is "no" if six is replaced by 10, 14, or any number congruent to 2 mod 4. He was completely wrong about this, but this was not discovered until the 1960s.

Example 3: Kirkman's schoolgirls In 1843, Thomas Kirkman asked:

> *Fifteen schoolgirls go for a walk every day for a week in five rows of three. Is it possible to arrange the walks so that every two girls walk together exactly once during the week?*

This is certainly plausible. Each girl has to walk with fourteen others; every day there are two other girls in her row, so seven days would be the right number for the schedule. However, this does not prove that the arrangement is possible.

In fact, it can be done; Kirkman himself found a schedule satisfying the conditions.

Examples and reality The examples may give you the impression that combinatorics is a collection of charming puzzles of little relevance to our modern technological world. In fact this is completely wrong. The course is not really about applications, but in the digital world this subject is of enormous significance. People (and computers!) spend a lot of time sorting data, sending messages through networks, correcting faulty data or encoding data to keep it safe from unauthorised access, designing better networks, looking for new combinations of atoms to form molecules which will provide us with better drugs, and so on. We need to decide when such a problem has a solution, and to find the solution efficiently.

These notes These notes reflect the contents of the course in 2007. I have added a couple of proofs of major theorems not covered in the course. The notes have been provided with exercises (some of them with worked solutions) and an index.

The recommended textbook for the course was my own book *Combinatorics: Topics, Techniques, Algorithms*, first published in 1994; but rather than following the book I have written everything anew. The course covers roughly the first half of the book; if you enjoyed this, you may want to read more, or to look at my *Notes on counting* on the Web.

I am grateful to Volkan Yildiz who spotted a number of misprints in a preliminary version of the notes.

Further reading Either of the two level 4 courses at Queen Mary can be taken by students who have done the Combinatorics course:

- MAS408: Graphs, Colourings and Design
- MAS439: Enumerative and Asymptotic Combinatorics

I mentioned above my *Notes on counting* which are on the web in the same place as these notes.

Some other books which contain further material (including the recommended course text) are:

- Martin Aigner, *Combinatorial Theory*, Springer, 1979.
- Norman Biggs, *Discrete Mathematics* (2nd edition), Oxford University Press, 2002.
- Peter J. Cameron, *Combinatorics: Topics, Techniques, Algorithms* (2nd edition), Cambridge University Press, 1996.
- J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, 1992.
- Jiri Matousek and Jaroslav Nešetřil, *Invitation to Discrete Mathematics*, Oxford University Press, 1998.

Contents

Chapter 1

Subsets and binomial coefficients

One of the features of combinatorics is that there are usually several different ways to prove something: typically, by a counting argument, or by analytic methods. There are lots of examples below. If two proofs are given, study them both. Combinatorics is about techniques as much as, or even more than, theorems.

1.1 Subsets

Let n be a non-negative integer, and let X be a set with n elements. How many subsets does X have?

Proposition 1.1 *The number of subsets of an n-element set is 2^n.*

First proof We encode subsets by sequences $(e_1, e_2, \ldots, e_n)$, where each e_i is either 0 or 1. There are 2 choices for e_1, 2 choices for e_2, ..., 2 choices for e_n; so altogether 2^n sequences. So we are done if we can establish a bijection between subsets and sequences.

To each subset Y of X, we associate the sequence $(e_1, e_2, \ldots, e_n)$ where

$$e_i = \begin{cases} 1 & \text{if } i \in Y, \\ 0 & \text{if } i \notin Y. \end{cases}$$

It is easy to see that each sequence arises from a subset, and distinct sequences arise from distinct subsets; so the correspondence is a bijection.

Second proof This is a proof by induction. Let $f(n)$ be the number of subsets of $\{1, 2, \ldots, n\}$. We see that $f(0) = 1$ (the empty set has just one subset, namely itself). Also, $f(n+1) = 2f(n)$; for each subset Y of $\{1, 2, \ldots, n\}$ can be extended in two ways to a subset of $\{1, 2, \ldots, n+1\}$: we can choose whether or not to

include $n+1$ in the subset. Now we can easily prove by induction that $f(n) = 2^n$. The induction starts because $f(0) = 1 = 2^0$. For the inductive step, assume that $f(n) = 2^n$; then

$$f(n+1) = 2f(n) = 2 \cdot 2^n = 2^{n+1}.$$

So the induction goes through, and the proof is complete.

1.2 Subsets of fixed size

If n and k are integers satisfying $0 \leq k \leq n$, how many k-element subsets does an n-element set X have?

Define the *binomial coefficient* $\binom{n}{k}$ by

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

(There are k factors in both the numerator and the denominator, the i-th factors being $n-i+1$ and $k-i+1$.)

For $0 \leq k \leq n$, the number of k-element subsets of an n-element set is $\binom{n}{k}$.

Proof We choose k distinct elements of the n-element set X. There are n choices for the first element; $n-1$ choices for the second; ... $n-i+1$ choices for the i-th; ... and $n-k+1$ choices for the k-th. Multiply these numbers together to get that the total number of choices is the numerator of the fraction $\binom{n}{k}$.

This is not the answer, since choosing the same elements in a different order would give the same subset – for example, 1, then 4, then 3 would be the same as 3, then 1, then 4. So we have to divide by the number of different orders in which we could choose the k elements. There are k choices for the first; $k-1$ for the second; ... $k-i+1$ for the i-th; ... and $k-k+1 = 1$ choice (really no choice at all!) for the last. Multiplying these numbers gives the denominator of the fraction. So the result is proved.

It will sometimes be convenient to give a meaning to the symbol $\binom{n}{k}$ even if k is greater than n. We specify:

$$\text{If } k > n, \text{ then } \binom{n}{k} = 0.$$

This is a "reasonable" choice since, if $k > n$, there are no k-element subsets of an n-element set. You should check that our formula for $\binom{n}{k}$ remains correct in this case: if $k > n$, then one of the factors in the numerator is equal to 0.

1.3 Properties of binomial coefficients

1.3.1 Sum of binomial coefficients

The total number of subsets of an n-element set is 2^n. We know the number of subsets of size k, for each value of k: adding these up must give the total. In other words,

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

1.3.2 Binomial coefficients and factorials

Here is an alternative formula for the binomial coefficients. This uses the *factorial function*, defined by

$$n! = n(n-1)(n-2)\cdots 1,$$

the product of all the integers from 1 to n inclusive. Now we have

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

For if we take the definition of the binomial coefficient, and multiply top and bottom by $(n-k)!$, then in the numerator we have the product of all the integers from 1 to n, that is, $n!$; the denominator is $k!\,(n-k)!$.

In order to make this formula valid in the limiting cases $k = 0$ and $k = n$, we have to adopt the convention that $0! = 1$. This may seem strange, but if we want the recurrence $n! = n \cdot (n-1)!$ to hold for $n = 1$, then it is forced upon us! This then correctly gives $\binom{n}{0} = \binom{n}{n} = 1$, and in particular $\binom{0}{0} = 1$.

However, the formula does not work if $k > n$, since then $n - k < 0$ and we cannot define factorials of negative numbers.

1.3.3 A recurrence relation

There is a simple *recurrence relation* for the binomial coefficients, which enables big ones to be calculated from smaller ones by addition:

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

First proof Consider the problem of counting the k-element subsets of an n-element set X, which contains one special element called x.

First we count the sets which contain x. Each of these must have $k-1$ out of the remaining $n-1$ elements. So there are $\binom{n-1}{k-1}$ such sets.

Next we count the sets which do not contain x. Each of these must necessarily have k elements chosen from the $n-1$ elements different from x; so there are $\binom{n-1}{k}$ such sets.

Adding these numbers together gives all the $\binom{n}{k}$ sets.

Second proof We can prove the result by calculation, using our formula:

$$\begin{aligned}\binom{n-1}{k-1}+\binom{n-1}{k} &= \frac{(n-1)!}{(k-1)!\,(n-k)!}+\frac{(n-1)!}{k!\,(n-k-1)!}\\ &= \frac{(n-1)\cdot k}{k!\,(n-k)!}+\frac{(n-1)!\cdot(n-k)}{k!\,(n-k)!}\\ &= \frac{n\cdot(n-1)!}{k!\,(n-k)!}\\ &= \binom{n}{k},\end{aligned}$$

where we have used the facts that $n! = n\cdot(n-1)!$, $k! = k\cdot(k-1)!$, and $(n-k)! = (n-k)\cdot(n-k-1)!$.

I make no secret of the fact that I like the first proof better!

1.3.4 Symmetry

We have

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

For the first proof, we find a bijective correspondence between the k-element sets and the $(n-k)$-element sets in a set of size n; this is easily done by simply matching each set with its complement.

The second proof, using the formula in 2 above, is a simple exercise for the reader.

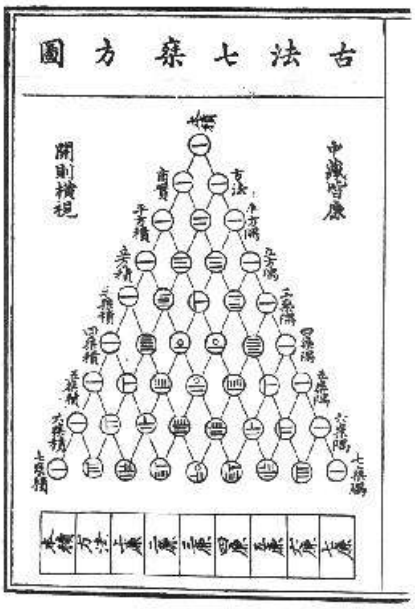
1.3.5 Pascal's Triangle

It is possible to arrange the binomial coefficients in a symmetrical triangular pattern, in which the $(n+1)$-st row contains the $n+1$ numbers $\binom{n}{0}, \ldots, \binom{n}{n}$.

The triangle begins as follows:

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
$$

Although we call this Pascal's Triangle, Pascal was not the first person to write it down. Below is a version due to Chu-Shi-Chieh (Zhu Shijie), taken from work of Yang Hui, in his book *Ssu Yuan Yü Chien*, dated 1303. Jia Xian knew it about 250 years earlier. Other people who knew about it at roughly the same time include Halayudha in India, and Al-Karaji and Omar Khayyam in Iran. We don't know who invented it!



The property in 1.3.4 above shows that the triangle has left-right symmetry. The recurrence relation 1.3.3 shows that each entry of the triangle is the sum of the two entries immediately above it. This gives a very quick method to generate as much of the triangle as required.

1.4 The Binomial Theorem

We now come to the *Binomial Theorem*, a generalisation of the property 1 of the preceding paragraph (put $x = y = 1$ to see this).

Theorem 1.2 (Binomial Theorem)

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

First proof We have

$$(x+y)^n = (x+y)(x+y)\cdots(x+y),$$

where there are n factors on the right-hand side of the equation. If all the brackets are expanded, we get a sum of very many terms; but each term is obtained by choosing x from some of the brackets and y from the remaining ones. If we choose x from k brackets and y from the remaining $n-k$, we obtain a term $x^k y^{n-k}$. So the coefficient of this term is the number of ways we can do this, in other words, the number of choices of k out of the n brackets from which x is selected. This number is $\binom{n}{k}$. So the theorem is proved.

Second proof We prove the theorem by induction on n. For $n = 0$, the left-hand side is $(x+y)^0 = 1$, while the right-hand side has just the single term $k = 0$, which is $\binom{0}{0} x^0 y^0 = 1$. So the induction starts.

Suppose that the Binomial Theorem holds for a value n. Then

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

For $k = m$, second term gives us a contribution $\binom{n}{m} x^m y^{n+1-m}$. What is the contribution to of the first term to the coefficient of $x^m y^{n+1-m}$? To get this term, we must put $k = m-1$, and the coefficient is $\binom{n}{m-1}$.

So the coefficient of $x^m y^{n+1-m}$ in $(x+y)^{n+1}$ is

$$\binom{n}{m-1} + \binom{n}{m} = \binom{n+1}{m},$$

which is just what we require to make the induction work. So the proof is complete.

Sometimes it is convenient to have a one-variable form of the Binomial Theorem. Putting $y = 1$, we obtain

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

1.5 Further properties of binomial coefficients

1.5.1 Even and odd

We know that, for fixed n, the sum of the binomial coefficients $\binom{n}{k}$ over all values of k from 0 to n is 2^n. What if we add them up just for even k, or just for odd k?

For $n > 0$,

$$\sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} = \sum_{i=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2i+1} = 2^{n-1}.$$

Proof Let S_e and S_o be the sums of the even and odd binomial coefficients respectively. Then $S_e + S_o$ is the sum of all the binomial coefficients; in other words,

$$S_e + S_o = 2^n.$$

If we put $x = -1$ in the one-variable Binomial Theorem, we obtain $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = (-1+1)^n = 0$. Now in this sum, the even binomial coefficients have coefficient $+1$ and the odd ones have coefficient -1; so the equation says that

$$S_e - S_o = 0.$$

The two displayed equations show that $S_e = S_o = 2^n/2 = 2^{n-1}$.

1.5.2 Binomial identities

There are a huge number of other equations connecting binomial coefficients. Here is one.

Let m, n, k be positive integers. Then

$$\sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} = \binom{m+n}{k}.$$

(This result is sometimes called the *Vandermonde convolution*.)

First proof Suppose a school class consists of m girls and n boys, and we need to choose a team of k children. In how many ways can this be done? We count the number of teams containing i girls: there are $\binom{m}{i}$ ways to choose the girls, and $\binom{n}{k-i}$ ways of choosing the remaining $k-i$ team members from the n boys. Multiplying these numbers gives us the number of possible teams containing i girls, and summing over i gives the total number of teams. But we know that the total is $\binom{m+n}{k}$.

Second proof Consider the equation

$$(1+x)^m \cdot (1+x)^n = (1+x)^{m+n}.$$

What is the term in x^k? On the right, it is $\binom{m+n}{k}$, by the Binomial Theorem. On the left, we could choose the term x^i from the first factor and x^{k-i} from the second and multiply them. The coefficients of these two terms are $\binom{m}{i}$ and $\binom{n}{k-i}$; so we multiply these numbers, and then sum over i.

Putting $m = n = k$, and noting that $\binom{n}{i} = \binom{n}{n-i}$, the equation reduces to

$$\sum_{i=0}^{n} \binom{n}{i}^2 = \binom{2n}{n}.$$

1.5.3 Sum of sizes of sets

Here are a some further results and proof techniques.

First result: $n\binom{n-1}{k-1} = k\binom{n}{k}$.

First proof From a class of n children, we have to choose a team of k members, and a captain for the team. There are $\binom{n}{k}$ teams, and k choices of a captain for any team; altogether $k\binom{n}{k}$ choices. But we could proceed differently: we could choose the captain first (in n ways), and then the remaining $k-1$ team members from the remaining $n-1$ children (in $\binom{n-1}{k-1}$ ways), giving $k\binom{n-1}{k-1}$ in all.

Second proof

$$\begin{aligned} k\binom{n}{k} &= \frac{k \cdot n!}{k!\,(n-k)!} \\ &= \frac{n \cdot (n-1)!}{(k-1)!\,(n-k)!} \\ &= n\binom{n-1}{k-1}. \end{aligned}$$

Second result: $\sum_{k=1}^{n} k\binom{n}{k} = n \cdot 2^{n-1}$.

First proof

$$\begin{aligned} \sum_{k=1}^{n} k\binom{n}{k} &= n\sum_{k=1}^{n} \binom{n-1}{k-1} \\ &= n\sum_{l=0}^{n-1} \binom{n-1}{l} \\ &= n \cdot 2^{n-1}. \end{aligned}$$

Second proof We have

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

Differentiating gives

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

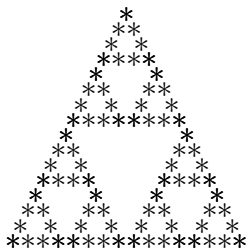
(We omit the $k = 0$ term since it is zero.) Now put $x = 1$ to get the result.

Third proof There are $\binom{n}{k}$ subsets of size k of an n-element set X; so the sum on the left simply adds up the sizes of all these subsets. But we can calculate this sum another way. Pair up each subset A with its complement $X \setminus A$; these two sets contain n elements between them. There are 2^n subsets, and so they fall into $2^n/2 = 2^{n-1}$ pairs. Thus the value of the sum is $n2^{n-1}$.

1.5.4 Congruences

Here is a picture of part of Pascal's triangle. I have put $*$ to mean that the entry is odd, and left a blank if the entry is even. Notice the fractal structure of the

diagram: If we know the triangle formed from the first 2^n rows, we obtain the first 2^{n+1} rows by putting two copies of the triangle side by side below the first one, and leaving the positions in the middle triangle blank.



This is explained by a result called *Lucas' Theorem*:

Theorem 1.3 (Lucas' Theorem) *Let p be a prime number. Write n and k to the base p:*

$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_d p^d, \qquad k = b_0 + b_1 p + b_2 p^2 + \cdots + b_d p^d,$$

where $0 \le a_i, b_i \le p-1$. Then

$$\binom{n}{k} \equiv \prod_{i=0}^{d} \binom{a_i}{b_i} \pmod{p}.$$

In particular, $\binom{n}{k}$ is divisible by p if and only if $a_i < b_i$ for some value of i.

The proof of the theorem is in the next section. You should try to explain how this justifies the fractal shape of the diagram showing the parities in Pascal's triangle.

1.6 Appendix: Proof of Lucas' Theorem

Recall the statement of Lucas' Theorem:

Theorem (Lucas' Theorem) *Let p be a prime number. Write n and k to the base p:*

$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_d p^d, \qquad k = b_0 + b_1 p + b_2 p^2 + \cdots + b_d p^d,$$

where $0 \le a_i, b_i \le p-1$. Then

$$\binom{n}{k} \equiv \prod_{i=0}^{d} \binom{a_i}{b_i} \pmod{p}.$$

The proof comes from the following lemma:

Lemma *Let p be prime, and let $n = cp + a$, $k = dp + b$, with $0 \le a, b \le p-1$. Then*

$$\binom{n}{k} \equiv \binom{c}{d}\binom{a}{b} \pmod{p}.$$

Proof Here is a short proof using the Binomial Theorem. The key is the fact that, if p is prime, then

$$(1+x)^p \equiv 1 + x^p \pmod{p}.$$

For each binomial coefficient $\binom{p}{i}$, for $1 \le i \le p-1$, is a multiple of p, so all intermediate terms in the Binomial Theorem vanish mod p. (We have $\binom{p}{i} = p!/i!(p-i)!$, and p divides the numerator but not the denominator.) Thus (congruence mod p):

$$\begin{aligned}(1+x)^n &= (1+x)^{cp}(1+x)^a \\ &\equiv (1+x^p)^c(1+x)^a \\ &= \sum_{i=0}^{c}\binom{c}{i}x^{pi} \cdot \sum_{j=0}^{a}\binom{a}{j}x^j.\end{aligned}$$

Since $0 \le a, b < p$, the only way to obtain a term in $t^k = t^{dp+b}$ in this expression is to take the term $i = d$ in the first sum and the term $j = b$ in the second; this gives

$$\binom{n}{k} \equiv \binom{c}{d}\binom{a}{b} \pmod{p},$$

as required.

Proof of the theorem The proof is by induction on d. The induction starts with $d = 1$ since, then $n = a_0$, $k = b_0$, and there is nothing to prove.

Suppose that the theorem holds with $d-1$ replacing d. As in the statement of the theorem, let

$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_d p^d, \qquad k = b_0 + b_1 p + b_2 p^2 + \cdots + b_d p^d,$$

where $0 \le a_i, b_i \le p-1$. Put $a = a_0$, $c = a_1 + a_2 p + \cdots + a_d p^{d-1}$, $b = b_0$, $d = b_1 + b_2 p + \cdots + b_d p^{d-1}$. Then $n = cp + a$, $k = dp + b$, and we have (with congruences mod p):

$$\binom{n}{k} \equiv \binom{c}{d}\binom{a}{b} \text{ (by the Lemma)}$$

$$\equiv \left(\prod_{i=1}^{d}\binom{a_i}{b_i}\right)\cdot\binom{a_0}{b_0} \text{ (by the induction hypothesis)}$$
$$= \prod_{i=0}^{d}\binom{a_i}{b_i}.$$

Corollary *With the hypotheses of the theorem,* $\binom{n}{k}$ *is divisible by* p *if and only if* $a_i < b_i$ *for some* i *with* $0 \le i \le d$.

Proof If $a_i < b_i$, then $\binom{a_i}{b_i} = 0$. So one of the factors on the right-hand side of the theorem is zero, whence the product is zero.

If $a_i \ge b_i$, then the binomial coefficient $\binom{a_i}{b_i}$ is not divisible by p (it is non-zero and there are no factors p in the numerator since $a_i \le p-1$. Now a product of d numbers not divisible by p is itself not divisible by p.

Example Let $n = 2^m - 1$. Then all the digits a_i of n in base 2 are equal to 1, so we have $b_i \le a_i$ for any k. This means that every entry in row n of Pascal's triangle is odd.

Exercises

1. Write 1001 as a binomial coefficient $\binom{n}{k}$ with $n \le 20$.

2. If X is a set of 8 elements, then the number of 3-element subsets of X is twice the number of 2-element subsets. Is there any other size of the set X for which this holds?

3. Calculate $\displaystyle\sum_{k=0}^{n} k^2\binom{n}{k}$.

4. Let X be a n-element set. Find a bijection F between the set of k-element subsets of X and the set of all $(n-k)$-element subsets of X. Deduce that

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

5. This exercise extends the result about "summation of even and odd binomial coefficients" in 1.5.1. Similar methods can deal with sums of binomial coefficients where k lies in any fixed congruence class of positive integers.

Let i denote the square root of -1, and note that $1+\mathrm{i} = \sqrt{2}\mathrm{e}^{\pi/4}$. Hence find the real and imaginary parts of $(1+\mathrm{i})^n$ for any natural number n. (You will probably find it convenient to consider the different congruence classes mod 8 separately.)

Expanding $(1+\mathrm{i})^n$ by the Binomial Theorem, find expressions for

$$\sum_{j=0}^{\lfloor (n-t)/4 \rfloor} \binom{n}{4j+t},$$

for $t = 0,1,2,3$, again separating the congruence classes mod 8. (This involves a lot of repetitious work. You should at least do all the calculations for $n \equiv 0 \pmod 8$.)

6. By calculating the coefficient of x^n on the two sides of the identity

$$(1+x)^n \cdot (1-x)^n = (1-x^2)^n,$$

or otherwise, prove that

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

7.

(a) Prove that

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

(b) Prove that, if $n > 2k+1$, then $\binom{n}{k+1} > \binom{n}{k}$; if $n = 2k+1$, then $\binom{n}{k+1} = \binom{n}{k}$; and if $n < 2k+1$, then $\binom{n}{k+1} < \binom{n}{k}$.

(c) Hence show that, for fixed n and $k = 0,1,\ldots,n$, the binomial coefficients increase, then remain constant for one step (if k is odd), then decrease. (Such a sequence is said to be *unimodal*).

(d) Show further that the largest binomial coefficient is $\binom{2m}{m}$ if $n = 2m$ is even, while if $n = 2m+1$ is odd, then $\binom{2m+1}{m}$ and $\binom{2m+1}{m+1}$ are equal largest.

(e) Deduce that, if $n = 2m$, then

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

8.

(a) Show that the binomial coefficient $\binom{2m}{m}$ is divisible by every prime p satisfying $m+1 \leq p \leq 2m$.

(b) Using the estimate on Problem Sheet 1, Question 2, show that the number of primes between $m+1$ and $2m$ is *at most* $\dfrac{2m}{\log_2 m}$.

Remark: This is a weak version of the famous *Prime Number Theorem*, which says that the number of prime numbers p satisfying $1 \leq p \leq n$ is asymptotically $\dfrac{n}{\log n}$.

Chapter 2

Selections and arrangements

2.1 The formulae

We have a hat containing n names, and we are going to draw out k names. In how many ways can we do this?

To answer the question, we have to clarify the strategy a bit. First, do we care about the order in which the names are drawn, or not? Second, when we have drawn a name, do we put it back in the hat and shake it up before the next draw, or do we discard it? The answers to this question correspond to sampling with order significant or not, and with repetition allowed or not allowed. If the order is significant, we have a k-tuple of names; if not, we have a set (if repetition is not allowed), or what might be called a "multiset" if repetition is allowed. We will write multisets in square brackets to distinguish them from sets.

For example, if the names are a, b, and we draw two of them, then

- order important, repetition allowed: there are four possibilities, (a,a), (a,b), (b,a) and (b,b).
- order important, repetition not allowed: there are two possibilities, (a,b) and (b,a).
- order unimportant, repetition allowed: there are three possibilities, $[a,a]$, $[a,b]$, and $[b,b]$. (Choosing a then b is the same as choosing b then a.)
- order unimportant, repetition not allowed: just one possibility, namely $\{a,b\}$.

In general, the numbers of selections are given by the entries in the following table. We use the notation $(n)_k$ for the number $n(n-1)\cdots(n-k+1)$. This is the numerator in our definition of $\binom{n}{k}$, and is often called the *falling factorial*.

	Order significant	Order not significant
Repetition allowed	n^k	$\binom{n+k-1}{k}$
Repetition not allowed	$(n)_k$	$\binom{n}{k}$

Note that the numerator in the top right entry is $n(n+1)\cdots(n+k-1)$, the so-called *rising factorial*.

2.2 Proofs

Order significant, repetition allowed: We get to make k choices, and there are n names to choose at each step. So there are n^k possibilities.

Order significant, repetition not allowed: This time, there are n names to choose at the first step; $n-1$ at the second step (since we discarded the first name after we chose it); $n-2$ at the third step; ... and $n-k+1$ at the k-th step. Multiplying these numbers gives the answer.

Order not significant, repetition not allowed: We simply choose a set with k elements from the n elements in the hat. The number of ways of doing this is $\binom{n}{k}$, by definition. Alternatively choose with order significant, and repetition allowed, and note that each unordered sample has $k!$ different orderings; so the answer is $(n)_k/k!$.

Order not significant, repetition allowed: This case is the most difficult. But note, before we begin, that we cannot just use the argument in the preceding paragraph to get $n^k/k!$. [WHY NOT?]

Step 1: The number of choices of k objects from n, with order not significant and repetition allowed, is equal to the number of ways of choosing n non-negative integers $x_1,\ldots,x_n$ satisfying $x_1+\cdots+x_n=k$. For given the selection, we can let x_i be the number of times that the ith name was selected; clearly $x_1,\ldots,x_n$ satisfy the stated conditions. Conversely, given $x_1,\ldots,x_n$ satisfying the conditions, form a selection in which the i-th name is chosen x_i times.

Thus, for example, suppose that $n=3$ and $k=6$. If the names are a,b,c, then the selection $[a,a,b,b,b,c]$ corresponds to $x_1=2$, $x_2=3$, $x_3=1$.

Step 2: So we have to count the number of choices of non-negative integers $x_1,\ldots,x_n$ with sum k. To do this, take a row of $n+k-1$ cells; choose $n-1$ of them and put markers in them. There are $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$ ways of making this choice. Having made the choice, define $x_1,\ldots,x_n$ as follows:

- Let x_1 be the number of cells before the first marked cell.
- Let x_2 be the number of cells between the first and second marked cell.
- ...
- Let x_{n-1} be the number of cells between the $n-1$-st and n-th marked cell.
- Let x_n be the number of cells after the n-th marked cell.

Then clearly the numbers $x_1,\ldots,x_n$ are non-negative integers; they add up to the number of unmarked cells, which is $(n+k-1)-(n-1)=k$.

Moreover, every way of choosing n non-negative integers adding up to k is represented uniquely by such a marking of $n-1$ out of $n+k-1$ boxes. So the result is proved.

For example, our choice $x_1=2$, $x_2=3$, $x_3=1$ would come from a marking of the following $n-1=2$ out of $n+k-1=8$ boxes:

□□⊠□□□⊠□

To make this clearer, here is a table which gives both steps in the case $n=3$, $k=2$. Let the names in the hat be a,b,c. The first column gives a selection of two names (with repetition allowed and order unimportant). The second gives three numbers adding up to 2. The third gives four boxes with a choice of two of them marked.

aa	$(2,0,0)$	□□⊠⊠
ab	$(1,1,0)$	□⊠□⊠
ac	$(1,0,1)$	□⊠⊠□
bb	$(0,2,0)$	⊠□□⊠
bc	$(0,1,1)$	⊠□⊠□
cc	$(0,0,2)$	⊠⊠□□

2.3 Balls in urns

There is another way to look at the main result of the last section. Suppose that we have n urns, or vases, $U_1,\ldots,U_n$. We have k indistinguishable balls. How many ways can we put the balls in the urns? [Of course this problem can be put into

many disguises. I have k identical sweets. In how many ways can I distribute them to a class of n children?]

If x_i is the number of balls I put into the ith urn [or the number of sweets I give to the ith child], then $x_1,\ldots,x_n$ are non-negative integers which add up to k. So the number of ways of putting the balls into the urns is $\binom{n+k-1}{k}$.

The conditions can be varied in many ways. Suppose, for example, that I have to distribute k balls among n urns as above, but with the requirement that no urn should be empty. This asks that $x_i \geq 1$ for all i. If we define new variables $y_1,\ldots,y_n$ by $y_i = x_i - 1$, then the sum of the ys is $k-n$; so the number of choices of the y's is

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

The simple way to think about this is: Suppose each urn is to be non-empty. Then I first take n balls and put one in each urn. Then I distribute the remaining $k-n$ balls into the urns in any way. This gives the same result as above.

Example How many ways can I distribute 100 sweets to a class of 30 boys and 20 girls, if it is required that each boy has at least one sweet and each girl has at least two sweets?

To solve this, I first give one sweet to each boy and two to each girl, using up $30+2\cdot 20 = 70$ sweets. Then I distribute the remaining 30 sweets among the 50 children, which can be done in $\binom{30+50-1}{30} = \binom{79}{30}$ ways.

2.4 Making words from letters

How many ways can we arrange n distinct objects in order? By the formula in the bottom left of the box, the answer is simply $(n)_n = n!$. Another way of seeing this is as follows. Let $F(n)$ be this number. Then

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

(We take $F(0) = 1$ because there is just one list with no entries, the empty or blank list. To get the second equation, we choose one of the n objects to be first on the list (there are n ways of doing this), and then we have to put the remaining $n-1$ in order after the first one.) Now an easy induction argument shows that $F(n) = n!$ for all n.

Now we make the question a bit harder. How many ways of arranging some (possibly all) of the n objects in a list?

Example How many words can I make from the letters of the word FACETIOUS? (A word is simply a string of letters chosen from those in the word; we do not require that it makes sense in English or any other language. The order of the original letters in the word is irrelevant; a better analogy is that you are playing Scrabble and you have these letters. By convention we include the "empty word", which is the string containing no letters.)

In the case given, the letters are all distinct; this makes life easier, so we start with this case. Suppose that we are given n letters, all different. How many k-letter words can we construct? These words are just selections of k letters from the given n, with order important and repetition not allowed; so the number is $(n)_k = n(n-1)\cdots(n-k+1)$.

So the total number $W(n)$ of words is

$$W(n) = \sum_{k=0}^{n} (n)_k.$$

We can express this another way. Note that $n_{(k)} = n!/(n-k)!$. So

$$W(n) = n! \left(\sum_{k=0}^{n} \frac{1}{(n-k)!} \right) = n! \left(\sum_{m=0}^{n} \frac{1}{m!} \right).$$

Now recall from calculus that

$$\sum_{m=0}^{\infty} \frac{1}{m!} = \mathrm{e}.$$

Inside the brackets of the formula for $W(n)$, we see the sum of the reciprocals of the factorials from 0 to n, in other words, the sum of the first $n+1$ terms of the infinite series. So we see that $W(n)$ is approximately $\mathrm{e} \cdot n!$.

We can be more precise:

$$\begin{aligned}
\mathrm{e} \cdot n! - W(n) &= \sum_{m=n+1}^{\infty} \frac{n!}{m!} \\
&= \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \frac{1}{(n+1)(n+2)(n+3)} + \cdots \\
&< \frac{1}{n+1} + \frac{1}{(n+1)^2} + \frac{1}{(n+1)^3} + \cdots \\
&= \frac{1}{n}.
\end{aligned}$$

(In the last term we summed a geometric series.)

In other words, $\mathrm{e}\,n!$ is bigger than the integer $W(n)$ but smaller than $W(n)+1/n$; so we get $W(n)$ by calculating $\mathrm{e}\cdot n!$ and rounding down to the integer below. So finally we conclude, using the "floor" or "integer part" function, that

$$W(n) = \lfloor \mathrm{e}\cdot n! \rfloor.$$

For the word FACETIOUS, we have $n = 9$, and

$$W(n) = \lfloor \mathrm{e}\cdot 9! \rfloor = 986410.$$

Remark Although $W(n) = \lfloor \mathrm{e}\cdot n! \rfloor$ is a beautiful simple formula, and gives us a very good estimate for the size of $W(n)$, it is not so good for the purpose of calculation. For example, 70! is a number with about 100 digits, so in order to decide whether $W(70)$ is odd or even we would need to know e to 100 places of decimals (at least). For exact calculation it is better to use the formula

$$W(n) = \sum_{k=0}^{n} n_{(k)} = 1 + n + n(n-1) + n(n-1)(n-2) + \cdots$$

We can also find $W(n)$ by a recurrence method. We have

$$W(0) = 1, \qquad W(n) = 1 + nW(n-1) \text{ for } n > 0.$$

(The condition $W(0) = 1$ is because of the empty word. In general, to form a word of from n letters, we choose one letter to go first (in n ways), and make a word from the remaining $n-1$ letters (in $W(n-1)$ ways) to follow it; but we have missed out one word, namely the empty word, so we need to add 1.) An easy induction now gives the formula for $W(n)$.

If the letters we are given contain repetitions, it is more difficult to write down a formula. Here, we will simply do an example.

Example How many words can be made from the letters of SYZYGY?

For the case when we use all the letters, the answer is not too hard. There are 6! ways of arranging the six letters, but any rearrangement of the three Ys will give the same word. So the number of arrangements is $6!/3! = 120$.

If we allow words of arbitrary length, it is a bit more difficult. To solve it, we subdivide the words according to the number of occurrences of the letter Y.

At most one Y We have to make words out of the four letters S, Z, G and Y. This can be done in $W(4) = 1+4+12+24+24 = 65$ ways.

Two Ys Temporarily label the Ys as Y_1 and Y_2 so we can distinguish them. Now we have five letters S, Z, G, Y_1 and Y_2, but we must use the two Ys. Choose some of the other three letters, order all letters including the two Ys in any way, and add up all possibilities; finally divide by 2 since the Ys are really indistinguishable. We get

$$\left(\binom{3}{0}2! + \binom{3}{1}3! + \binom{3}{2}4! + \binom{3}{3}5!\right) \Big/ 2 = 106.$$

All three Ys Similarly the total for this case is

$$\left(\binom{3}{0}3! + \binom{3}{1}4! + \binom{3}{2}5! + \binom{3}{3}6!\right) \Big/ 3! = 193.$$

So the total is $65 + 106 + 193 = 364$.

Exercises

1. (a) How many ordered sequences of length 5 can be made using the elements $\{1,2,3,4,5,6,7\}$ if repetitions are allowed? How many of these contain exactly two of the numbers $1,2,3$? In how many of them do even and odd numbers alternate?

(b) What are the answers to these questions if repetitions are not allowed?

2. How many words can be made using the letters of the word STARTS? How many of these are palindromes (that is, read the same backward as forward)?

3. Let X and Y be sets with $|X| = n$ and $|Y| = m$.

(a) Determine the number of functions f mapping X into Y.

(b) How many of these functions are injections, i.e. one-to-one?

(c) How many of these functions are bijections, i.e. one-to-one and onto?

(d) (much harder) How many of these functions are surjections, i.e. onto?

4. How many permutations of the set $\{1,2,\ldots,n\}$ are there? How many of these are *cyclic permutations*, that is, their cycle decomposition consists of a single cycle of length n?

5. For which values of n is $W(n)$ odd?

Chapter 3

Power series

A lot of combinatorics is about sequences of numbers:

$$(a_0, a_1, a_2, \ldots)$$

We'll see such sequences as

$$(1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots)$$

(Fibonacci numbers), or

$$(1, 1, 2, 6, 24, 120, 720, 5040, \ldots)$$

(factorials). A very useful device to represent such a sequence of numbers is to take the numbers to be the coefficient in a power series

$$\sum_{n \geq 0} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$$

We call this power series the *generating series* or *generating function* for the sequence of numbers.

In this chapter we look at power series and some of their uses in combinatorics.

Example We saw that the number of subsets of an n-element set is 2^n. This gives us a sequence of numbers, namely

$$(2^0 = 1, 2^1 = 2, 2^2 = 4, 2^3 = 8, \ldots)$$

whose generating function is

$$\sum_{n \geq 0} 2^n x^n = \frac{1}{1 - 2x}$$

using the formula for the sum of a geometric series.

3.1 Power series

You've met power series in calculus, and maybe in analysis also. So how do they compare with combinatorics:

First, the good news. We are not doing calculus here, so we don't have to worry whether the sequences converge or not. For us, a power series is just a bookkeeping device, to wrap up infinitely many terms into a single mathematical object. For example, if our sequence is the factorials above, then the power series is

$$\sum_{n\ge 0} n!\,x^n$$

and if you remember the ratio test from calculus, you should be able to show that this series never converges unless $x = 0$. (The ratio of successive terms is $(n+1)!\,x^{n+1}/n!\,x^n = (n+1)x$, which tends to infinity as $n \to \infty$.) But this power series might still be useful!

Second, the good news. If a power series does converge, and if you know something about the properties of the function $A(x)$ it defines, then you can use those properties in combinatorics also! We'll see some examples later. In the example above, the sum of the series is $1/(1-2x)$; the series converges if $|x| < 1/2$.

We denote the set of all power series with integer coefficients by $\mathbb{Z}[[x]]$. This should remind you of the notation $\mathbb{Z}[x]$ for the set of polynomials with integer coefficients; power series are very similar to polynomials, but can have infinitely many coefficients. Similarly, if we want the coefficients to be real numbers, we write $\mathbb{R}[[x]]$, with similar modifications for the other number systems.

3.2 Operations on power series

There are various operations that can be done to power series. If you studied Algebra, you have met the idea of a *ring*; the first two operations below (addition and multiplication) make $R[[x]]$ into a ring, for any ring R (though we won't stop to prove this).

Addition We add two power series term by term:

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

Multiplication We multiply power series in the same way as we multiply polynomials. To get a term in x^n in the product, we multiply the term in x^k in the first factor by the term in x^{n-k} in the second, and sum over all values of k from 1 to n. Thus

$$\begin{aligned}\left(\sum_{n\geq 0} a_n x^n\right)\cdot\left(\sum_{n\geq 0} b_n x^n\right) &= \sum_{n\geq 0} c_n x^n,\\ \text{where}\quad c_n &= \sum_{k=0}^{n} a_k b_{n-k}.\end{aligned}$$

Substitution Let $A(x)=\sum_{n\geq 0} a_n x^n$ and $B(x)=\sum_{n\geq 0} b_n x^n$. Suppose that $a_0=0$. Then we can substitute $A(x)$ for x in the second series:

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

where $A(x)^n$ is calculated using the multiplication rule.

Why do we need the constant term of $A(x)$ to be zero? Consider the constant term of the series $B(A(x))$. It would be $b_0+b_1a_0+b_2a_0^2+\cdots$, and we would have an infinite series of *numbers*, and would have to worry about convergence. But if $a_0=0$, then the smallest power of x occurring in $A(x)^n$ is at least x^n; so when we come to calculate the coefficient of x^n in $B(A(x))$, we only have to consider finitely many terms $b_kA(x)^k$ for $0\leq k\leq n$. In other words, we only need finitely many additions and multiplications to work out any term.

Differentiation We can also differentiate power series. If

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

then

$$\frac{\mathrm{d}}{\mathrm{d}x}A(x)=\sum_{n\geq 1} n a_n x^{n-1}=\sum_{m\geq 0}(m+1)a_{m+1}x^n.$$

Notice what has happened here. The term $n=0$ is zero, so we leave it out in the first step; then we use a new summation variable $m=n-1$, so that as n runs from 1 to infinity, m runs from 0 to infinity.

3.3 The Binomial Theorem

We saw the Binomial Theorem, a formula for $(1+x)^n$ for positive integers n. Here is a generalisation of it, first proved by Isaac Newton.

We need to generalise the definition of binomial coefficients first. Let a be any number, positive or negative, rational or irrational, real or complex. Let k be a natural number (a positive integer or zero). Define

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

This has the properties

- if a is a natural number, then $\binom{a}{k} = 0$ for $k \geq n$;
- otherwise, $\binom{a}{k} \neq 0$ for all a.

For the only way we can have $\binom{a}{k} = 0$ is for one of the factors in the numerator to be zero, that is, $a - i = 0$ (that is, $a = i$) for some $i \leq k-1$.

Now we have:

Theorem 3.1 (The Binomial Theorem) *For any complex number a,*

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

There are two ways to interpret this theorem. In terms of calculus: the series on the right converges for $|x| < 1$, and its sum is $(1+x)^a$. Second, in terms of combinatorics: The usual rules of exponents hold. A "calculus proof" of the Binomial Theorem (without all the tricky details about convergence) is given in an appendix.

Example 1 The first law of exponents says that

$$(1+x)^a(1+x)^b = (1+x)^{a+b}.$$

By the Binomial Theorem,

$$\left(\sum_{k\geq 0} \binom{a}{k} x^k\right) \cdot \left(\sum_{k\geq 0} \binom{b}{k} x^k\right) = \sum_{k\geq 0} \binom{a+b}{k} x^k.$$

Now by the rule for multiplication of power series,

$$\sum_{i=0}^{k} \binom{a}{i}\binom{b}{k-i} = \binom{a+b}{k}.$$

This is the Vandermonde convolution. We saw it for natural numbers a and b in Section 1.5.2; but now we know that it holds for any a and b at all.

Example 2 We get some interesting examples by choosing exponents which are not natural numbers.

Case $a = -1$ We have

$$\binom{-1}{k} = \frac{(-1)(-2)\cdots(-k)}{k(k-1)\cdots 1} = (-1)^k,$$

so

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

so we have the formula for the sum of a geometric series. We already used this in calculating the generating function for the powers of 2.

Case $a = -1/2$ We have

$$\begin{aligned}
\binom{-1/2}{k} &= \frac{(-1/2)(-3/2)\cdots(-(2k-1)/2)}{k(k-1)\cdots 1} \\
&= \left(\frac{-1}{2}\right)^k \frac{(2k-1)(2k-3)\cdots 1}{k(k-1)\cdots 1} \\
&= \left(\frac{-1}{4}\right)^k \frac{1}{k!}\frac{2k(2k-1)\cdots 1}{k!} \\
&= \left(\frac{-1}{4}\right)^k \binom{2k}{k},
\end{aligned}$$

where we have used the fact that $2k(2k-2)\cdots 2 = 2^k k!$. Thus

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

So the generating function for the *central binomial coefficients* $\binom{2k}{k}$ is $1/\sqrt{1-4x}$.

Exampe 2, continued We can use what we just learned to prove the following identity for the central binomial coefficients:

$$\sum_{k=0}^{n}\binom{2k}{k}\binom{2(n-k)}{n-k} = 4^n.$$

Proof We start from the identity

$$(1-4x)^{-1/2}(1-4x)^{-1/2} = (1-4x)^{-1}.$$

Now the coefficient of x^n on the left is obtained by taking the coefficient of x^k in the first factor $(1-4x)^{-1/2}$, multiplying by the coefficient of x^{n-k} in the second factor, and summing over k from 0 to n. This gives precisely the left-hand side of the result we are proving.

On the right,

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

so the coefficient of x^n is 4^n, and we are done.

Example 3 Here is a simple example of the use of power series to solve a recurrence. We will have more complicated examples later.

Suppose that a sequence of numbers a_0, a_1, a_2 satisfy $a_0 = 1$ and $a_n = 2a_{n-1}$ for $n \geq 1$. Of course it is clear that these numbers are the powers of 2. But let us see this another way. The generating function is

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

(Check that you can follow all these steps. In the third step we have used a new summation variable $m = n-1$.) This equation can be rearranged to give

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

Now if two power series are equal then their coefficients must be the same; so we have $a_n = 2^n$ for all n.

3.4 Other power series

Apart from the Binomial Theorem, there are a couple of other famous power series which crop up from time to time:

The exponential function In calculus this is usually written as e^x. I will usually write it as $\exp(x)$; this means the same thing. The power series is

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

The most important properties are

- $\dfrac{\mathrm{d}}{\mathrm{d}x}\exp(x) = \exp(x)$. This is easy to prove from the power series since

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\frac{x^n}{n!} = \frac{x^{n-1}}{(n-1)!}.$$

- $\exp(x+y) = \exp(x)\exp(y)$. (We prove this below.)

The logarithm function The function $\log(x)$ is not defined at $x = 0$ so we cannot write it as a power series. Instead, we have

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

If we differentiate term by term we get

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

The logarithm is the inverse of the exponential:

$$\exp(\log(1+x)) = 1+x, \qquad \log(\exp(x)) = x.$$

(Remember that we can substitute one power series in another if the first one has constant term zero. This is OK for the first result above. In the second case, it is really $\log(1+y)$, where $y = \exp(x)-1$, which does indeed have constant term zero.)

Example Consider the equation $\exp(x+y) = \exp(x)\exp(y)$. The left hand side is

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

(In the second line we used a dummy variable $\ell = n-k$. We have to check the ranges of summation: n taking all values and k running from 0 to n is the same as k and ℓ independently taking all non-negative values.)

We could have reversed the procedure and derived the Binomial Theorem from the property of the exponential function.

Actually there is a lot of very interesting combinatorics hidden in the power series for the exponential and logarithm functions. If you are interested in this, see my *Notes on Counting* on the Web.

3.4.1 Appendix: Proof of the Binomial Theorem

This proof is a bit of a cheat, since all the hard work is in the calculus.

Suppose we have a power series $\sum_{k\geq 0} a_k x^k$ whose sum is a known function $f(x)$. How do we work out the coefficients a_k? If we differentiate the series n times, we get

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x) = \sum_{k\geq n} a_k k(k-1)\cdots(k-n+1)x^{k-n}.$$

(We start the sum at $k=n$ because the n-th derivative of any smaller power of x is zero.) Then if we put $x=0$, we find

$$\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x)\right]_{x=0} = n!\, a_n,$$

so that $a_n = [(\mathrm{d}^n/\mathrm{d}x^n) f(x)]_{x=0}/n!$.

Taking $f(x) = (1+x)^a$, when we differentiate n times we get

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(1+x)^a = a(a-1)\cdots(a-n+1)(1+x)^{a-n}.$$

Putting $x=0$, we get

$$\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}(1+x)^a\right]_{x=0} = a(a-1)\cdots(a-n+1).$$

So the coefficient of x^n in the power series for $(1+x)^a$ is

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

so that $(1+x)^a = \sum_{n\geq 0} \binom{a}{n} x^n$.

Exercises

1. The purpose of this exercise is to show you that, even when a power series fails to converge, algebraic manipulations on it can still give us something interesting.

(a) Let π be a permutation of the set $\{1,\ldots,n\}$. We say that π is *decomposable* if there is a number k, with $1 \leq k \leq n-1$, such that π maps the numbers $1,\ldots,k$ to themselves. If no such k exists then π is *indecomposable*.

There are $n!$ permutations of the set $\{1,\ldots,n\}$. Suppose that $g(n)$ of them are indecomposable. (By convention we take $0! = 1$ but we do not define $g(0)$.)

For any permutation π, let k be the smallest number such that π maps $1,\ldots,k$ to themselves (so that $k = n$ if π is indecomposable). Show that there are $g(k)(n-k)!$ permutations with any given value of k. Hence show that

$$\sum_{k=1}^{n} g(k)(n-k)! = n!.$$

Now let $F(x) = \sum_{n\geq 0} n!\,x^n$ and $G(x) = \sum_{n\geq 1} g(n)x^n$ be the generating functions for the factorial numbers and the numbers $g(n)$ respectively. Note that $G(x)$ has constant term zero since we start at 1. Prove that

$$F(x)(1-G(x)) = 1.$$

Note that this equation makes sense even though the power series do not converge for any non-zero value of x.

Chapter 4

Recurrence relations

Recurrence relations are a very powerful method of calculating combinatorial numbers. But there are not many general methods for dealing with them, so mostly we will just look at a few important examples. The main idea is that we can turn a recurrence relation for a sequence of numbers into an equation (algebraic or differential) for the generating function.

4.1 Fibonacci numbers

Leonardo Fibonacci was an Italian mathematician of the 13th century. His most important work was the introduction of the Arabic numerals $0,1,2,3,4,5,6,7,8,9$ to Europe. In order to show how much easier it is to calculate with these than with the Roman numerals previously used, he posed the following problem as an exercise in his book *Liber Abaci* (The Book of Calculation):

> A pair of rabbits do not breed in their first month of life, but at the end of the second and every subsequent month they produce one pair of offspring. If I acquire a new-born pair of rabbits at the beginning of the year, how many pairs of rabbits will I have at the end of the year?

Under these conditions, the number of pairs of rabbits after n months is called the nth *Fibonacci number* F_n. How do we calculate these numbers?

First, we have

$$F_0 = 1, \qquad F_1 = 1.$$

For we are given that we have one pair of rabbits at the start of month 0, and they do not produce offspring in month 1.

Next,

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

To show this, let G_n be the number of pairs of rabbits which are old enough to breed at the end of month n. Now by the conditions of the problem, we have $G_n = F_{n-2}$ (since the rabbits breeding in month n are all those born in month $n-2$ or earlier). Also, $F_n - F_{n-1} = G_n$, since G_n pairs are born in month n and are those contributing to F_n but not to F_{n-1}. Eliminating G_n from these two equations gives the result.

So the answer to Fibonacci's exercise can be found by a dozen additions, a simple job using Arabic numerals.

Fibonacci did not invent these numbers, which had been known to Indian mathematicians including Pingala, Virahanka and Hemachandra for nearly 1500 years when he wrote his book.

The condition $F_n = F_{n-1} + F_{n-2}$ is an example of a *recurrence relation*. This is a relation which enables any term of the sequence to be calculated if the earlier terms are known. In this case we only need to know the two preceding terms. Usually, a recurrence relation needs to be supplemented with initial conditions, telling us how the sequence starts. In this case the recurrence relation only applies for $n \geq 2$, so we need to be given the values of F_0 and F_1 separately.

In the next section, we will solve this recurrence relation to find an explicit formula for the nth Fibonacci number. First, though, we give a couple more counting problems for which the Fibonacci numbers are the solution.

Example I have a staircase with n steps. At a single stride, I can go up either one or two of the steps. In how many different ways can I walk up the staircase?

Let a_n be this number. Then $a_0 = 1$ (since if there are no steps, then there is only one way to do nothing!) and $a_1 = 1$ (obviously). We claim that $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$. For let S be the set of all ways of walking up the steps. The last step we use before we reach the top is either number $n-1$ or number $n-2$ (since we ascend either one or two steps in the last stride); so let S_1 be the set of ways in which the penultimate step is number $n-1$, and S_2 those in which it is number $n-2$. Then S_1 and S_2 are disjoint and have union S. Moreover, clearly we have $|S_1| = a_{n-1}$ while $|S_2| = a_{n-2}$, and $|S| = a_n$. So the recurrence relation holds. Now a straightforward induction shows that $a_n = F_n$ for all natural numbers n.

This representation of the Fibonacci numbers was discussed by Virahanka in the 6th century, in connection with Sanskrit poetry. A vowel in Sanskrit can be long or short. If we assume that a long vowel is twice as long as a short vowel, in how many ways can we make a line of poetry of length n out of long and short vowels? Clearly this is the same problem, and the answer is the nth Fibonacci number F_n.

From this theorem we get a curious formula for F_n:

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

For another way of stating our result is that the number of ways of writing n as an ordered sum of ones and twos is F_n. Now we can count these expressions another way. Suppose that we have k twos in the sum. Then we must have $n-2k$ ones, so there are $n-k$ terms altogether (and we see that $k \leq \lfloor n/2 \rfloor$). So the number of expressions with k twos is the number of selections of k elements from $n-k$ (the positions in the sequence where the 2s occur), of which there are $\binom{n-k}{k}$. Summing over k gives the result.

For example, when $n = 4$, we have

- $\binom{4}{0} = 1$ (corresponding to $1+1+1+1$);
- $\binom{3}{1} = 3$ (corresponding to $2+1+1$, $1+2+1$ and $1+1+2$);
- $\binom{2}{2} = 1$ (corresponding to $2+2$).

Summing, we have $F_4 = 5$.

Example How many sequences of length n are there consistsing of zeros and ones with no two consecutive ones? (Call such a sequence *admissible*.) Let b_n be this number. Clearly $b_0 = 1$ (only the empty sequence), and $b_1 = 2$ (the sequences 0 and 1 are both admissible).

Partition the set T of all admissible sequences into two subsets T_0 and T_1, where T_0 is the set of sequences ending in 0, and T_1 is the set of sequences ending in 1. Now given any admissible sequence of length $n-1$, we can add a zero to it to get an admissible sequence of length n; so $|T_0| = b_{n-1}$. But we may only add a 1 to an admissible sequence if it ends in zero; so $|T_1|$ is the number of admissible sequences of length $n-1$ ending in zero, which by the preceding argument is b_{n-2}. Thus, $b_n = b_{n-1} + b_{n-2}$.

We have the same recurrence relation as for the Fibonacci numbers, but different initial conditions. However, we have $b_n = F_{n+1}$ for all n. The proof is by induction. We have $b_0 = 1 = F_1$, $b_1 = 2 = F_2$, and for $n \geq 2$,

$$b_n = b_{n-1} + b_{n-2} = F_n + F_{n-1} = F_{n+1}.$$

4.2 Linear recurrences with constant coefficients

In this section we will find a formula for the nth Fibonacci number. The two methods we use can be extended to a wider class of recurrence relations.

Method 1 We are trying to solve the recurrence relation with initial conditions

$$F_0 = 1, \quad F_1 = 1, \qquad F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 2.$$

We begin by observing that there is a unique solution. For F_0 and F_1 are given, and then the recurrence determines $F_2, F_3, \ldots$. (This is really an argument by induction!) So, if we can find by any method at all a solution, then we know it is the unique solution.

We will consider just the recurrence relation $a_n = a_{n-1} + a_{n-2}$, and worry about the initial conditions later. The next observation that we make is that the recurrence relation is *linear*. That means that, if two sequences (a_n) and (b_n) satisfy it, then so does any linear combination (c_n) with $c_n = pa_n + qb_n$ for any numbers p and q. So we concentrate on finding specific solutions.

We *try* a solution of the form $a_n = \alpha^n$ for some number α. [Why? One answer is that it works, as we will see. A better answer is that, if you consider a "one-term recurrence relation" like $a_n = \alpha a_{n-1}$, it is obvious that there will be a solution $a_n = \alpha^n$.]

Now $a_n = \alpha^n$ will satisfy the recurrence relation if and only if

$$\alpha^n = \alpha^{n-1} + \alpha^{n-2} \text{ for } n \geq 2.$$

This will be the case if and only if $\alpha^2 = \alpha + 1$.

The quadratic equation $x^2 = x + 1$ has two solutions

$$\alpha = \frac{1+\sqrt{5}}{2}, \qquad \beta = \frac{1-\sqrt{5}}{2}.$$

So $a_n = \alpha^n$ and $a_n = \beta^n$ both satisfy our recurrence relation, and by the linearity principle, so does

$$a_n = p\alpha^n + q\beta^n$$

for any p and q.

Finally, we try to choose p and q such that this solution also satisfies the initial conditions $a_0 = a_1 = 1$. This gives us two equations

$$\begin{array}{rcrcl} p & + & q & = & 1, \\ p\alpha & + & q\beta & = & 1. \end{array}$$

Solving these equations we find that

$$p = \frac{1+\sqrt{5}}{2\sqrt{5}}, \qquad q = \frac{-1+\sqrt{5}}{2\sqrt{5}}.$$

So we conclude that

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}.$$

Now this is not a very good formula for calculation, since we need to know $\sqrt{5}$ to a high degree of accuracy to use it. But it has one advantage. We have $\alpha = 1.618\ldots$ and $\beta = -0.618\ldots$. So $\alpha > 1$ while $|\beta| < 1$. This means that the nth power of α grows exponentially, while the nth power of β tends exponentially to zero. So we get a very good approximation

$$F_n \approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}.$$

The number $(1+\sqrt{5})/2$ is called the *golden ratio*. It has a long history in Western art, music and botany.

Method 2 This method works with the generating function $f(x) = \sum_{n\geq 0} F_n x^n$. Recall our conditions:

$$F_0 = 1, \quad F_1 = 1, \qquad F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 2.$$

We claim that

$$(1 - x - x^2) f(x) = 1.$$

For the constant term in $(1-x-x^2)f(x)$ is $F_0 = 1$, and the coefficient of x is $F_1 - F_0 = 0$; while, for $n \geq 2$, the coefficient of x^n is $F_n - F_{n-1} - F_{n-2} = 0$.

So

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

To proceed we use the method of *partial fractions*. First, we factorise the denominator:

$$1 - x - x^2 = (1-\alpha x)(1-\beta x),$$

where α and β are as in the last section. Then we write

$$\frac{1}{1-x-x^2} = \frac{p}{1-\alpha x} + \frac{q}{1-\beta x}.$$

Multiplying up by the denominator,

$$1 = p(1-\beta x) + q(1-\alpha x),$$

giving two equations

$$p+q = 1, \qquad p\beta + q\alpha = 0.$$

Solving these two equations gives the same values of p and q as we found in the last section.

Finally, we use the fact that

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

and similarly for β, using the formula for a geometric series. So we have

$$f(x) = \sum_{n\geq 0} (p\alpha^n + q\beta^n) x^n,$$

so that $F_n = p\alpha^n + q\beta^n$, exactly as we found by the other method.

The methods used here work more generally. A *kth order linear recurrence with constant coefficients* is a relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k},$$

for fixed constants $c_1, \ldots, c_k$, connecting the terms of a sequence (a_n). In order to specify the terms completely, we need to specify the values of $a_0, a_1, \ldots, a_{k-1}$; then the recurrence expresses all later terms uniquely.

We can use either of the above methods. The numbers α and β earlier are replaced by the solutions $\alpha_1, \ldots, \alpha^k$ of the equation

$$x^k = c_1 x^{k-1} + c_2 x^{k-2} + \cdots + c_k.$$

There is one complication. If this polynomial has repeated roots, we don't find enough solutions of the form $a_n = \alpha^n$ to use the first method. Instead, if the number α is an r-fold root, then the r functions

$$a_n = \alpha^n, n\alpha^n, \ldots, n^{r-1}\alpha^n$$

all turn out to be solutions.

Example The number $f(n)$ steps required to solve the "Chinese rings puzzle" with n rings satisfies the recurrence

$$f(1) = 1, \quad f(2) = 2, \qquad f(n) = f(n-1) + 2f(n-2) + 1 \text{ for } n \geq 3.$$

There is an awkward 1 on the right, which can be removed by putting $g(n) = f(n) + 1/2$; we find that

$$g(1) = 3/2, \quad g(2) = 5/2, \qquad g(n) = g(n-1) + 2g(n-2) \text{ for } n \geq 3.$$

Now the equation for α and β is $x^2 = x + 2$, with solutions $\alpha = 2$, $\beta = -1$. So

$$g(n) = p \cdot 2^n + q(-1)^n.$$

The initial values give

$$2p - q = 3/2, \qquad 4p + q = 5/2,$$

with solution $p = 2/3$, $q = -1/6$. So the solution to the original problem is

$$f(n) = (2/3)2^n - (1/6)(-1)^n - (1/2).$$

4.3 Linear recurrences with non-constant coefficients

The next complication is that a recurrence relation can have coefficients which are not constants but functions of n. The simplest example is the recurrence for the factorial numbers $n!$:

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

A closely related example concerns $W(n)$, the number of words that can be formed of n distinct letters. We saw that

$$W(0) = 1, \qquad W(n) = 1 + nW(n-1) \text{ for } n \geq 1.$$

There are general methods for solving recurrences of this type (if the coefficients are polynomials in n) in terms of so-called *hypergeometric functions*. Here I will simply discuss one example, which illustrates another technique.

Derangements A *permutation* of $\{1,\ldots,n\}$ is a bijective function from the set $\{1,\ldots,n\}$ to itself. The number of permutations is $n!$. We will say more about permutations later. Here we look at a special type of permutation.

A *derangement* is a permutation which leaves no point fixed. That is, the permutation π is a derangement if $\pi(i) \neq i$ for $i = 1,\ldots,n$. How many derangements are there?

Let this number be $d(n)$. Trivially $d(0) = 1$, since there are no points to fix. Also, $d(1) = 0$ (there is only one permutation of $\{1\}$, and it obviously fixes the point 1), and $d(2) = 1$ (the unique derangement being the permutation which swaps 1 and 2).

We show that the following recurrence relation holds:

$$d(n) = (n-1)(d(n-1)+d(n-2)) \text{ for } n \geq 2.$$

To see this, consider derangements π of $\{1,\ldots,n\}$. Since the point n is not fixed, we must have $\pi(n) = i$ for some i, with $1 \leq i \leq n-1$. Now by symmetry, the number of derangements satisfying $\pi(n) = i$ is independent of i; so we only have to count the derangements with a fixed value of i, and multiply the number of these by $n-1$.

We divide the derangements satisfying $\pi(n) = i$ into two types:

Type 1: Those with $\pi(i) = n$, that is, swapping i with n. Such a permutation π is a derangement of the $n-2$ points different from i and n. There are $d(n-2)$ such derangements; each of them can be extended to the whole set so that it swaps i and n.

Type 2: Those with $\pi(i) \neq n$. Then $\pi(j) = n$ for some $j \neq i$. Now π maps $j \mapsto n \mapsto i$. We can take a short-cut by going straight from j to i, giving a permutation of $\{1,\ldots,n-1\}$; this permutation is a derangement, so there are $d(n-1)$ choices. Given any derangement of $\{1,\ldots,n-1\}$, we can extend it to $\{1,\ldots,n\}$ by interpolating n just before i.

So the number of derangements mapping n to i is $d(n-1)+d(n-2)$, and the total number of derangements is $(n-1)(d(n-1)+d(n-2))$.

It is possible to use this recurrence to find a formula for the numbers $d(n)$, or to find a generating function for them; and there is a completely different approach using the Inclusion–Exclusion Principle that I will discuss later in the notes. Here I will merely quote the formula:

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

Let us look at this formula. It is very similar to the formula for $W(n)$, and can be analysed similarly. Note that, from the exponential series, we see that

$$\mathrm{e}^{-1} = \sum_{k\geq 0} \frac{(-1)^k}{k!},$$

so

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

Just as for $W(n)$, the right-hand side of this equation has modulus smaller than 1, indeed, smaller than $1/2$ if $n \geq 1$. We conclude that

$d(n)$ is the integer nearest to $n!/\mathrm{e}$

the difference being alternately positive and negative.

This has an interesting interpretation. Suppose that n people go to the theatre, and leave their hats at the cloakroom. After the performance, when they go to collect their hats, the cloakroom attendant gives them out at random. Then the probability that nobody gets his or her correct hat is very close to $1/\mathrm{e} = 0.367879\ldots$. For we can regard the allocation of the hats as a random permutation of the correct allocation; and the event that nobody gets the correct hat is just that the random permutation is a derangement.

4.4 Non-linear recurrences

A recurrence relation is really any expression, however complicated, which expresses the nth term of a sequence in terms of smaller terms. There is no general method for solving an arbitrary recurrence relation. Here I will just consider one important example.

Catalan numbers The Catalan numbers appear as the solution of many different counting problems. For example, suppose that we have to calculate a product

$$x_1 \cdot x_2 \cdots x_n.$$

If we can only multiply two factors at a time, we have to put in brackets to make the expression well-defined. how many ways can we bracket such a product? Let C_n be this number. If $n = 1$, no brackets are needed, and $C_1 = 1$. If $n \geq 2$, then we can bracket together the first k terms in C_k ways, and the last $n-k$ terms in C_{n-k} ways, and finally multiply together these two expressions and sum over k; so

$$C_n = \sum_{k=1}^{n-1} C_k C_{n-k}.$$

For example, there are five bracketings for $n = 4$, namely

$$((ab)c)d, \quad (a(bc))d, \quad (ab)(cd), \quad a((bc)d), \quad a(b(cd)).$$

Thus, the *Catalan numbers* are the numbers $C_1, C_2, \ldots$ satisfying

$$C_1 = 1, \qquad C_n = \sum_{k=1}^{n-1} C_k C_{n-k} \text{ for } n \geq 2.$$

We have $C_2 = 1 \cdot 1 = 1$; $C_3 = 1 \cdot 1 + 1 \cdot 1 = 2$; $C_4 = 1 \cdot 2 + 1 \cdot 1 + 2 \cdot 1 = 5$ (as illustrated above); and so on. Each value C_n for $n \geq 1$ is uniquely determined by these conditions.

Let $c(x)$ be the generating function:

$$c(x) = \sum_{n \geq 1} C_n x^n.$$

We claim that

$$c(x)^2 = c(x) - x.$$

For consider the coefficient of x^n on the left-hand side. If $n \geq 2$, we obtain a contribution to this term by taking the term in x^k in the first factor $c(x)$, and the term in x^{n-k} in the second; multiplying the coefficients (giving $C_k C_{n-k}$); and summing over k. The sum runs from 1 to $n-1$ since the lowest degree of a term is 1, not 0.

For $n = 1$, this argument is wrong; there is no term in x on the left, where as $c(x)$ starts with the term x. So we have to subtract x to make the coefficients equal. This gives the stated relation.

We can write this relation as

$$c(x)^2 - c(x) + x = 0.$$

Think of this as a quadratic in the unknown $c(x)$. The solution is

$$c(x) = \frac{1 \pm \sqrt{1-4x}}{2}.$$

Now we seem to have two solutions, whereas there should only be one. But we know that $c(0) = 0$, since the series has no constant term. If we took the plus sign, we would get $c(0) = 1$. So we have to take the minus sign:

$$c(x) = \frac{1 - \sqrt{1-4x}}{2}.$$

We can use this to find a formula for the Catalan numbers, using the Binomial Theorem:

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

We have

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

(We used the fact that $2\cdot 4\cdot(2n-2) = 2^{n-1}(n-1)!$.)

Now C_n is the coefficient in x^n in this series multiplied by $-1/2$; so we have

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

Exercises

1. Let $s(n)$ be the number of expressions for n as a sum of positive integers. For example,

$$4 = 3+1 = 1+3 = 2+2 = 1+1+2 = 1+2+1 = 2+1+1 = 1+1+1+1,$$

so $s(4) = 8$.

(a) Show that $s(1) = 1$ and

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

for $n \geq 2$.

(b) Deduce that $s(1) = 1$ and $s(n) = 2s(n-1)$ for $n \geq 2$.

(c) Hence show that $s(n) = 2^{n-1}$ for $n \geq 1$.

Notice how we have converted a rather complicated recurrence relation into a much simpler one!

2. Solve the recurrence relation and initial conditions

$$a_0 = 2, a_1 = 4, a_2 = 7, \qquad a_n = 4a_{n-1} - 5a_{n-2} + 2a_{n-3} \text{ for } n \geq 3.$$

3. I purchase an item costing n pence. I have a large number of 1 and 2 pence coins at my disposal. In how many ways can I pay for the item

(a) if I am buying it from a machine and have to insert the coins one at a time;

(b) if I am buying it in a shop and can hand the money over all at once?

4. Solve the recurrence relation and initial conditions

$$a_0 = 1, \quad a_1 = 1, \qquad a_n = 3a_{n-1} - 2a_{n-2} \text{ for } n \geq 2.$$

5. Solve the recurrence relation and initial conditions

$$a_0 = 2, \qquad a_n = a_{n-1}^2 \text{ for } n \geq 1.$$

6. Let

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Prove by induction that

$$A^{n+1} = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}$$

for $n \geq 1$, where F_n is the nth Fibonacci number.

7. (a) Use the recurrence relation in the text to prove that the derangement numbers $d(n)$ satisfy the simpler recurrence

$$d(0) = 1, \qquad d(n) = nd(n-1) + (-1)^n \text{ for } n \geq 1.$$

(b) Now put $f(n) = d(n)/n!$. Show that

$$f(0) = 1, \qquad f(n) = f(n-1) + \frac{(-1)^n}{n!} \text{ for } n \geq 1.$$

Hence show that $f(n) = \sum_{k=0}^{n} (-1)^k/k!$, and deduce the formula in the text for $d(n)$.

(c) Use this formula to show that

$$\sum_{n \geq 0} \frac{d_n x^n}{n!} = \frac{\mathrm{e}^{-x}}{1-x}.$$

The series on the left is the *exponential generating function* of the derangement numbers.

8. Take a circle and put $2n$ points $P_1, Q_1, P_2, Q_2, \ldots, P_n, Q_n$ equally spaced around the circumference. A *matching* is a set of n chords to the circle such that

- each chord joins a point P_i to a point Q_j, for some i, j;
- each of the $2n$ points lies on exactly one of the chords;
- no two chords cross.

Let A_n be the number of different matchings.

(a) Show that $A_0 = 1$, $A_1 = 2$, $A_2 = 3$, $A_3 = 5$.

(b) Show that, for $n \geq 1$, we hvae

$$A_n = \sum_{i=1}^{n} A_{i-1} A_{n-i}.$$

[*Hint:* Consider the matchings in which P_1 is joined to the point Q_i, and show that there are $A_{i-1}A_{n-i}$ of these.]

(c) Hence show by induction that A_n is the $(n+1)$st Catalan number C_{n+1}.

Chapter 5

Partitions and permutations

It can be argued that combinatorics is about three things: subsets, partitions, and permutations. In the first section of the notes we counted the subsets of a set. In this section we count partitions and permutations.

5.1 Partitions: Bell numbers

A *partition* of a set X is a set P of subsets of X with the properties:

- any set in P is non-empty;
- any two sets in P are disjoint;
- the union of all the sets in P is X.

In other words, the sets of the partition cover X without any overlap.

By the *Equivalence Relation Theorem*, if R is an equivalence relation on X (a reflexive, symmetric and transitive relation), then the equivalence classes of R form a partition of X. Conversely, any partition is the set of equivalence classes of a (unique) equivalence relation. So the number of partitions of X is equal to the number of equivalence relations on X.

Let $B(n)$ be the number of partitions of an n-element set, say $\{1,2,\ldots,n\}$. The number $B(n)$ is the nth *Bell number*. It is easy to see that $B(0) = B(1) = 1$, $B(2) = 2$, and $B(3) = 5$. The five partitions of $\{1,2,3\}$ are

$$\{123\}, \{12,3\}, \{13,2\}, \{23,1\}, \{1,2,3\},$$

where we have written 12 instead of $\{1,2\}$ to avoid a proliferation of curly brackets.

Proposition 5.1 *The Bell numbers satisfy the recurrence*

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

Proof We have seen that the initial condition holds. For the recurrence, we ask: how many partitions are there such that the part containing n has exactly k elements? We must have $1 \le k \le n$. We have to choose this part, which involves choosing $k-1$ of the remaining $n-1$ elements to go in a part with n; this can be done in $\binom{n-1}{k-1}$ ways. Then we must partition the remaining $n-k$ points, which can be done in B_{n-k} ways. Multiplying, and summing over k, gives the result.

This recurrence can be used to find a generating function for the Bell numbers. The type of generating function we use is called an *exponential generating function*, or e.g.f. for short. This has the form

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

The name is because of the relation to the exponential function

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

We claim that

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

For on the left we have

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x) = \sum_{n \ge 1} \frac{B(n)x^{n-1}}{(n-1)!}$$

on cancelling the n from the derivative into $n!$. So the coefficient of x^{n-1} is $B(n)/(n-1)!$.

On the right, to obtain the coefficient of x^{n-1}, we take the coefficient of x^{k-1} in $\exp(x)$ (which is $1/(k-1)!$, multiply by the coefficient of x^{n-k} in $F(x)$ (which is $B(n-k)$), and sum, obtaining

$$\sum_{k=1}^{n} \frac{1}{(k-1)!} \cdot \frac{B(n-k)}{(n-k)!} = \frac{1}{(n-1)!} \sum_{k=1}^{n} \binom{n}{k} B(n-k) = \frac{B(n)}{(n-1)!}$$

So the two sides are equal.

Now the differential equation can be separated as

$$\frac{1}{F(x)}\frac{\mathrm{d}}{\mathrm{d}x}F(x) = \exp(x).$$

Integrating, we obtain

$$\log F(x) = \exp(x) + c,$$

so $F(x) = \mathrm{e}^c \exp(exp(x))$. But $F(0) = 1$ (since $B(0) = 1$), so $c = -1$, and we conclude that

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

Unfortunately this simple formula for the e.g.f. doesn't help us find a formula for $B(n)$. Even its asymptotic behaviour for large n is very complicated.

5.2 Partitions: Stirling numbers

We saw that the subsets of an n-element set (which are 2^n in number) can be split up according to the number of elements they contain. There are $\binom{n}{k}$ k-element subsets, and so we have

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

In the same way, the partitions of a set can be split up. For $1 \leq k \leq n$, let $S(n,k)$ be the number of partitions of an n-element set having k parts. The numbers $S(n,k)$ are called *Stirling numbers of the second kind*. (We meet Stirling numbers of the first kind later.) Thus, we have

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

In the last section we listed the partitions of a 3-element set; from the list we see that $S(3,1) = 1$, $S(3,2) = 3$, $S(3,3) = 1$.

There is a recurrence relation for Stirling numbers, similar to that for binomial coefficients:

Proposition 5.2 *$S(n,1) = S(n,n) = 1$ and*

$$S(n,k) = S(n-1,k-1) + kS(n-1,k)$$

for $1 < k < n$.

Proof The initial values are clear: there is a unique partition with a single part, and a unique partition with n parts (each part has one element).

Now consider the partitions of $\{1,\ldots,n\}$ with k parts, and divide them into two classes:

- Those in which $\{n\}$ is a part. These are obtained by adding the set $\{n\}$ to a partition of $\{1,\ldots,n-1\}$ with $k-1$ parts. So there are $S(n-1,k-1)$ of them.
- Those in which n belongs to a part of size bigger than 1. If we delete n from this part, we get a partition of $\{1,\ldots,n-1\}$ with k parts. But now, to go back, we have to choose a partition, and also choose one of its k parts to add the element n to. So there are $kS(n-1,k)$ of these.

Adding gives the result.

We can arrange the Stirling numbers in a trinagle like Pascal's, except that we will line it up on the left. $S(n,k)$ is the kth number in the nth row, starting both counts at 1.

1					
1	1				
1	3	1			
1	7	6	1		
1	15	25	10	1	
1	31	90	65	15	1

The rule is a little different from Pascal's. To find the next element in column k, we multiply the number immediately above it by k and add the number above and to the left.

The Stirling numbers have a remarkable property. Recall the falling factorial:

$$(x)_k = x(x-1)(x-2)\cdots(x-k+1).$$

Proposition 5.3 *For $n \geq 1$,*

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

First proof We use the fact that $(x)_{k+1} = (x)_k(x-k)$. Now our proof is by induction, the result being clear for $n = 1$. Assuming the result for $n-1$, we have

$$x^n = x^{n-1} \cdot x$$

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

For $k \leq n-1$, the coefficient of $(x)_k$ is $S(n-1,k-1)+kS(n-1,k) = S(n,k)$. (We have to shift the argument k down by one in the first term.) For $k = n$, the coefficient of $(x)_n$ is $S(n-1,n-1)$; but this is equal to $S(n,n)$, since both are 1. The induction is complete.

Second proof Here is a completely different proof. Let m and n be positive integers. We know that m^n is the number of ordered selections of n objects from a set of m objects, with repetitions allowed; if repetitions are not allowed, the number is $(m)_n$. Let us count the selections with repetitions allowed in another way. Take any such selection, say $(x_1,x_2,\ldots,x_n)$. Define a relation $\sim$ on the set $\{1,\ldots,n\}$ by the rule that $i \sim j$ if $x_i = x_j$. This is an equivalence relation, so it corresponds to a unique partition of the set $\{1,\ldots,n\}$. If this partition has k classes, say, then there are k distinct elements among $x_1,\ldots,x_n$, which we can regard as a selection of k things from a set of m with repetitions not allowed.

Now given a partition of $\{1,\ldots,n\}$ with k parts, and a selection $(y_1,\ldots,y_k)$ of k from m with repetitions not allowed, we can recover the original selection: put $x_i = y_j$ if i belongs to the jth part of the partition. So the number of selections with k distinct elements is $S(n,k)m^k$, and we conclude that the total number (which we know to be m^n) is the sum of all these values:

$$m^n = \sum_{k=1}^{n} (m)_k.$$

Now consider the two polynomials x^n and $\sum_{k=1}^{n}(x)_k$. We know that they take the same value if any positive integer m is substituted for x. So they are equal as polynomials. For their difference is a polynomial of degree at most n; if it is not identically zero, it could have at most n roots. So we have

$$x^n = \sum_{k=1}^{n} (x)_k,$$

as required.

In Exercise 1 at the end of this chapter, we will see that some values of $S(n,k)$ can be calculated. A general formula will be given in a later chapter of the notes.

5.3 Permutations: cycle decomposition

A *permutation* is a one-to-one and onto function from a set to itself; in other words, an arrangement of the elements of the set. In this section and the next we will be counting permutations. The total number of permutations of a set of size n is $n!$. But we will subdivide the permutations, much as we did for partitions, and count the parts.

Recall the *cycle decomposition* of a permutation, which we do by example. Consider the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 3 & 5 & 7 & 2 & 6 & 8 & 10 & 9 & 1 \end{pmatrix}$$

of the set $\{1,\ldots,10\}$. (This *two-line notation* indicates the permutation which maps 1 to 4, 2 to 3, 3 to 5, and so on.) To compute the cycle decomposition, we start anywhere (say 1), and follow the iterates of the function applied to our starting point until we return there. If we have used every point, we are finished; otherwise, we close the cycle, and start another one at an unused point. Continue until every point has been used. We write a cycle as a list of points in brackets, separated by commas. For our example above, the cycle decomposition is

$$\pi = (1,4,7,8,10)(2,3,5)(6)(9).$$

(Note that points fixed by the permutation show up as cycles of length i.)

The cycle decomposition of a permutation is not unique. We could start each cycle at any point, and write the cycles in any order. For example, the permutation above could also be written

$$\pi = (3,5,2)(9)(4,7,8,10,1)(6).$$

5.4 Permutations: Stirling numbers

The *parity* of a permutation π is the parity (odd or even) of the number $n - c(\pi)$, where $c(\pi)$ is the number of cycles of π (including fixed points). A permutation is called *even* or *odd* according to its parity. Sometimes we talk about the *sign* of π: this is $(-1)^{n-c(\pi)}$, that is, $+1$ if π is even and -1 if π is odd.

Now the *unsigned Stirling number of the first kind*, $u(n,k)$, is defined to be the number of permutations of $\{1,\ldots,n\}$ which have k cycles; the *Stirling number of the first kind* is $s(n,k) = (-1)^{n-k}u(n,k)$. (The sign we put in front is the sign of the permutations we are counting.)

We have $|s(n,k)| = u(n,k)$ and

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

(since the sum counts all permutations).

Proposition 5.4 *$s(n,1) = (-1)^{n-1}(n-1)!$, $s(n,n) = 1$, and*

$$s(n,k) = s(n-1,k-1) - (n-1)s(n-1,k).$$

Proof How many permutations of $\{1,\ldots,n\}$ have a single cycle? Since the cycle can start anywhere, we may as well begin with 1. Then we can visit the other $n-1$ numbers in any order. So the number of cyclic permutations is $(n-1)!$. Each has sign $(-1)^{n-1}$, so $s(n,1) = (-1)^{n-1}(n-1)!$.

There is only a single permutation with n cycles, namely the identity permutation which fixes every point; it has sign $+1$. So $s(n,n) = 1$.

For the recursion, take the set of permutations with k cycles, and divide into two classes:

- Those which fix the point n (that is, which have a cycle (n)). Deleting this cycle gives a permutation of $\{1,\ldots,n-1\}$ with $k-1$ cycles. Clearly the procedure reverses. Since $(-1)^{(n-1)-(k-1)} = (-1)^{n-k}$, the contribution to $s(n,k)$ from these permutations is $s(n-1,k-1)$.

- Those which move the point n, (that is, in which n is in a cycle of length greater than 1). Deleting n from the cycle containing it gives a permutation of $\{1,\ldots,n-1\}$, also with k cycles. When we reverse the construction, for each permutation of $\{1,\ldots,n-1\}$, there are $n-1$ places in which we could insert n. Since $(-1)^{(n-1)-k} = -(-1)^{n-k}$, the contribution of these permutations to $s(n,k)$ is $-(n-1)s(n-1,k)$.

Adding these two terms gives the result.

We can write these Stirling numbers, like the others, in a triangular array. This time, ignoring signs, a given entry is obtained by multiplying the entry above by its row number (rather than its column number, as before) and adding the entry

above and to the left. Then put in signs in a chessboard fashion. We obtain:

$$\begin{array}{cccccc} 1 & & & & & \\ -1 & 1 & & & & \\ 2 & -3 & 1 & & & \\ -6 & 11 & -6 & 1 & & \\ 24 & -50 & 35 & -10 & 1 & \\ -120 & 274 & -225 & 85 & -15 & 1 \end{array}$$

From Proposition 5.4 we can prove a result which is the "inverse" of Proposition 5.3:

Proposition 5.5 *For $n \geq 1$,*

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

For example, since $(x)_3 = x(x-1)(x-2) = x^3 - 3x^2 + 2x$, we have

$$s(3,3) = 1, \qquad s(3,2) = -3, \qquad s(3,1) = 2.$$

Proof Again the proof is by induction. For $n = 1$, both sides of the equation are equal to x. So suppose that the result is true for $n-1$. Then we have

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

The coefficient of x^k for $k < n$ is $s(n-1,k-1) - (n-1)s(n-1,k)$ (moving the index down by one in the first term). The coefficient of x^n is $s(n-1,n-1) = 1 = s(n,n)$.

One consequence of Propositions 5.3 and 5.5 is the following result.

Proposition 5.6 *The lower triangular matrices of Stirling numbers of the first and second types are inverses of each other.*

This applies whether we regard them as "infinite matrices" or chop them off after a fixed number of rows. Because the matrices are lower triangular, even multiplying the infinite matrices only involves finite sums.

Proof The polynomials with constant term zero form a vector space V. The following two sequences are bases for V:

- $x, x^2, x^3, \ldots$
- $(x)_1, (x)_2, (x)_3, \ldots$

Propositions 5.3 and 5.5 show that the two matrices of Stirling numbers are the transition matrices between these two bases.

Another consequence of Proposition 5.5 is:

Proposition 5.7 *For $n \geq 2$, the numbers of even and odd permutations of $\{1, \ldots, n\}$ are equal.*

Proof Because $n \geq 2$, we see that $(x)_n$ has a factor $(x-1)$, and so is zero when we put $x = 1$. Substituting $x = 1$ into Proposition 5.5, we have

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

for $n \geq 2$. Now an even permutation contributes $+1$ to this sum, and an odd permutation contributes -1; the contributions must match.

Remark For those who have done some abstract algebra, here is a completely different proof of this result. The set of all permutations of $\{1, \ldots, n\}$ forms a group (with the operation of composition), called the *symmetric group* and written S_n. The mapping that takes a permutation to its sign is a homomorphism from S_n to the multiplicative group $\{\pm 1\}$ of order 2; the even permutations form the kernel of this homomorphism, and therefore comprise a normal subgroup of S_n of index 2, called the *alternating group* and written A_n. The odd permutations form a coset of A_n. Now Lagrange's Theorem tells us that the subgroup and its coset have equally many elements.

Exercises

1. (a) Prove each of the following statements (i) by directly counting the partitions, (ii) by using the recurrence relation:

- $S(n,2) = 2^{n-1} - 1$;
- $S(n,n-1) = \binom{n}{2}$.

(b) Find a formula for $S(n, n-2)$.

2. (a) Prove (i) by directly counting the permutations, (ii) by using the recurrence relation, that $s(n, n-1) = -\binom{n}{2}$.

(b) Find a formula for $s(n, n-2)$.

3. Calculate the number of permutations of $\{1,2,3,4,5,6\}$ with three cycles

(a) by using the recursion formula for the appropriate Stirling numbers;

(b) by listing the possible cycle lengths of such a permutation and counting the number of permutations with each possible cycle structure.

4. Let k be given, and let $p_k(n)$ be the probability that a randomly-chosen permutation of $\{1,\ldots,n\}$ has exactly k fixed points. Show that $p_k(n) = \binom{n}{k} d(n-k)/n!$, where $d(n-k)$ is the $(n-k)$th derangement number, and hence show that

(a) $\displaystyle\sum_{k=0}^{n} \binom{n}{k} d(n-k) = n!$;

(b) $\displaystyle\lim_{n\to\infty} p_k(n) = \frac{\mathrm{e}^{-1}}{k!}$.

[If you have studied some probability theory, the last statement says that the number of fixed points of a random permutation of the set $\{1,\ldots,n\}$ approaches a Poisson distribution with parameter 1 as $n \to \infty$.]

Chapter 6

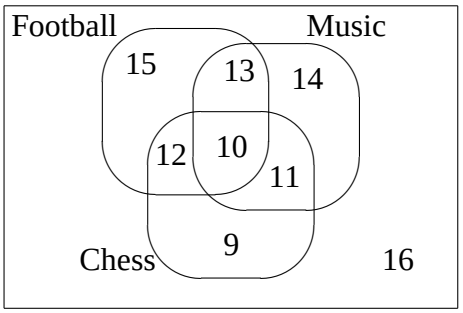
The Principle of Inclusion and Exclusion

Suppose that, in a class of 100 pupils, we are given the following information:

- 50 play football, 48 play music and 42 play chess;
- 23 play football and music, 22 play football and chess, and 21 play music and chess;
- 10 play all three.

How many pupils do none of the three activities?

We can illustrate the eight possible combinations of activities with a Venn diagram. Starting from the inside and working out, it is possible to see that the numbers are as shown in the diagram. So the answer is 16.



In this section we are going to develop a formula for this number, so that the answer can be calculated directly from the given data. We will then use this formula to count derangements and to find a formula for the Stirling numbers of the second kind.

6.1 PIE

The set-up is as follows. We have a "universal" set X, and a collection $A_1, A_2, \ldots, A_n$ of subsets of X. (In the example, X is the set of 100 pupils, $n = 3$, and A_1, A_2, A_3 are the sets of pupils who take part in each of the three activities. We are given $|X|$, $|A_1|$, $|A_2|$ and $|A_3|$, and also the sizes of the intersections $A_1 \cap A_2$, $A_1 \cap A_3$, $A_2 \cap A_3$ and $A_1 \cap A_2 \cap A_3$. For example, $A_1 \cap A_2$ is the set of pupils who play both football and music.

In order to simplify the notation, we will denote $A_1 \cap A_2$ by $A_{\{1,2\}}$. More generally, for every subset I of the index set $\{1, 2, \ldots, n\}$, we let

$$A_I = \bigcap_{i \in I} A_i.$$

Thus, A_I is the set of elements belonging to all the sets A_i for which the index i belongs to I, and possibly some others. By convention, $A_{\{i\}} = A_i$, and $A_\emptyset = X$.

Theorem 6.1 (Principle of Inclusion and Exclusion) *The number of elements of X which lie in none of the sets $A_1, \ldots, A_n$ is equal to*

$$\sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} |A_I|.$$

In our example, the number of children taking part in no activity is

$$100 - 50 - 48 - 42 + 23 + 22 + 21 - 10 = 16,$$

agreeing with what we found directly.

Proof We look at the expression in the theorem. It is the sum of cardinalities of various subsets of X with plus and minus signs. We evaluate this by looking at each element $x \in X$ and seeing how much it contributes to the sum.

An element x which lies in none of the sets A_i gives a contribution $+1$ from $A_\emptyset = X$, and no contribution from any of the other sets.

Now consider an element x which lies in some of the A_i, and let

$$J = \{i \in \{1, \ldots, n\} : x \in A_j\}$$

be the set of indices of sets containing x. Then $x \in A_I$ if and only if $I \subseteq J$. So the contribution of X is

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

Let $|J| = j$. Then there are $\binom{j}{i}$ subsets of J of size i, and each of them gets the sign $(-1)^i$. So the contribution is

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

by the Binomial Theorem.

So the elements in none of the sets A_i contribute $+1$ to the sum, while the elements lying in some of these sets contribute 0. Hence the sum is just equal to the number of elements lying in none of the A_i, as was to be proved.

As a corollary we have the following result:

Proposition 6.2 *Suppose that $A_1, \ldots, A_n$ are subsets of a set X. Suppose that $|X| = m_0$ and that $|A_i| = m_1$ for $i = 1, \ldots, n$. Suppose further that the intersection of any j of the sets A_i has cardinality m_j. Then the number of elements in none of the sets is*

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

For in the sum in Theorem 6.1, there are $\binom{n}{j}$ sets A_I with $|I| = j$; each of them has cardinality m_j and contributes $(-1)^j m_j$ to the sum.

6.2 Surjections and Stirling numbers

Let $|A| = n$ and $|B| = k$. How many functions are there from A to B? To specify a function f, we simply have to define the n values $f(a)$ for $a \in A$, which can be arbitrary; so the number is k^n.

How many of these functions are injective (one-to-one)? To count these, we proceed as above, making sure that the values are all distinct; that is, we sample without replacement. The answer is $(k)_n = k(k-1)\cdots(k-n+1)$. Note that this number is zero if $n > k$; there can be no injective function from a set to a smaller set.

How many of the functions are surjective (onto)? This is more difficult to count by elementary means; but PIE allows us to find the answer.

Let X be the set of all functions from $\{a_1, \ldots, a_n\}$ to $\{b_1, \ldots, b_m\}$. Then $|X| = k^n$.

For $i = 1, \ldots, k$, let A_i be the set of all those functions which never take the value b_i. (We are trying to count functions that take all values; to apply PIE we need to remove all the functions that miss at least one value). Then the functions in A_i take $k-1$ possible values, so there are $(k-1)^n$ of them.

Now suppose that $I \subseteq \{1, \ldots, k\}$ with $|I| = j$. Then A_I is the intersection of the sets A_i for $i \in I$, so it consists of all the functions which take no value b_i for $i \in I$. These functions have $k-i$ possible values, so there are $(k-i)^n$ of them.

Since the surjections are the functions lying in none of the sets A_i, Proposition 6.2 gives:

Theorem 6.3 *The number of surjections from a n-element set to an k-element set is*

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

Remark This formula is useful but has its drawbacks. For example, it should give zero when $k > n$, since there cannot be a surjection from a set to a larger set. But this is quite hard to show directly – have a try! Also, it is not obvious that

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

though of course this must hold since if $k = n$ then the surjections are bijections and there are $n!$ of them.

This theorem allows us to find a formula for the Stirling number $S(n,k)$ of the second kind. Remember that $S(n,k)$ is the number of partitions of an n-element set into k parts. Given such a partition of $\{1, \ldots, n\}$, say, we can define a surjection from $\{1, \ldots, n\}$ to $\{1, \ldots, k\}$ as follows: if $P_1, \ldots, P_k$ are the parts, map the elements of part P_i to the value i. Since $P_i \neq \emptyset$, some element is mapped to i for all $i \in \{1, \ldots, k\}$, so we do have a surjection. In fact, a given partition gives $k!$ surjections, since we can order the parts in any way we like. Conversely, any surjection f gives a partition into k parts, where the ith part is the inverse image of i under f. Thus the number of surjections is $k!$ times the number of partitions, and so we have:

Proposition 6.4

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

6.3 Derangements

We can use a similar method to find a formula for the number of derangements of $\{1,\ldots,n\}$ (permutations with no fixed points).

Let X be the set of all permutations of $\{1,\ldots,n\}$. Then $|X| = n!$. Now for $i = 1,\ldots,n$, let A_i be the set of permutations which fix the point i. There are $(n-1)!$ such permutations, since they can permute the other $n-1$ elements arbitrarily. For any $I \subseteq \{1,\ldots,n\}$, A_I consists of permutations fixing every point in I. If $|I| = j$, these permutations move the other $n-j$ points arbitrarily, so $|A_I| = (n-j)!$. Now a permutation is a derangement if and only if it does not fix any point; so, if d_n is the number of derangements, then Proposition 6.2 gives:

Proposition 6.5

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

This is the formula we saw in Chapter 5.

Exercises

1. An opinion poll reports that the percentage of voters who would be satisfied with each of three candidates A, B, C for President is 65%, 57%, 58% respectively. Further, 28% would accept A or B, 30% A or C, 27% B or C, and 12% would be content with any of the three. What do you conclude?

2. I have 25 sweets to distribute to a class of 10 children.

(a) In how many ways can I distribute the sweets?

(b) In how many ways can I distribute the sweets if I give Alice at least four sweets?

(c) In how many ways can I distribute the sweets if I give both Alice and Bob at least four sweets?

(d) Use the Principle of Inclusion and Exclusion to count the number of ways I can distribute the sweet if no child is to have more than three sweets.

Chapter 7

Families of sets

We know now how many subsets of the set $\{1,2,\ldots,n\}$ there are altogether (namely 2^n), and the number of subsets of fixed size k (namely $\binom{n}{k}$). We are now going to turn to collections of sets satisfying various other conditions. Of course we are really talking about "sets of sets" here, but to avoid confusion we will refer to them as "families of sets". Typically we will denote a family of sets by script capital F, thus: $\mathscr{F}$.

The set of all subsets of a set X is called the *power set* of X, and denoted by $\mathscr{P}(X)$. Thus, a family of sets is a subset of $\mathscr{P}(\{1,\ldots,n\})$.

The main questions will be: Suppose that we place some restriction on the relationships between sets in a family. What is the largest number of sets that we can have? Which families reach this upper bound? We will examine one case in detail, and state and prove *Sperner's Theorem*. Then we will look more briefly at intersecting families and at families where any two sets intersect in a single element.

7.1 Sperner's theorem

A family $\mathscr{F}$ of subsets of $\{1,\ldots,n\}$ is called a *Sperner family* if the following is true:

> For any two distinct sets $A,B \in \mathscr{F}$, neither of them contains the other; that is, $A \not\subseteq B$ and $B \not\subseteq A$.

Our first question is: What is the largest Sperner family?

The cheapest way to build a Sperner family is to take all the subsets of some fixed size k. This gives us $\binom{n}{k}$ such sets, and clearly no such set can contain

another. We saw in Assignment 1, Question 2, that for fixed n the binomial coefficient $\binom{n}{k}$ is greatest when $k = n/2$ (if n is even) or when $k = (n-1)/2$ or $k = (n+1)/2$ (when n is odd – these two binomial coefficients are equal).

Other Sperner families can be constructed: $\{\{1,2\},\{1,3,4\},\{2,3,4\}\}$ is an example. Perhaps it is possible to find a larger family containing sets of different sizes? The first part of *Sperner's Theorem* tells us that it is not:

Theorem 7.1 *Let $\mathscr{F}$ be a Sperner family of subsets of $\{1,\ldots,n\}$. Then*

$$|F| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

Proof The following ingenious proof is known as the LYM method, since it was invented by Lubell, Yamamoto and Meshalkin (and also by Bollobás; we have seen many instances of mis-named theorems in mathematics!)

A *chain* is a sequence of subsets of $\{1,2,\ldots,n\}$, in which each set is contained in the next set in the sequence. The maximal number of sets in a chain is $n+1$; in such a maximal chain $(A_0,A_1,\ldots,A_n)$, the set A_k has k elements. Such a chain starts with the empty set and adds one new element each time. Thus, any maximal chain is described by a permutation π of $\{1,\ldots,n\}$; we have $A_i = \{\pi(1),\ldots,\pi(i)\}$.

It follows that the number of maximal chains is equal to the number of permutations of $\{1,\ldots,n\}$, namely $n!$.

Next we ask: how many maximal chains contain a given set A? If $|A| = k$, then the first k numbers $\pi(1),\ldots,\pi(k)$ in the permutation must be the elements of A, and the last $n-k$ must be the remaining elements; so the number of maximal chains containing A is $k!\,(n-k)!$.

Now, let $\mathscr{F}$ be a Sperner family. Then, for $A,B \in \mathscr{F}$, no maximal chain can contain both A and B. For if so, such a maximal chain would be $(\ldots,A,\ldots,B,\ldots)$, say, and then $A \subseteq B$, contradicting the definition of a Sperner family. So if we add up the numbers of maximal chains containing all the sets in $\mathscr{F}$, the total cannot be more than the total number of maximal chains:

$$\sum_{A\in\mathscr{F}} |A|!\,(n-|A|)! \leq n!,$$

from which we deduce (dividing both sides by $n!$) that

$$\sum_{A\in\mathscr{F}} \frac{1}{\binom{n}{|A|}} \leq 1.$$

This result (which is valid for any Sperner family) is known as the *LYM inequality*.

Now we saw, the largest binomial coefficient is $\binom{n}{\lfloor n/2 \rfloor}$. So if we replace the denominators on the left by larger quantities, we make the sum smaller, and we find

$$\frac{|\mathscr{F}|}{\binom{n}{\lfloor n/2 \rfloor}} = \sum_{A \in \mathscr{F}} \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \leq 1,$$

and so we conclude that

$$|\mathscr{F}| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

Our second question is: What are the families which meet this bound? We have seen that, if n is even, the set of all n-element subsets meets the bound; if n is odd, the set of all $(n-1)/2$-element subsets meets the bound, and so does the set of all $(n+1)/2$-element subsets. The second part of Sperner's Theorem tells us that there are no others:

Theorem 7.2 *Suppose that $\mathscr{F}$ is a Sperner family of subsets of $\{1,\ldots,n\}$ with $|F| = \binom{n}{\lfloor n/2 \rfloor}$. Then*

(a) if n is even, then $\mathscr{F}$ consists of all the $n/2$-element subsets;

(b) if n is odd, then either $\mathscr{F}$ consists of all the $(n-1)/2$-element subsets, or it consists of all the $(n+1)/2$-element subsets.

Proof We have to look back at the preceding proof; in particular, the step from the LYM inequality to the next line. This involved replacing the denominators $\binom{n}{|A|}$ by the possibly larger denominators $\binom{n}{\lfloor n/2 \rfloor}$. If it ever occurred that the new denominator was strictly larger, then we would have strict inequality in the next step, and we would conclude that $|\mathscr{F}| < \binom{n}{\lfloor n/2 \rfloor}$. So all the sets in the family $\mathscr{F}$ must have size $n/2$ (if n is even), or size $(n-1)/2$ or $(n+1)/2$ (if n is odd).

So the theorem is proved in the case where n is even.

If n is odd, however, we have one more job to do: we must rule out the possibility that there are sets of both possible sizes in $\mathscr{F}$. So let $n = 2m+1$.

Looking at the proof again we see that, if the bound is met, then every maximal chain must contain a set of $\mathscr{F}$. If A and B are sets of sizes m and $m+1$ respectively,

then there is a maximal chain containing both A and B; so $\mathscr{F}$ must contain exactly one of these two sets. So if $A \in \mathscr{F}$ then $B \notin \mathscr{F}$ and conversely.

Now if C and D are two sets of size m, then it is possible to find a sequence of sets of sizes alternately m and $m+1$ from C to D. If $C \in \mathscr{F}$, then we find that the sets of size $m+1$ don't belong to $\mathscr{F}$, while the sets of size m all do. So every set of size m belongs to $\mathscr{F}$. This implies that no set of size $m+1$ can belong to $\mathscr{F}$, so $\mathscr{F}$ is the set of all m-element sets.

In a similar way, if $\mathscr{F}$ contains an $(m+1)$-element set, then it consists of all $(m+1)$-element sets.

Here is an example to illustrate this proof. Suppose that $n = 7$, $m = 3$, and $\mathscr{F}$ contains the set $\{1,2,3\}$. We want to show that it contains $\{4,5,6\}$. We produce the sequence

$$(\{1,2,3\},\{1,2,3,4\},\{2,3,4\},\{2,3,4,5\},\{3,4,5\},\{3,4,5,6\},\{4,5,6\}).$$

Now $\{1,2,3\} \in \mathscr{F}$; so $\{1,2,3,4\} \notin \mathscr{F}$; so $\{2,3,4\} \in \mathscr{F}$; and so on. Finally, $\{4,5,6\} \in \mathscr{F}$, as required.

7.2 Intersecting families

We say that two sets A and B *intersect* if their intersection is not empty: $A \cap B \neq \emptyset$. A family of subsets of $\{1,\ldots,n\}$ is an *intersecting family* if any two of its members intersect.

Theorem 7.3 *The maximum size of an intersecting family of subsets of* $\{1,\ldots,n\}$ *is* 2^{n-1}.

Proof First we note that there do exist intersecting families containing 2^{n-1} sets. For consider all the subsets of $\{1,\ldots,n\}$ which contain the element n. Such a subset has the form $\{n\} \cup A$, where A is an arbitrary subset of $\{1,\ldots,n-1\}$; there are 2^{n-1} choices for A. Clearly the resulting family is intersecting since any two members have at least the element n in common.

We have to show that it is not possible to have more than 2^{n-1} sets in an intersecting family. To see this, we list the sets in complementary pairs $\{A,A'\}$, where $A' = \{1,\ldots,n\} \setminus A$. There are $2^n/2 = 2^{n-1}$ such pairs. Now an intersecting family $\mathscr{F}$ can contain at most one set from each pair, since A and A' are disjoint. So $|\mathscr{F}| \leq 2^{n-1}$, as required.

Following the pattern of Sperner's Theorem, we should now go on to find all the intersecting families which meet this bound. But this is not possible; there are many different intersecting families meeting the bound. Here are a few examples.

- As in the theorem, if we take all the subsets which contain some fixed element of $\{1,\ldots,n\}$, we obtain an intersecting family of size 2^{n-1}.
- If n is odd, choose all the subsets which have size strictly greater than $n/2$. Any two such subsets must intersect; and there are 2^{n-1} of them, since out of each complementary pair we take the larger one.
- If n is even we can modify this argument. Take all subsets with more than $n/2$ elements, and out of the sets of size $n/2$ pick one of each complementary pair. For example, for $n = 4$, we could have either of the following (where I write $\{1,2,3\}$ as 123 for brevity):

$$\{1234, 123, 124, 134, 234, 12, 13, 14\}$$
$$\{1234, 123, 124, 134, 234, 12, 13, 23\}$$

- Many other examples are possible. For $n = 7$, it can be showed that there are $2^6 = 64$ sets which contain at least one of $123, 145, 167, 246, 257, 347, 356$. Since these seven sets intersect, any sets containing them will also intersect.

Let's ask a different question. What is the maximum size of an intersecting family of k-element subsets of $\{1,\ldots,n\}$?

If $n < 2k$, then any k-element set contains more than half of $\{1,\ldots,n\}$, so any two of them intersect. The answer then is $\binom{n}{k}$, which is not very interesting. So we assume that $n \geq 2k$.

In this case, let $\mathscr{F}_k(i)$ consist of all the k-element subsets of $\{1,\ldots,n\}$ which contain the element i, for $i \in \{1,\ldots,n\}$. Then $|\mathscr{F}_k(i)| = \binom{n-1}{k-1}$, since we have to pick $k-1$ more elements from $\{1,\ldots,n\} \setminus \{i\}$. A famous theorem called the *Erdős–Ko–Rado theorem* shows that this is the best we can do:

Theorem 7.4 *(a) If $n \geq 2k$, then an intersecting family of k-element subsets of $\{1,\ldots,n\}$ has size at most $\binom{n-1}{k-1}$.*

(b) If $n > 2k$, then the only intersecting families which have size $\binom{n-1}{k-1}$ are the sets $\mathscr{F}_k(i)$, for $i \in \{1,\ldots,n\}$.

The proof of this theorem is beyond the scope of the course, but I have written out the main part in the final section of this chapter. If you are interested in combinatorics, you are encouraged to study this proof, which has a lot of important ideas in it.

I will give here just the proof of part (a) of the theorem in the case when $n = 2k$. This is similar to what we did before. Arrange the k-element subsets of $\{1,\ldots,2k\}$ into complementary pairs $\{A,A'\}$, where $A' = \{1,\ldots,2k\} \setminus A$. Then an intersecting family contains at most one of each complementary pair, so the size of such a family is at most

$$\frac{1}{2}\binom{2k}{k} = \binom{2k-1}{k-1}.$$

7.3 The de Bruijn–Erdős theorem

Now we will be even more specific. What is the maximum size of a family of subsets of $\{1,\ldots,n\}$ having the property that any two of them have exactly one point in common?

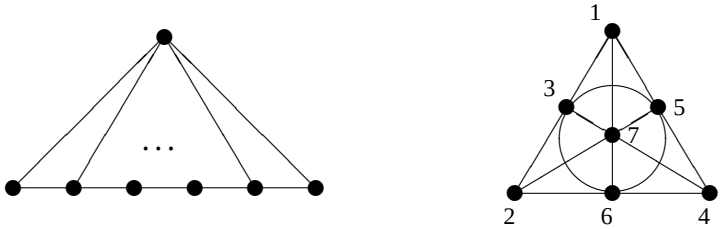
First we give some examples of such families.

- For $i \in \{1,\ldots,n\}$, let $\mathscr{A}(i)$ consist of the set $\{i\}$ together with all sets of the form $\{i,j\}$ for $j \neq i$. Then $|\mathscr{A}(i)| = n$ and any two members of $\mathscr{A}(i)$ meet in the point $\{i\}$.

- For $i \in \{1,\ldots,n\}$, let $\mathscr{B}(i)$ consist of the set $\{1,\ldots,n\} \setminus \{i\}$ together with all sets of the form $\{i,j\}$ for $j \neq i$. Then $|\mathscr{B}(i)| = n$, and any two members of $\mathscr{B}(i)$ meet in one point.

- For $n = 7$, the seven sets $123, 145, 167, 246, 257, 347, 356$ have the required property. This configuration of seven sets is known as the *Fano plane*.

The diagram shows the second and third examples.



The *de Bruijn–Erdős theorem* shows that n is the maximum size, and describes the families which meet the bound:

Theorem 7.5 *Let $\mathscr{F}$ be a family of subsets of $\{1,\ldots,n\}$ with the property that $|A \cap B| = 1$ for all $A,B \in \mathscr{F}$. Then $|\mathscr{F}| \leq n$. Equality holds if and only if one of the following occurs:*

(a) $\mathscr{F} = \mathscr{A}(i)$ or $\mathscr{F} = \mathscr{B}(i)$ for some $i \in \{1,\ldots,n\}$;

(b) there is a positive integer q such that $n = q^2+q+1$, every set in $\mathscr{F}$ contains $q+1$ elements, and every element of $\{1,\ldots,n\}$ is contained in $q+1$ of the sets of $\mathscr{F}$.

Note that the third example in our list above satisfies conclusion (c), with $q = 2$: we have $n = 2^2+2+1 = 7$, each set has $2+1 = 3$ elements and each point lies in $2+1 = 3$ sets of $\mathscr{F}$.

The proof will not be given here: it can be found in the recommended textbook for the course. We will end this chapter by a brief look at some families satisfying conclusion (c) of the theorem.

7.4 Finite fields and projective planes

We first do a little bit of linear algebra. Recall that a *field* is an algebraic structure in which addition, subtraction, multiplication and division (except by zero) are all possible, and the commutative, associative and distributive laws apply. If F is a field, then we can talk about vector spaces over F; the standard m-dimensional vector space is the set F^m of all m-tuples of elements of F, with coordinatewise operations.

We are interested in finite fields. Do any exist? Yes, the integers mod p form a field whenever p is a prime number; this field is denoted by $\mathbb{Z}_p$. There are others too, as we will see later.

Theorem 7.6 *Let F be a finite field with q elements, and V an m-dimensional vector space over F. Then*

(a) $|V| = q^m$;

(b) the number of 1-dimensional subspaces of V is $(q^m-1)/(q-1)$;

(c) the number of 2-dimensional subspaces of V is

$$\frac{(q^m-1)(q^{m-1}-1)}{(q-1)(q^2-1)}.$$

Proof (a) Any m-dimensional vector space can be coordinatised by choosing a basis; so we can take it to be F^m. The result is clear.

(b) Any 1-dimensional subspace is spanned by a non-zero vector, of which there are q^m-1 in V. But, if c is any non-zero element of F, then v and cv span the same 1-dimensional subspace. So each such subspace has $q-1$ spanning vectors, and the number of such subspaces is $(q^m-1)/(q-1)$.

(c) Any 2-dimensional subspace is spanned by two linearly independent vectors v and w. There are $q^m - 1$ choices for v (since it must be non-zero) and $q^m - q$ choices for w (since it cannot be a multiple of v). Thus there are $(q^m - 1)(q^m - q)$ pairs (v, w). But by the same argument (putting $m = 2$) we see that any 2-dimensional subspace contains $(q^2 - 1)(q^2 - q)$ spanning pairs of vectors. So we get the number of subspaces by dividing, and cancelling a factor q.

We see that a 3-dimensional vector space has $(q^3 - 1)/(q - 1) = q^2 + q + 1$ 1-dimensional subspaces, and

$$\frac{(q^3 - 1)(q^2 - 1)}{(q - 1)(q^2 - 1)} = q^2 + q + 1$$

2-dimensional subspaces. Furthermore, if V is 3-dimensional, then

- any 2-dimensional subspace contains $q + 1$ 1-dimensional subspaces;
- any 1-dimensional subspace is contained in $q + 1$ 2-dimensional subspaces;
- any two 2-dimensional subspaces intersect in a 1-dimensional subspace.

The first part follows from the case $m = 2$ in the theorem; the second is proved by a similar argument. For the third, let $\dim(V) = 3$ and let W_1, W_2 be subspaces with $\dim(W_1) = \dim(W_2) = 2$. By the dimension formula,

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2).$$

Now the right-hand side is equal to 4. We must have $\dim(W_1 + W_2) = 3$ (it cannot be larger since $\dim(V) = 3$, and it cannot be smaller since $W_1 + W_2$ properly contains W_1). So $\dim(W_1 \cap W_2) = 1$, as required.

Now let $n = q^2 + q + 1$, and number the 1-dimensional subspaces of $V = F^3$ as $U_1, \ldots, U_n$, and the 2-dimensional subspaces as $W_1, \ldots, W_n$. Now let A_i be the set $\{j : U_j \leq W_i\}$, the set of indices of the 1-dimensional subspaces in W_i. Then

- $A_1, \ldots, A_n$ are subsets of $\{1, \ldots, n\}$, each having size $q + 1$, and any element lying in $q + 1$ of them;
- any two of $A_1, \ldots, A_n$ intersect in just one element.

So we have a configuration satisfying case (c) of the de Bruijn–Erdős theorem.

A family of sets satisfying case (c) of the de Bruijn–Erdős theorem is called a *projective plane*. The number q is called the *order* of the projective plane. So the example in the last section (the Fano plane) is a projective plane of order 2; and the construction of this section shows that there exists a projective plane of

any order q for which there is a finite field with q elements, in particular, if q is a prime number.

A theorem of Galois (which we will not prove here) says that there exists a finite field with q elements if and only if q is a power of a prime number; and moreover, such a finite field is unique. Here, for example, are the addition and multiplication tables for a field with 4 elements, called $0, 1, \alpha, \beta$:indexfield!of order 4

$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

$\cdot$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

Here is an outline of the construction. We build this field from the field $\mathbb{Z}_2$ (integers mod 2) by adding the root of an irreducible polynomial, in the same way that we construct the complex numbers from the real numbers by adding i, a root of the polynomial $x^2+1=0$. Some trial and error shows that the polynomial x^2+x+1 is irreducible over $\mathbb{Z}_2$; let α be a root of this polynomial, so that $\alpha^2+\alpha+1=0$, or $\alpha^2=\alpha+1$. (Remember that $1+1=0$ in $\mathbb{Z}_2$, and so $x+x=0$ for any x.) Then the elements of the field are all linear combinations of 1 and α, i.e. $0, 1, \alpha, \alpha+1$. We have put $\beta=\alpha+1$ in the tables. For example,

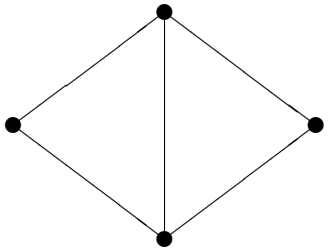
$$\begin{aligned} \alpha+\beta &= \alpha+(\alpha+1)=(\alpha+\alpha)+1=1, \\ \alpha\beta &= \alpha(\alpha+1)=\alpha^2+\alpha=\alpha+1+\alpha=1. \end{aligned}$$

For which numbers q does there exist a projective plane of order q? This seems to be one of the hardest questions in combinatorics. We have seen that they exist whenever q is a prime power. No example is known if q is not a prime power. A famous theorem called the *Bruck–Ryser Theorem* says that, if q is congruent to 1 or 2 mod 4 and q is not the sum of two squares, then there is no projective plane of order q. Thus, there is no projective plane of order 6, since 6 is congruent to 2 mod 4 but is not the sum of two squares. The number 10 is also congruent to 2 mod 4, but is the sum of two squares ($10=1^2+3^2$), so the theorem doesn't tell us whether there is a projective plane or not. But a huge computation, including two years on a Cray supercomputer, showed in 1989 that there is no projective plane of order 10. The next number in doubt is $n=12$ (this is congruent to 0 mod 4, so the Bruck–Ryser theorem does not apply to it). We have no idea whether a projective plane of order 12 exists or not!

7.5 Appendix: Proof of the Erdős–Ko–Rado Theorem

We will prove this theorem using a result about graphs. First we develop some terminology.

A *graph* consists of a set V of vertices, and a set E of edges; each edge is a set of two vertices, and is regarded as connecting those two vertices. Here is a picture of a graph.



A *clique* in a graph is a set of vertices, any two of which are joined by an edge. A *coclique* is a set of vertices, no two of which are joined by an edge. In the example drawn above, the largest clique has three vertices, and the largest coclique has two vertices.

An *automorphism* of a graph is a permutation of the vertices which maps every edge to an edge. The set of all automorphisms is a group (a subgroup of the symmtric group), called the *automorphism group* of the graph. In our example, the left-to-right reflection is an automorphism; the automorphism group contains four elements (including the identity).

A graph is said to be *vertex-transitive* if, for any two vertices, there is an automorphism of the graph which carries the first to the second. More generally, for any group G acting on any set X, we say that G acts *transitively* on X if we can carry any element of X to any other by some element of the group. Our example graph is not vertex-transitive since no automorphism can take the left-hand vertex (which lies on two edges) to the top vertex (which lies on three edges).

We use the fact that, if a group G acts transitively on a set X, with $|X| = n$, then for any $x, y \in X$, the number of elements of G mapping x to y is equal to $|G|/n$. For it is clear that the average number (if x is fixed and y varies over all points in X) is $|G|/n$; and the elements of the group which map x to y form a coset of a subgroup H, where H is the *stabiliser* of x, the set of all elements of G fixing x. By Lagrange's Theorem, all cosets contain the same number of elements.

We first prove a theorem which is too weak for our purposes, but illustrates the proof technique. Then we strengthen it to get the result we want.

Theorem 7.7 *Let Γ be a vertex-transitive graph on n vertices. Let C be a clique and D a coclique in G. Then $|C| \cdot |D| \leq n$.*

Remark Our earlier example shows that the assumption of vertex-transitivity is necessary for this theorem. The graph there has four vertices and has a 3-vertex clique and a 2-vertex coclique.

Proof Let G be the automoorphism group of the graph Γ. Count triples (x,y,g), where $x \in C$, $y \in D$, $g \in G$, and g maps x to y. There are $|C|$ choices of x, $|D|$ choices of y, and (by the remark before the theorem) $|G|/n$ choices for G, so $|C| \cdot |D| \cdot |G|/n$ such triples.

Now count another way. For each element $g \in G$, there is at most one element of C which can be mapped into D by g. For any two elements of C are joined; any two elements of D are not joined; so if g mapped two elements of C into D it would change an edge into a non-edge, which is impossible for an automorphism. So for every element of G there is at most one triple of the type we are counting.

So $|C| \cdot |D| \cdot |G|/n \leq |G|$, from which the result follows.

Theorem 7.8 *Let Γ be a vertex-transitive graph on n vertices. Let Y be a subset of the vertex set such that any clique contained in Y has size at most $m|Y|$. Then any clique in G has size at most mn.*

To see that the preceding theorem follows from this one, take X to be a coclique. Any clique contained in D has at most one vertex, i.e. size at most $m|D|$, where $m = 1/|D|$; so any clique C in G satisfies $|C| \leq mn = n/|D|$, from which the required result follows.

Proof Again let G be the automorphism group of the graph Γ. Let C be a clique. As in the preceding proof, we count triples (x,y,g), where $x \in C$, $y \in Y$, $g \in G$, and g maps x to y. As before there are $|C| \cdot |Y| \cdot |G|/n$ such triples.

Now choose first the element g. Suppose that it maps k points of C into Y. Their images form a clique, so $k \leq m|Y|$. Thus there are at most $m|Y| \cdot |G|$ such triples.

So $|C| \cdot |Y| \cdot |G|/n \leq m|Y| \cdot |G|$, whence $|C| \leq mn$.

Theorem 7.9 (Erdős–Ko–Rado) *Suppose that n and k are given, with $n \geq 2k$. Then the size of an intersecting family of k-subsets of $\{1,\ldots,n\}$ is at most $\binom{n-1}{k-1}$.*

Proof We make a graph Γ as follows: the vertices are the k-element subsets of $\{1,\ldots,n\}$; two vertices are joined if the corresponding subsets have non-empty intersection. Then an intersecting family is just a clique in this graph. Moreover, it is clear that the symmetric group on $\{1,\ldots,n\}$ acts vertex-transitively on this

graph. (This says, given any two k-element subsets A and B of $\{1,\ldots,n\}$, there is a permutation of $\{1,\ldots,n\}$ which carries A to B. Note that any permutation maps an intersecting pair of k-element sets to an intersecting pair, so is an automorphism of the graph Γ.)

Now we find a collection of n subsets such that an intersecting family contains at most k of them. To do this, we take n points $p_0, p_1, \ldots, p_{n-1}$ equally spaced around a circle. Number the intervals between the points as $I_0, \ldots, I_{n-1}$, where $I_j = (p_j, p_{j+1})$. (We read the subscripts modulo n where necessary.) The set Y consists of all "intervals of length k", that is, sets of the form $\{j, j+1, \ldots, j+k-1\}$. There are n such sets; we have to show that an intersecting family contains at most k of them. (Note that we have replaced $\{1,\ldots,n\}$ by $\{0,\ldots,n-1\}$; this does not affect the argument.)

So suppose that Z is a subset of Y, any two of whose sets intersect. Each point p_j is the endpoint of two intervals in Y, namely $\{j, j+1, \ldots, j+k-1\}$ and $\{j-k, \ldots, j-2, j-1\}$. These two sets are disjoint, because of our assumption $n \geq 2k$; so Z contains at most one of them.

Now take a set in Z, say $A = \{j, j+1, \ldots, j+k-1\}$. Any other set in Z intersects this one, so must have an end point in the set $\{p_{j+1}, \ldots, p_{j+k-1}\}$ (the set of p's which are interior points of the interval corresponding to A). Since each of these points can be the end point of at most one interval corresponding to sets in Z, there are at most $k-1$ more such sets, that is, at most k altogether.

Now it follows from Theorem 7.8 that the size of any intersecting family of k-sets (that is, any clique in the graph G) is at most

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

Remark It is possible to show that, if $n > 2k$, then any intersecting family of k-sets attaining the bound of the theorem must consist of all k-sets containing some given point of $\{1,\ldots,n\}$.

Exercises

1. Let $n = 2k$. Show that there are $2^{\binom{2k-1}{k-1}}$ intersecting families of k-element subsets of $\{1,\ldots,n\}$ having the maximum number $\binom{2k-1}{k-1}$ of members. Show that only $2k$ of them have the form $\mathscr{F}_k(i)$ for $i \in \{1,\ldots,n\}$. Hence show that the second part of the Erdős–Ko–Rado Theorem goes badly wrong when $n = 2k$.

2. Identify the right-hand picture before Theorem 7.5 with the example constructed using the finite field $F = \mathbb{Z}_2$. [*Hint*: the subspace spanned by the vector $v = (a,b,c)$ corresponds to the point with label $4a+2b+c$ in the figure.]

3. Construct a finite field with 8 elements.

4. Let $X = \{1,\ldots,n\}$. Show that, for every non-empty subset A of X, there is an intersecting family $\mathscr{F}$ of subsets of X of size 2^{n-1} with $A \in \mathscr{F}$. Show further that any two subsets A,B with $A \cap B \neq \emptyset$ are contained in a family with these properties. What about three pairwise intersecting sets?

5. Show that the largest Sperner family of subsets of $\{1,2,3,4,5\}$ containing the sets $\{1,2\}$ and $\{3,4,5\}$ contains eight sets. How does this compare with Sperner's Theorem?

6. Let $\mathscr{S}$ be the family

$$\{\{1,2,3\},\{1,4,5\},\{1,6,7\},\{2,4,6\},\{2,5,7\},\{3,4,7\},\{3,5,6\}\}$$

of subsets of $\{1,2,3,4,5,6,7\}$.

Let $\mathscr{F}$ be the family of all subsets of $\{1,\ldots,7\}$ which contain a member of $\mathscr{S}$. Show that

(a) $\mathscr{F}$ is an intersecting family;

(b) $\mathscr{F}$ contains 7 sets of size 3, 28 of size 4, 21 of size 5, 7 of size 6, and one of size 7: in all, 64 sets.

7. Let $X = \{1,2,\ldots,n\}$ and $\mathscr{S} = \{A_1,A_2,\ldots,A_b\}$ be a family of distinct subsets of X such that $|A_j \cap A_k| = 1$ for all $k \neq j$. For each $i \in X$, let r_i be the number of subsets of $\mathscr{S}$ which contain i. Prove that

$$\sum_{i=1}^{n} r_i(r_i-1) = b(b-1).$$

[*Hint*: Count the number of ordered triples (i,A_j,A_k), where A_j,A_k are distinct sets in $\mathscr{S}$ and $A_j \cap A_k = \{i\}$, in two different ways.]

Chapter 8

Systems of distinct representatives

The Students' Union has n affiliated clubs. Each club elects a delegate to the Executive Committee. The delegate must be a member of the club (s)he represents, and no person can represent more than one club.

The two natural questions for a combinatorialist are: is it possible to choose the representatives according to these rules? and In how many ways can this be done? We will answer the first question here; the second is more difficult.

Of course the answer depends on the membership of the clubs. If, for example, Sid and Doris are the only members of the Football Club, the Music Club, and the Chess Club, then the election is clearly not possible. More generally, we see that if any m clubs contain altogether less than m members, the election is not possible. So a **necessary** condition for the election is that any m clubs have between them at least m members. Surprisingly, this obvious condition also turns out to be **sufficient**; this is the content of *Hall's Theorem*.

8.1 Hall's Theorem

Let us express these ideas more mathematically. Let $A_1, A_2, \ldots, A_m$ be sets. (Suppose that they are all subsets of a universal set X.) We allow here the possibility that some of the sets are equal. A *system of distinct representatives* for the sets $(A_1, A_2, \ldots, A_n)$ is an n-tuple $(x_1, x_2, \ldots, x_n)$ of elements of X with the properties

(a) $x_i \in A_i$ for $i = 1, \ldots, n$ (this says that the elements are 'representatives' of the sets);

(b) $x_i \neq x_j$ for $i \neq j$ (this says that the representatives are 'distinct').

We abbreviate 'system of distinct representatives' to SDR.

For $J \subseteq \{1,2,\ldots,n\}$, define

$$A(J) = \bigcup_{i \in J} A_i.$$

(Do not confuse this with A_J, which we met in the chapter on PIE; A_J is the intersection of the sets with index in J, $A(J)$ is their union.) We say that the family of sets satifsies *Hall's Condition* if the following holds: for any subset J of $\{1,2,\ldots,n\}$,

$$|A(J)| \geq |J|.$$

Theorem 8.1 *Let $(A_1,\ldots,A_n)$ be a family of subsets of X. Then there exists a system of distinct representatives for $A_1,\ldots,A_n)$ if and only if Hall's condition holds.*

Proof First suppose that there is a SDR, say $(x_1,\ldots,x_n)$ for the sets. Take any subset J of $\{1,\ldots,n\}$. Then $A(J)$ contains all the sets A_i for $i \in J$, and hence contains all the representatives x_i for $i \in J$; since the representatives are distinct, we have $|A(J)| \geq |J|$. So Hall's condition holds. (This is the argument we saw earlier. If the choice of representatives is possible, then any m sets must contain at least enough members to act as their representatives.)

Now we prove the converse, which is more difficult. Let $(A_1,\ldots,A_n)$ be a family of sets satisfying Hall's condition. We have to show that an SDR can be found. Our proof will be by induction on n; we assume that a family of fewer than n sets which satisfies Hall's condition has a SDR. The induction begins with $n = 1$ since Hall's condition guarantees that A_1 is not empty [WHY?]

We say that a set $J \subseteq \{1,\ldots,n\}$ of indices is *critical* if $|A(J)| = |J|$. (Then all of the members of the sets A_i for $i \in J$ must be used as their representatives.) We divide the proof into two cases:

Case 1: No set is critical except for $\emptyset$ and possibly $\{1,\ldots,n\}$. This means that $|A(J)| > |J|$ for every non-empty proper set of indices.

Choose any element x_n of A_n to be its representative. Then x_n cannot be the representative of any other set; so we remove it. Let $A'_i = A_i \setminus \{x_n\}$ for $i = 1,\ldots,n-1$. Now for any non-empty subset J of $\{1,\ldots,n-1\}$, we have

$$|A'(J)| \geq |A(J)| - 1 > |J| - 1,$$

where the strict inequality holds by the case assumption. This means that $|A'(J)| \geq |J|$, so that the family $(A'_1,\ldots,A'_{n-1})$ satisfies Hall's condition. By the inductive hypothesis, this family has a SDR, say $(x_1,\ldots,x_{n-1})$.

But then $(x_1,\ldots,x_{n-1},x_n)$ is a SDR for the original family; for $x_n \in A_n$, and $x_n \neq x_i$ for $i < n$, since x_i belongs to A_i' but x_n doesn't.

Case 2: There is a critical set, say J. By the induction hypothesis, we can choose an SDR $(x_j : j \in J)$ for the sets indexed by J.

For $k \notin J$, let $A_k^* = A_k \setminus A_J$. (We must remove the elements of A_J: they have all been used as representatives.) We check Hall's condition for this family. For $K \subseteq \{1,\ldots,n\} \setminus J$, we have

$$\begin{aligned} |A^*(K)| &= |A(J\cup K)| - |A(J)| \\ &\geq |J\cup K| - |J| \\ &= |K|, \end{aligned}$$

where in the second line, $|A(J\cup K)| \geq |J\cup K|$ by Hall's condition, and $|A(J)| = |J|$ by the case assumption. So the sets A_k^* for $k \notin J$ satisfy Hall's Condition. By induction we can find a SDR for them, say $(x_k : k \notin J)$. Putting this together with the previously chosen SDR for the sets A_j for $j \in J$ gives a SDR for all of the sets.

This concludes the inductive step and so the proof.

There is a lot of checking to do to verify Hall's condition, though it is possible to do this in a systematic way quite efficiently. But there is one nice situation in which we can guarantee that it holds.

Proposition 8.2 *Let $(A_1,\ldots,A_n)$ be a famly of subsets of the set $\{1,\ldots,n\}$. Suppose that there is a positive number k such that*

(a) $|A_i| = k$ for $k = 1,\ldots,n$;

(b) each element of $\{1,\ldots,k\}$ is contained in exactly k of these sets.

Then the family has a SDR.

Proof We show that Hall's condition holds. Take $J \subseteq \{1,\ldots,n\}$ and count pairs (i,x) with $i \in J$ and $x \in A_i$. Clearly there are $|J|\cdot k$ such pairs. On the other hand, x can be any element of $A(J)$, and for each x there are at most k sets A_i containing x with $i \in J$ (since there are just k suchj sets altogether). So the number of pairs is at most $|A(J)|\cdot k$. Thus

$$|J|\cdot k \leq |A(J)|\cdot k,$$

and so $|A(J)| \geq |J|$.

For example, the Fano plane discussed in the last chapter of the notes consists of seven 3-element subsets of $\{1,\ldots,7\}$, so that each element of $\{1,\ldots,7\}$ lies in exactly three of them. So it has an SDR.

8.2 How many SDRs?

If the sets $A_1, \ldots, A_n$ do not satisfy Hall's condition, they have no SDR. But if they do, and they are not too small, then they have many different SDRs. This is the content of the next result.

Proposition 8.3 *Let $A_1, \ldots, A_n$ be subsets of a set X which satisfy Hall's condition. Suppose that $|A_i| \geq k$ for $i = 1, \ldots, n$, where k is a positive integer. Then the number of SDRs of the family is at least*

$$\begin{cases} k! & \text{if } k \leq n, \\ (k)_n & \text{if } k > n, \end{cases}$$

where $(k)_n$ is the falling factorial $k(k-1)\cdots(k-n+1)$.

Proof The proof follows closely the proof of Hall's Theorem. We prove the result by induction on n. If $n = 1$, there is only one set, having k elements; there are at least $(k)_1 = k$ SDRs. (Note that $k \geq n$ in this case, so we are in the second case in the statement of the result.)

So let $A_1, \ldots, A_n$ be sets satisfying the hypothesis. We divide into two cases as in the proof of Hall's Theorem.

Case 1: There are no non-empty critical sets except possibly $\{1, \ldots, n\}$. Choose any element x of A_n as its representative (there are at least k choices for x). Then the family $(A'_1, \ldots, A'_{n-1})$ consists of $n-1$ sets, each of cardinality at least $k-1$. So the number of SDRs of this family is at least

$$\begin{cases} (k-1)! & \text{if } k-1 \leq n-1, \\ (k-1)_{n-1} & \text{if } k-1 > n-1, \end{cases}$$

by the induction hypothesis. Multiplying by k gives the correct lower bound for the number of SDRs of the original family. (Note that $k \cdot (k-1)! = k!$ and $k \cdot (k-1)_{n-1} = (k)_n$.)

Case 2: There is a non-empty proper critical set, say J. We have $k \leq |A_J| = |J| \leq n$, so by the induction hypothesis the family $(A_j : j \in J)$ has at least $k!$ SDRs. As in the proof of Hall's Theorem, any such SDR can be extended to an SDR for the whole family. So there are at least $k!$ SDRs, and the induction is complete.

As a consequence, we can improve Proposition 8.2:

Proposition 8.4 *Suppose that the hypotheses of the preceding Proposition are satisfied. Then there are at least $k!$ distinct SDRs of the family of sets.*

This follows immediately from Propositions 8.3 and 8.2.

For example, the family $(\{1,2\},\{1,3\},\{2,3\})$ of sets has two SDRs, namely $(1,3,2)$ and $(2,1,3)$; the seven lines of the Fano plane have at least $3! = 6$ SDRs.

8.3 Sudoku

Hall's Marriage Theorem, and in particular the idea of a "critical set" which we met in the proof, is relevant to solving Sudoku puzzles. This is something which every Sudoku player knows to some degree.

Look at the empty cells in any row, column or subsquare of a Sudoku puzzle. Let A_i be the set of entries which could appear in the ith empty cell (i.e. those which do not already appear in the same row, column or subsquare). Then the entries which we put there must form a SDR for the sets A_i. Moreover, if we can find a critical set, then as in the proof of Hall's Theorem, we can remove its elements from the other sets, which simplifies the search for a SDR.

Here is an example.

The Times, 14 September 2005

Rating: Fiendish

					2		4	8
						2	9	
		1	9					
1				9	5	3		
		3				4		
		8	3	1				6
					8	7		
	1	5						
2	3		5					

Look at the 3×3 square in the bottom left of the puzzle. It has five empty cells, whose row and column numbers are $(1,1)$, $(1,2)$, $(1,3)$, $(2,1)$ and $(3,3)$.

Cell $(1,1)$ has 8 and 7 in the same row, 1 and 2 in the same column, and $1,2,3,5$ in the same subsquare. So the number we put there must be one of $4,6,9$. Similarly we find the possibilities for $(1,2)$ are $4,6,9$, for $(1,3)$ also $4,6,9$, for $(2,1)$ are $4,6,7,8,9$, and for $(3,3)$ are $4,6,7,9$.

Cells $(1,1)$, $(1,2)$ and $(1,3)$ form a critical set, since between them they only contain three elements $4,6,9$. So we can delete $4,6,9$ from the other sets. We find that 7 must go in $(3,3)$ and then 7 or 8 in $(2,1)$. So it must be 8 in cell $(2,1)$.

In fact, this entire puzzle can be solved by this method of finding critical sets and removing their elements from other sets.

Exercises

1. (a) Write down a SDR for the Fano plane.

(b) How many different SDRs can you find?

2. Let $A_1,\ldots,A_n$ be subsets of $\{1,\ldots,n\}$. Let M be the $n\times n$ matrix whose (i,j) entry is 1 if $j\in A_i$, and 0 otherwise. Prove that the number of SDRs of $(A_1,\ldots,A_n)$ is at least $|\det(M)|$. [*Hint:* Use the formula for the determinant as a sum over permutations. Each SDR contributes a term ±1 to the sum.]

Deduce that the Fano plane has at least 24 SDRs.

3. Construct five families, $\mathscr{F}_1$, $\mathscr{F}_2$, $\mathscr{F}_3$, $\mathscr{F}_4$, $\mathscr{F}_6$, each consisting of three subsets of the set $\{1,2,3\}$, such that $\mathscr{F}_i$ has exactly i different SDRs, for each $i\in\{1,2,3,4,6\}$.

Does there exist a family of three subsets of $\{1,\ldots,6\}$ with five SDRs?

4. This exercise gives the *deficit form of Hall's Theorem*. It is a generalisation of Hall's Theorem, but can be deduced from Hall's Theorem.

Theorem *Let $A_1,\ldots,A_n$ be subsets of a set X. Suppose that, for some positive integer m, we have*

$$|A(J)|\geq|J|-m \text{ for all } J\subseteq\{1,\ldots,n\},$$

where $A(J)=\bigcup_{j\in J}A_j$. Then it is possible to find $n-m$ of the sets $A_1,\ldots,A_n$ which have a SDR.

Prove this. [*Hint:* Take m 'dummy' elements $z_1,\ldots,z_m$, and add them to all the sets A_i.]

Chapter 9

Latin squares

A *Latin square* of order n is a $n \times n$ array, each cell containing one entry from the set $\{1, \ldots, n\}$, with the property that each element of $\{1, \ldots, n\}$ occurs once in each row and once in each column of the array.

Here is an example.

1	2	3	4	5
2	3	1	5	4
3	4	5	1	2
5	1	4	2	3
4	5	2	3	1

Notice that each row and each column is a permutation of $\{1, \ldots, n\}$.

Sometimes we use a different set of n symbols as the entries of a Latin square. The definition is just the same. For example, the *Cayley table* of a group of order n is a Latin square whose symbols are the group elements.

We will sometimes look at more general structures. Two of these are:

- A *Latin rectangle* is a $k \times n$ array (where $k \leq n$) with entries from $\{1, \ldots, n\}$ such that each symbol occurs once in each row and at most once in each column.

- A *partial Latin square* is a $n \times n$ array (where $k \leq n$) with each cell either empty or containing an symbol from $\{1, \ldots, n\}$ such that each symbol occurs at most once in each row and at most once in each column.

Thus, the first k rows of a Latin square form a Latin rectangle; and if we take a Latin square and blank out some of the entries we obtain a partial Latin square. There is no shortage of partial Latin squares; the newspapers publish examples every day!

Latin squares occur in algebra as Cayley tables of groups. If $G = \{g_1, \ldots, g_n\}$ is a finite group of order n, then its *Cayley table* is the $n \times n$ matrix whose (i, j)

entry is $g_i \circ g_j$ (the product of g_i and g_j in the group). This matrix is a Latin square – this follows from the group axioms, and is not discussed here.

9.1 Row by row

It is not difficult to show that there exists a Latin square of order n for each n. We are going to show something stronger, namely, a Latin square can be constructed by adding a row at a time; it is not possible to get stuck. The proof uses Hall's Theorem.

Proposition 9.1 *Let L be a $k \times n$ Latin rectangle, where $k < n$. Then L can be extended to a $(k+1) \times n$ Latin rectangle.*

Proof For $i = 1, \ldots, n$, let A_i be the set of symbols which do *not* occur in the ith column of L. Then $|A_i| = n - k$, since there are k distinct symbols in any column of L. Pick a symbol j. How many sets A_i contain j? We know that j occurs k times in L, once in each row; these occurrences are in k different columns, so there are $n - k$ columns in which j does not occur, that is, $n - k$ sets A_i containing j.

We have now verified the hypotheses of Proposition 8.2. From that proposition we conclude that the family $(A_1, \ldots, A_n)$ has a system of distinct representatives, say $(x_1, \ldots, x_n)$.

We claim that we can add the row $(x_1, \ldots, x_n)$ to L to obtain a larger Latin rectangle. This is true because the x_i are all distinct, so no symbol is repeated in the new row; and $x_i \in A_i$, so x_i doesn't occur in column i of L, so no repeated symbol is intriduced in any column. So the result is proved.

We can do more; we can give a lower bound for the number of Latin squares of order n.

Theorem 9.2 *The number of Latin squares of order n is at least*

$$n! \cdot (n-1)! \cdots 2! \cdot 1!.$$

Proof In the preceding proof, we replace Proposition 8.2 by the stronger Proposition 8.4, to conclude that the number of SDRs of the sets $(A_1, \ldots, A_n)$ is at least $(n-k)!$. So the number of ways of extending a $k \times n$ Latin rectangle to a $(k+1) \times n$ Latin rectangle is at least $(n-k)!$. (Each SDR gives an extension.)

Now there are $n!$ $1 \times n$ Latin rectangles (these are just permutations); each can be extended to a $2 \times n$ Latin rectangle in at least $(n-1)!$ ways; each of these can be extended to a $3 \times n$ Latin rectangle in at least $(n-2)!$ ways; and so on. The result follows.

Can we put an upper bound on the sumber of Latin squares? There is a trivial upper bound, namely n^{n^2}; this is just the number of ways we can put symbols into the n^2 cells without worrying about the rules. This can be improved slightly. Each row is a permutation, so the number of Latin squares is at most $(n!)^n$. And, having chosen the first row, each subsequent row is a derangement of it, so the number of Latin squares is at most $n! \cdot (d(n))^{n-1}$, where $d(n)$ is the derangement number (remember that $d(n)$ is approximately $n!/\mathrm{e}$).

The exact answer has been calculated for $n \leq 11$ by exhaustive search. The literature on this contains a lot of mistakes; the most reliable paper is by McKay, Meynert and Wanless in the *Journal of Combinatorial Designs* in 2007, which gives these values:

n	Number of Latin squares
1	1
2	2
3	12
4	576
5	161280
6	812851200
7	61479419904000
8	108776032459082956800
9	5524751496156892842531225600
10	9982437658213039871725064756920320000
11	776966836171770144107444346734230682311065600000

Beyond this nobody knows the exact value. Even the best known upper and lower bounds are quite a long way apart!

9.2 Youden 'squares'

Youden 'squares' form a class of designs used in statistics. As we will present them (and as they were first described by Youden) they are Latin rectangles; the name comes from a different representation used by Fisher, in which they are partial Latin squares (but we won't go into that). Essentially, a Youden 'square' is a way of representing a family of sets satisfying the conditions of Proposition 8.2 as a Latin rectangle. Strictly speaking, statisticians only use the term when an extra condition is satisfied by the family of sets, but that does not affect the result below.

Proposition 9.3 *Let $A_1, \ldots, A_n$ be subsets of $\{1, \ldots, n\}$. Suppose that, for some $k > 0$, every set A_i has k elements, and every element $x \in \{1, \ldots, n\}$ lies in k of*

the sets A_i. Then there is a $k \times n$ Latin rectangle M such that the entries in the ith column of M form the set A_i.

Example The Fano plane $123, 145, 167, 246, 257, 347, 365$ can be represented as

1	4	6	2	5	7	3
2	5	1	4	7	3	6
3	1	7	6	2	4	5

Proof By Proposition 8.2, the family of sets has an SDR $(x_1, \ldots, x_n)$, which we can take to be the first row of the rectangle. Now let $A_i' = A_i \setminus \{x_i\}$ for $i = 1, \ldots, n$. Clearly $|A_i'| = k-1$. Aos, given any $x \in \{1, \ldots, n\}$, we have used x as the representative for one of the sets A_i, so it lies in just $k-1$ of the sets A_j'. So the new family satisfies the conditions of the proposition with $k-1$ replacing k. Continue the process, with each SDR forming a new row, until $k = 0$.

9.3 Orthogonal Latin squares

Two Latin squares $A = (a_{ij})$ and $B = (b_{ij})$ are said to be *orthogonal* if they have the following property: given a pair (k, l) of symbols from the set $\{1, \ldots, n\}$, there is exactly one cell (i, j) such that $a_{ij} = k$ and $b_{ij} = l$.

Here is an example:

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

1	2	3	4
3	4	1	2
4	3	2	1
2	1	4	3

Sometimes a pair of orthogonal Latin squares is called a *Graeco-Latin square*. If we replace the symbols in the first square by Latin letters and those in the second by Greek letters, then the orthogonality condition says that every pair consisting of a Latin and a Greek letter occurs exactly once in the array. For our above example, we would get the following:

$A\alpha$	$B\beta$	$C\gamma$	$D\delta$
$B\gamma$	$A\delta$	$D\alpha$	$C\beta$
$C\delta$	$D\gamma$	$A\beta$	$B\alpha$
$D\beta$	$C\alpha$	$B\delta$	$A\gamma$

Euler posed the following question in 1782.

> Of 36 officers, one holds each combination of six ranks and six regiments. Can they be arranged in a 6×6 square on a parade ground, so that each rank and each regiment is represented once in each row and once in each column?

This question asks whether there exists a Graeco-Latin square of order 6. If so, then taking the ranks as Latin letters and regiments as Greek letters we could produce the required parade.

Euler invented Graeco-Latin squares in his researches on magic squares. A *magic square* is a $n \times n$ square containing the numbers $1, \ldots, n^2$ each once, so that the sum of the numbers in any row, column or diagonal is the same (necessarily $n(n^2+1)/2$. Euler noticed that it is often possible to construct a magic square from a Graeco-Latin square as follows:

- Replace each of the two sets of symbols by the numbers $0, 1, \ldots, n-1$.
- Regard a pair of numbers ij as being the base n representation of a single number $in + j$.
- Now the entries run from 0 to $n-1$. Add one to each so that the range is 1 to n.

The resulting square has all row and column sums constant [WHY?]. With some extra care it is possible to make the diagonal sums constant as well. Here is an example.

$C\beta$	$A\alpha$	$B\gamma$
$A\gamma$	$B\beta$	$C\alpha$
$B\alpha$	$C\gamma$	$A\beta$

21	00	12
02	11	20
10	22	01

7	0	5
2	4	6
3	8	1

8	1	6
3	5	7
4	9	2

Euler knew how to construct two orthogonal Latin squares of any order not congruent to 2 mod 4. We now outline an algebraic construction which is similar to the one Euler used.

First we give the construction using modular arithmetic. Remember that $\mathbb{Z}_n$ denotes the *integers modulo n*. We can take the elements to be $0, 1, \ldots, n-1$. To add or multiply two elements, we add or multiply in the usual way as integers, and then divide by n and take the remainder. So, in $\mathbb{Z}_7$, we have $4+5=2$, $4 \cdot 5 = 6$.

An element $a \in \mathbb{Z}_n$ is a *unit* if there exists $b \in \mathbb{Z}_n$ such that $a \cdot b = 1$. The following fact is proved in elementary algebra:

> The element a is a unit in $\mathbb{Z}_n$ if and only if $\gcd(a, n) = 1$.

Now, for any $a \in \mathbb{Z}_n$, let $L(a)$ be the matrix whose (x,y) entry is $a \cdot x + y$. Here are the matrices $L(1), L(2), L(3)$ over $\mathbb{Z}_4$:

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

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

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

We see that $L(1)$ and $L(3)$ are Latin squares but $L(2)$ is not. Moreover $L(1)$ and $L(3)$ are not orthogonal: the pair $(0,0)$ occurs twice, the pair $(0,1)$ not at all.

Proposition 9.4 *(a) $L(a)$ is a Latin square if a is a unit in $\mathbb{Z}_n$.*

(b) $L(a_1)$ and $L(a_2)$ are orthogonal if $a_1 - a_2$ is a unit.

Proof (a) In $L(a)$, if we look for symbol c in row x, we find it in column y where $ax + y = c$. This has a unique solution $y = c - ax$. Similarly, if we look for c in column y, it will be in row x is $ax + y = c$. This implies $ax = c - y$, so $x = b(c - y)$, where b is the inverse of a.

(b) Suppose we are looking for a cell in which the first square has the entry c and the second square has the entry d. Then we have to solve the simultaneous equations

$$\begin{aligned} a_1 x + y &= c, \\ a_2 x + y &= d. \end{aligned}$$

From these equations we deduce that $(a_1 - a_2)x = (c - d)$. So $x = b(c - d)$, where b is the inverse of $a_1 - a_2$. Then either of the equations can be used to find y.

This theorem is no use for constructing orthogonal Latin squares of even order. For suppose that n is even. If a is a unit in $\mathbb{Z}_n$, then $\gcd(a,n) = 1$, so a must be odd. But if a_1 and a_2 are both odd, then $a_1 - a_2$ is even, and $\gcd(a_1 - a_2, n \not= 1$.

However, for n odd, $\gcd(1,n) = \gcd(2,n) = 1$, so we get a pair of orthogonal Latin squares for every odd n.

Euler knew that it is possible to construct orthogonal Latin squares of order n also if n is a multiple of 4. Here's why. This argument shows one of the benefits of abstract algebra.

If we look at the construction we just gave, we see that there is nothing special about the integers mod n. The construction works for any *commutative ring with identity*, that is, any structure in which we can add, subtract, and multiply, so that the associative, commutative, distributive and identity laws hold. Proposition 9.4

holds in this generality. Thus, if R is a commutative ring with identity, then we can define the matrix $L(a)$ for any $a \in R$; and $L(a)$ is a Latin square if a is a unit in R, while $L(a_1)$ and $L(a_2)$ are orthogonal if $a_1 - a_2$ is a unit.

Now we use the following facts:

(a) The direct product of two commutative rings with identity is a commutative ring with identity. [The *direct product* of R_1 and R_2 is the set of all ordered pairs (r_1, r_2), with $r_1 \in R_1$ and $r_2 \in R_2$, with coordinatewise addition and multiplication.] If a_1 is a unit in R_1 with inverse b_1, and a_2 is a unit in R_2 with inverse b_2, then $(a_1, a_2)(b_1, b_2) = (1, 1)$, so (a_1, a_2) is a unit in $R_1 \times R_2$.

(b) There is a finite field (a commutative ring with identity in which every non-zero element is a unit) of every prime power order.

Using (b), we can construct orthogonal Latin squares of order 4, 8, and any larger power of 2. Then using (a), we can construct Latin squares of order $4m$, $8m$, ..., for any odd number m.

Here is an example for (b). We give first the addition and multiplication tables for a field with four elements $0, 1, \alpha, \beta$. (We saw this already in chapter 6 of the notes; there is a connection which we will see in the next section of this chapter.) Then we give the three Latin squares $L(1)$, $L(\alpha)$ and $L(\beta)$.

$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

$\cdot$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

0	1	α	β
1	0	β	α
α	β	0	1
β	α	1	0

1	0	β	α
α	β	0	1
β	α	1	0
1	0	β	α

0	1	α	β
β	α	1	0
1	0	β	α
α	β	0	1

Euler knew this, and he asked about the 36 officers because his constructions could not deal with the cases 2, 6, 10, or any number congruent to 2 mod 4. He conjectured that orthogonal Latin squares of these orders could not exist. He was right about 2 (this is easy to show directly) and 6 (this was proved by a long case-by-case argument by Tarry in 1900), but wrong about the rest. Bose, Shrikhande and Parker (the "Euler spoilers") showed in 1960 that orthogonal Latin squares of order n exist for every n except $n = 2$ and $n = 6$.

9.4 Sets of mutually orthogonal Latin squares

A *set of mutually orthogonal Latin squares* or *set of MOLS* of order n is a set $\{L_1, L_2, \ldots, L_r\}$ such that

(a) each L_i is a Latin square of order n;

(b) L_i and L_j are orthogonal if $i \neq j$.

We let $N(n)$ denote the maximum size of a set of MOLS of order n, for $n \geq 2$. [WHY NOT $n = 1$?] This is a very important function. Our conclusion from the last section can be expressed as follows:

$$N(2) = 1, \qquad N(6) = 1, \qquad N(n) \geq 2 \text{ for } n \neq 2, 6.$$

Proposition 9.5 $N(n) \leq n - 1$.

Proof Take a set $\{L_1, \ldots, L_r\}$ of MOLS of order n. We can change the symbols in each Latin square to be anything we like. So let us agree that each square has entry 1 in the top right-hand cell (in row 1 and column 1).

Now each square contains $n-1$ further entries 1. None of them can be in the first row or first column, since this would violate the Latin square condition. Also, no two of the entries 1 in different squares can be in the same cell. For any two cells L_i and L_j have entries $(1,1)$ in the top right-hand cell, so this combination cannot occur anywhere else.

So the $r(n-1)$ further entries 1 in the r squares fit in to the $(n-1)^2$ positions outside the first row and column without overlap. So we must have $r(n-1) \leq (n-1)^2$, whence $r \leq n-1$ (since $n > 1$).

Here is an important theorem of Bose about when equality can hold. Recall the definition of a *projective plane* of order n from Chapter 7: a set of m subsets of $\{1, \ldots, m\}$, where $m = n^2 + n + 1$, such that each subset contains $n+1$ elements, each element of $\{1, \ldots, m\}$ lies in $n+1$ subsets, and any two subsets intersect in exactly one point.

Theorem 9.6 *We have $N(n) = n - 1$ if and only if there is a projective plane of order n.*

The proof of this theorem is given in the next section.

If n is a prime power, we have seen that there exists a field F with n elements. Now every non-zero element of F is a unit (by definition of a field). So all the matrices $L(a)$ for $a \in F$, $a \neq 0$, are Latin squares; and any two of them are orthogonal, since if $a_1 \neq a_2$ then $a_1 - a_2$ is a unit. Thus:

Proposition 9.7 *If there exists a field of order n (that is, if n is a prime power), then* $N(n) = n-1$.

It is thought that the converse of this proposition is also true. But this is not proved, and seems to be one of the hardest open problems in combinatorics. We know that $N(6) = 2$, and that $3 \leq N(10) \leq 6$. But apart from $n = 6$, there is not a single non-prime-power value of n for which $N(n)$ is known!

9.5 Appendix: Proof of Bose's Theorem

This proof is closely connected with classical topics in projective geometry, going back to the invention of perspective by Renaissance artists and mathematicians.

Recall the statement of the theorem:

Theorem 9.8 *For an integer $n \geq 2$, the following are equivalent:*

(a) there exists a projective plane of order n;

(b) there exists a set of $n-1$ mutually orthogonal Latin squares of order n.

Proof The two constructions are just the reverse of each other.

(a) implies (b): We are given a projective plane of order n, consisting of $n^2 + n + 1$ points and the same number of $(n+1)$-element subsets called lines, so that any two lines intersect in a unique point.

Pick a line L. This will play a special role in our construction, corresponding to the 'line at infinity' where parallel lines 'meet'. Also, select two special points X and Y in L, and number the remaining points as $Z_1, \ldots, Z_{n-1}$.

The point X lies on $n+1$ lines, one of which is L. Number the others as $x_1, \ldots, x_n$. Similarly number the remaining lines through Y as $y_1, \ldots, y_n$, and the remaining lines through Z_i as $z_{i1}, \ldots, z_{in}$ for $i = 1, \ldots, n-1$.

Any point P not on L lies on a unique line x_j through X and a unique line y_k through Y. The lines x_j and y_k intersect just in $\{P\}$. We identify P with the cell (j,k) in row j and column k of an $n \times n$ grid.

Now we can define arrays $M(1), \ldots, M(n-1)$ as follows. Let P be the point corresponding to cell (j,k) as above. There is a unique ine joining P to Z_i. If this line is z_{is}, then we put symbol s in this cell iin the array $M(i)$.

Claim: $M(i)$ is a Latin square. For if the symbol s occurred twice in the same row, say in positions (j,k) and (j,l), then the lines z_{is} and x_j would have the corresponding two points in common, which is not the case.

Claim: $M(i)$ and $M(j)$ are orthogonal for $i \neq j$. For suppose we are looking for a cell containing entry s in $M(i)$ and entry t in $M(j)$. The corresponding point lies on the lines z_{is} and z_{jt}, and so is uniquely defined as their intersection. So there is just one such cell.

So, from a projective plane of order n, we have constructed a set of $n-1$ MOLS of order n.

(b) implies (a): Reverse the above construction. Here is a sketch, with the details left out. Suppose that $M(1), M(2), \ldots, M(n-1)$ be MOLS. We build a projective plane.

The *points* are of two types. First, the n^2 ordered pairs (j,k), for $1 \leq j,k \leq n$. Then $n+1$ special points $X, Y, Z_1, \ldots, Z_{n-1}$. This makes n^2+n+1 altogether, the right number.

The *lines* are of several types:

- n lines x_j for $j = 1, \ldots, n$: x_j contains the points (j,k) for $k = 1, \ldots, n$ and X.
- n lines y_k for $k = 1, \ldots, n$: y_k contains the points (j,k) for $j = 1, \ldots, n$ and Y.
- $n(n-1)$ lines z_{is} for $i = 1, \ldots, n-1$ and $s = 1, \ldots, n$: $z_{i,s}$ contains all points (j,k) for which $M(i)_{jk} = s$, and Z_i.
- Finally, a line $\{X, Y, Z_1, \ldots Z_{n-1}\}$.

One can check that this really is a projective plane of order n.

Here is an example. We start with a pair of orthogonal Latin squares of order 3 and construct a projective plane. The lines are written in the same order as in the above proof that (b) implies (a).

The Latin squares (which we have met before) are:

1	2	3
2	3	1
3	1	2

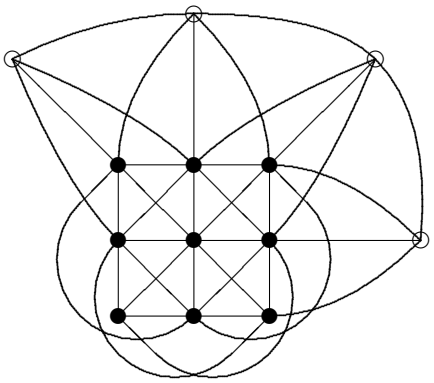
1	2	3
3	1	2
2	3	1

The 13 points of the projective plane are the pairs ij for $1 \leq i, j \leq 3$ together with

X, Y, Z_1, Z_2. The thirteen lines are:

11	12	13	X
21	22	23	X
31	32	33	X
11	21	31	Y
12	22	32	Y
13	23	33	Y
11	23	32	Z_1
12	21	33	Z_1
13	22	31	Z_1
11	22	33	Z_2
12	23	31	Z_2
13	21	32	Z_2
X	Y	Z_1	Z_2

Here it is in diagrammatic form. The lines through X and Y are the vertical and horizontal lines of the grid; the lines through Z_1 and Z_2 pass through the positions of the three symbols in the squares L_1 and L_2.



Exercises

1. Let A and B be orthogonal Latin squares of order n, which uses the symbols $0, 1, \ldots, n-1$. Construct a matrix S in which the ith entry consists of the pair (a_{ij}, b_{ij}), regarded as a two-digit number written in base n. Show that S has the properties

(a) its entries are all the integers from 0 to $n^2 - 1$ inclusive;

(b) the sum of the entries in any row or column is $n(n^2-1)$.

(This construction is due to Euler.)

2. Show that, up to permutations of rows and columns and changes in the names of the symbols, there are just two different Latin squares of order 4. Show that one, but not the other, has an *orthogonal mate* (a Latin square orthogonal to it).

3. Prove that the Latin square given by the addition table of the integers mod n has an orthogonal mate if and only if n is odd.

4. Let A be a Latin square of order n. Suppose that, for some positive integer $r < n$, only the numbers $1,\ldots,r$ occur in the first r rows and columns of A.

(a) Show that the submatrix of A formed by the first r rows and columns is a Latin square of order r.

(b) Show that $n \geq 2r$.

(c) Give an example with $r = 3$ and $n = 7$.

5. Let m be an integer greater than 1. Let $X(m)$ be the multiplication table of the non-zero integers mod m, that is, the $(m-1)\times(m-1)$ matrix defined as follows: rows and columns are indexed by $1,2,\ldots,m-1$, and the (i,j) entry is ij mod m.

Prove that $X(m)$ is a Latin square if and only if m is prime.

($**$) Does it have an orthogonal mate?

Chapter 10

Steiner triple systems

A Steiner triple system is a very special kind of family of sets. Here is the definition.

A *Steiner triple system* is a family $\mathscr{B}$ of subsets of the n-element set $X = \{1,\ldots,n\}$ with the properties

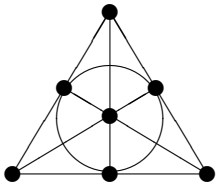
(a) every set in $\mathscr{B}$ has three elements;

(b) every two points of X are contained in exactly one member of $\mathscr{B}$.

We often call the elements of X "points" and the elements of $\mathscr{B}$ "blocks" or "triples".

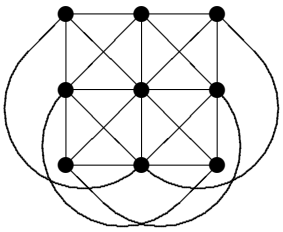
Examples Here are some examples. The first three are 'trivial', the last two are more interesting.

- Take $X = \emptyset$, $\mathscr{B} = \emptyset$.
- Take $X = \{1\}$, $\mathscr{B} = \emptyset$.
- Take $X = \{1,2,3\}$, $\mathscr{B} = \{\{1,2,3\}\}$.
- Take $X = \{1,2,3,4,5,6,7\}$ and $\mathscr{B}$ to be the Fano plane:



- Take $X = \{1,\ldots,9\}$ and arrange the points of X in a 3×3 grid. Now take $\mathscr{B}$ to consist of the horizontal and vertical lines and the positions of the six

terms in the formula for the determinant of a 3×3 matrix. Another way of saying the same thing is that these six sets are the positions of the symbols in a pair of orthogonal Latin squares of order 3.



Steiner triple systems have various uses. For example, a Latin square $L = (l_{ij})$ of order n is called *idempotent* if $l_{ii} = i$, and *totally symmetric* if $l_{ij} = k$ implies $l_{ji} = k$ and $l_{jk} = i$, for any i, j, k. Now, given a Steiner triple system $(X, \mathscr{B})$, where $X = \{1, \ldots, n\}$, we construct a matrix $L = (l_{ij})$, where $l_{ii} = i$ and, if $i \neq j$, then $l_{ij} = k$ if $\{i, j, k\} \in \mathscr{B}$. Conversely, any idempotent and totally symmetric Latin square arises in this way from a Steiner triple system.

10.1 Existence of $\mathrm{STS}(n)$

For which numbers n does there exist a $\mathrm{STS}(n)$?

Theorem 10.1 *Let $(X, \mathscr{B})$ be a Steiner triple system of order n, where $n > 0$. Then:*

(a) Any element of X is contained in $(n-1)/2$ members of $\mathscr{B}$.

(b) $|\mathscr{B}| = n(n-1)/6$.

(c) $n \equiv 1$ or $3 \bmod 6$.

Proof (a) Take $x \in X$, and let r be the number of members of $\mathscr{B}$ containing x. Count pairs (y, B) where $B \in \mathscr{B}$, $y \in X$, and $y \neq x$. There are $n-1$ points $y \neq x$, and for each such y, there is a unique set $B \in \mathscr{B}$ containing x and y; so there are $n-1$ such pairs. On the other hand, there are r choices of B containing x, and

then two choices of $y \in B$ with $y \neq x$ (since $|B| = 3$). So the number of pairs is $2r$. Thus $2r = n-1$, whence $r = (n-1)/2$.

(b) Now count pairs (x, B) where $x \in X$, $B \in \mathscr{B}$, and $x \in B$. There are n choices for x and $(n-1)/2$ choices for a set B containing it (by (a)). On the other hand, there are $|\mathscr{B}|$ choices for B, and 3 choices for a point $x \in B$. So

$$n(n-1)/2 = 3|\mathscr{B}|,$$

whence $|\mathscr{B}| = n(n-1)/6$.

(c) Since $(n-1)/2$ must be an integer, we see that n must be odd. So $n \equiv 1$, 3, or 5 mod 6. We have to exclude the last case. So suppose that $n = 6m+5$. Then, by (b),

$$|\mathscr{B}| = \frac{(6m+5)(6m+4)}{6} = \frac{(6m+5)(3m+2)}{3},$$

which is impossible since 3 does not divide $6m+5$ or $3m+2$.

We saw that there exist Steiner triple systems of orders 1, 3, 7 and 9. The above theorem shows that they do not exist for any other positive order less than 10. In order to show that there is a STS(n), we need to give a construction of one. To prove the next result, we have to give infinitely many constructions. So the proof is quite complicated! This theorem was first proved by Kirkman.

Theorem 10.2 *There exists a Steiner triple system of order n if and only if either $n = 0$ or $n \equiv 1$ or 3 mod 6.*

Proof We have seen the "only if" part already. So we have to take a number n congruent to 1 or 3 mod 6, and construct a STS(n).

For the cases where n is congruent to 3 mod 6, we can give a direct construction. For the other cases, we have to use a rather complicated recursive construction, building up large systems from smaller ones. I wish there were a simpler proof!

Case $n \equiv 3$ mod 6: Let $n = 3m$, where m is odd. We take the points of the STS to be symbols a_i, b_i, c_i, where $i = 0, \ldots, m-1$: there are $3m = n$ such points. The blocks are of two types:

(a) sets $\{a_i, b_i, c_i\}$, for $i = 0, 1, \ldots, m-1$;

(b) sets $\{a_i, a_j, b_k\}$, $\{b_i, b_j, c_k\}$ and $\{c_i, c_j, a_k\}$, where $i \neq j$ and $i + j \equiv 2k$ mod m.

All addition in this proof will be mod m, so the last condition will simply be written as $i+j=2k$. We note that, in this equation, any two of i,j,k determine the third. (Clearly i and k determine $j=2k-i$, and similarly j and k determine i. Suppose i and j are given. Since m is odd, we have $\gcd(2,m)=1$, so 2 is a unit in $\mathbb{Z}_m$; so there is a unique k satisfying $2k=i+j$, namely $k=h(i+j)$, where h is the inverse of 2 mod m.) Moreover, if the two given values are unequal, then the third value is different from both. (Clearly $i=k$ implies $j=k$. If $i=j$ then the equation reads $2i=2k$ which has the solution $i=k$.)

Clearly every block is a set of size 3, so condition (a) of the definition holds. We have to verify (b). So choose two points p and q; we have to show that there is a unique block containing them. There are several cases:

(a) $p=a_i$, $q=a_j$ ($j\neq i$). Now i and j determine a unique k such that $i+j=2k$; and $\{a_i,a_j,b_k\}$ is the unique block containing p and q.

(b) $p=b_i$, $q=b_j$, or $p=c_i$, $q=c_j$: the argument is similar.

(c) $p=a_i$, $q=b_i$: clearly there is a unique block $\{a_i,b_i,c_i\}$ of type (a) containing p and q.

(d) $p=b_i$, $q=c_i$, or $p=c_i$, $q=a_i$: the argument is similar.

(e) $p=a_i$, $q=b_k$ where $k\neq i$. There is a unique j satisfying $i+j=2k$; and $j\neq i$. So the unique block is $\{a_i,a_j,b_k\}$.

(f) $p=b_i$, $q=c_k$, or $p=c_i$, $q=a_k$, where $k\neq i$: the argument is similar.

For $n=9$, we obtain the twelve blocks

$$\{a_0,b_0,c_0\},\{a_1,b_1,c_1\},\{a_2,b_2,c_2\},\{a_0,a_1,b_2\},\{a_0,a_2,b_1\},\{a_1,a_2,b_0\},$$
$$\{b_0,b_1,c_2\},\{b_0,b_2,c_1\},\{b_1,b_2,c_0\},\{c_0,c_1,a_2\},\{c_0,c_2,a_1\},\{c_1,c_2,a_0\}.$$

Case $n\equiv 1 \bmod 6$: This case is much harder. I will give one example of a recursive construction here, and put the complete proof of the theorem in an appendix to this chapter.

Proposition 10.3 *Suppose there exists a* $\mathrm{STS}(n)$. *Then there exists a* $\mathrm{STS}(2n+1)$.

Proof Let $(X,\mathscr{B})$ be a STS(n), where $X=\{1,2,\ldots,n\}$. We take a new set

$$Y=\{a_1,\ldots,a_n,b_1,\ldots,b_n,c\},$$

with $|Y|=2n+1$, and construct a set $\mathscr{C}$ of blocks; we show that these blocks form a STS$(2n+1)$.

The blocks are of two types:

(a) $\{a_i,b_i,c\}$, for $i=1,\ldots,n$.

(b) For every block $\{i,j,k\}\in\mathscr{B}$, the four blocks $\{a_i,a_j,b_k\}$, $\{a_j,a_k,b_i\}$, $\{a_k,a_i,b_j\}$, and $\{b_i,b_j,b_k\}$.

Clearly every block contains three points, so (a) of the definition holds. Take two points p and q; we have to show that just one block contains them. There are several cases:

(a) $p=c$, $q=a_i$: a unique block $\{a_i,b_i,c\}\in\mathscr{C}$ contains p and q.

(b) $p=c$, $q=b_i$, or $p=a_i$, $q=b_i$: the argument is similar.

(c) $p=a_i$, $b=a_j$ with $i\neq j$: there is a unique block $\{i,j,k\}\in\mathscr{B}$ containing i and j, and then a unique block $\{a_i,a_j,b_k\}\in\mathscr{C}$ containing p and q.

(d) $p=a_i$, $q=b_j$, with $i\neq j$: the argument is similar.

(e) $p=b_i$, $q=b_j$, with $i\neq j$: there is a unique block $\{i,j,k\}\in\mathscr{B}$ containing i and j, and then a unique block $\{b_i,b_j,b_k\}\in\mathscr{C}$ containing p and q.

For $n=3$, starting with the single block $\{1,2,3\}$, we obtain seven blocks

$$\{a_1,b_1,c\},\{a_2,b_2,c\},\{a_3,b_3,c\},\{a_1,a_2,b_3\},\{a_2,a_3,b_1\},\{a_3,a_1,b_2\},\{b_1,b_2,b_3\}$$

of STS(7).

This method constructs Steiner triple systems of orders 7, 15, 19, 31, ..., but leaves several values undecided, such as 13, 25, 29, 37, ... These will be settled in the Appendix. The general principle is always the same (we build larger systems out of smaller ones) except in one case, where we have to give a direct construction: this is $n=13$, where we can take the point set to be $\{0,\ldots,12\}$ (the integers mod 13), and the blocks to be

$$\begin{aligned}&\{0,1,4\},\{1,2,5\},\{2,3,6\},\{3,4,7\},\{4,5,8\},\{5,6,9\},\{6,7,10\},\\&\{7,8,11\},\{8,9,12\},\{0,9,10\},\{1,10,11\},\{2,11,12\},\{0,3,12\},\\&\{0,2,8\},\{1,3,9\},\{2,4,10\},\{3,5,11\},\{4,6,12\},\{0,5,7\},\{1,6,8\},\\&\{2,7,9\},\{3,8,10\},\{4,9,11\},\{5,10,12\},\{0,6,11\},\{1,7,12\}\end{aligned}$$

Note that it is not necessary to remember all these blocks. We start with the two blocks $\{0,1,4\}$ and $\{0,2,8\}$, and produce the rest of the system by adding x to each of their elements mod 13, for $x = 1,\ldots,12$. (This process is called the *development* of the two blocks mod 13.)

There is a simple test for a starting set of blocks in this construction:

Proposition 10.4 *Let $B_1,\ldots,B_k$ be 3-element subsets of $\mathbb{Z}_n$. Then the development of $\{B_1,\ldots,B_k\}$ is a Steiner triple system if and only if every non-zero element of $\mathbb{Z}_n$ has a unique representation in the form $x-y$, where $x,y \in B_i$ for some i with $1 \le i \le k$. If this holds, then $n = 6k+1$.*

Proof We will prove it one way round. Suppose that the condition of the Proposition holds, and let u,v be distinct elements of $\mathbb{Z}_n$. We want to show that there are unique elements i,z with $1 \le i \le k$ and $z \in \mathbb{Z}_n$ such that $u,v \in B_i + z$. If this is to hold, we must have $u-z, v-z \in B_i$. But $(u-z)-(v-z) = u-v$, and by assumption there are unique i,x,y such that $u-v = x-y$ with $x,y \in B_i$. So we must have $u-z = x$ and $v-z = y$, whence z and i are uniquely determined.

Moreover, the $n-1$ non-zero elements must be given by the 6 differences for each of the k blocks, so $n-1 = 6k$, as required.

In the above example, we have the following expressions for elements of $\mathbb{Z}_{13}$ as differences from $\{0,1,4\}$ and $\{0,2,8\}$:

$$\begin{array}{cccc} 1=1-0 & 2=2-0 & 3=4-1 & 4=4-0 \\ 5=0-8 & 6=8-2 & 7=2-8 & 8=8-0 \\ 9=0-4 & 10=1-4 & 11=0-2 & 12=0-1 \end{array}$$

For a simpler example, we get $\mathrm{STS}(7)$ as the development of a single block $\{0,1,3\}$ in $\mathbb{Z}_7$:

$$\begin{array}{ccc} 1=1-0 & 2=3-1 & 3=3-0 \\ 4=0-3 & 5=1-3 & 6=0-1 \end{array}$$

10.2 Kirkman's schoolgirls

Despite their name, Steiner triple systems were invented, not by Steiner, but by Kirkman; he gave the definition and proved Theorem 10.2 several years before Steiner published a paper asking whether or not they exist. The reason we do not call them "Kirkman triple systems" is that this name has been used for something a bit different.

Kirkman posed the following problem:

Fifteen schoolgirls go for a walk every day for a week in five rows of three. Is it possible to arrange the walks so that every two girls walk together exactly once during the week?

Let $X = \{1,\ldots,15\}$, and let $\mathscr{B}$ be the set of all groups of three girls who walk together during the course of the week. Then the terms of the problem require that $(X,\mathscr{B})$ is a Steiner triple system. But there is more structure. The $15\cdot 14/6 = 35$ blocks of the Steiner triple system must have a partition into seven sets of five (corresponding to the days of the week) such that the five blocks in each part of the partition themselves form a partition of X.

We say that a Steiner triple system $(X,\mathscr{B})$ is *resolvable*, or is a *Kirkman triple system*, if the set $\mathscr{B}$ can be partitioned into subsets $\mathscr{B}_1,\ldots,\mathscr{B}_r$ such that $\mathscr{B}_i$ is a partition of X for $i = 1,\ldots,r$.

Here is an example of a Kirkman triple system of order 9. The twelve blocks are arranged into four rows forming the required partition.

$$\begin{aligned}
\mathscr{B}_1 &: \ \{\{1,2,3\},\{4,5,6\},\{7,8,9\}\}\\
\mathscr{B}_2 &: \ \{\{1,4,7\},\{2,5,8\},\{3,6,9\}\}\\
\mathscr{B}_3 &: \ \{\{1,5,9\},\{2,6,7\},\{3,4,8\}\}\\
\mathscr{B}_4 &: \ \{\{1,6,8\},\{2,4,9\},\{3,5,7\}\}
\end{aligned}$$

Since the n points must be partitioned by the blocks, each of which has size 3, we see that n must be divisible by 3 for such a system to exist. Since we know that n is odd, we conclude that n must be congruent to 3 mod 6.

Now each class $\mathscr{B}_i$ contains $n/3$ blocks. Since there are $n(n-1)/6$ blocks altogether, we see that the number of classes is equal to $(n-1)/2$. This can be verified another way. We know that each point lies in $(n-1)/2$ blocks; but exactly one of these blocks belongs to each class $\mathscr{B}_i$, so there must be $(n-1)/2$ classes.

Kirkman himself constructed a solution to the problem with $n = 15$ (his original 'schoolgirls problem'). It took 120 years before the general case was finally solved by Ray-Chaudhuri and Wilson in the 1970s. They proved the following theorem (which is too complicated for this course!)

Theorem 10.5 *For $n > 0$, there exists a Kirkman triple system of order n if and only if $n \equiv 3$ mod 6.*

10.3 Appendix: Proof of Kirkman's Theorem

Kirkman's Theorem states that a Steiner triple system of order n exists if and only if $n = 0$ or n is congruent to 1 or 3 mod 6. We have seen that this condition is necessary, and we have to show that it is sufficient: in other words, if n satisfies the congruence condition, then we can construct a STS of order n.

This is a good example of a combinatorial construction. It contains both a direct part (for numbers congruent to 3 mod 6, which we saw already) and a recursive part; the recursive part shows that one small case ($n = 13$) remains to be dealt with (we already gave a construction for this value); and as we will see below, it requires the construction of an 'auxiliary' structure for use in the main construction (this is a STS containing a subsystem of order 7, which we now define).

Let $(X, \mathscr{B})$ be a Steiner triple system, and let Y be a subset of X. We say that Y is a *subsystem* if, for any two points $y_1, y_2 \in Y$, the unique block of $(X, \mathscr{B})$ containing y_1 and y_2 is contained in Y. If Y is a subsystem, and $\mathscr{C}$ is the set of blocks which are contained in Y, then $(Y, \mathscr{C})$ is a Steiner triple system in its own right. Given any Steiner triple system $(X, \mathscr{B})$ with $|X| \geq 3$, there are some obvious subsystems which always exist: the empty set; any 1-element set $\{x\}$; and any block of $\mathscr{B}$.

Our main recursive construction is given by the following result:

Theorem 10.6 *Let $(X, \mathscr{B})$ be a Steiner triple system of order v containing a subsystem Y of order u, and let $(Z, \mathscr{D})$ be a Steiner triple system of order w. Then there exists a Steiner triple system of order $u + w(v-u)$. Moreover, if $0 < u < v$ and $w > 1$, then we may assume that this system contains a subsystem of order 7.*

Remark If we take $v = 3$ and $u = 1$ (that is, $(X, \mathscr{B})$ consists of a single block and the subsystem is a single point), then we obtain a Steiner triple system of order $1 + 2w$. This is precisely the construction of Proposition 10.3. So the above theorem generalises that proposition.

Proof We take the point set of the new system to be $Y \cup ((X \setminus Y) \times Z)$, which does indeed have $u + (v-u)w$ points, since $|Y| = u$, $|X \setminus Y| = v - u$, and $|Z| = w$. We set $m = v - u = |X \setminus Y|$, and number the points of $X \setminus Y$ with the elements of $\mathbb{Z}_m$.

The blocks of the new system are of the following types:

(a) All blocks contained in Y.

(b) All blocks of the form $\{y, (x,z), (x',z)\}$, where $y \in Y$, $x, x' \in X \setminus Y$, $z \in Z$, and $\{y, x, x'\} \in \mathscr{B}$.

(c) All blocks of the form $\{(x,z), (x',z), (x'',z)\}$, where $x, x', x'' \in X \setminus Y$, $z \in Z$, and $\{x, x', x''\} \in \mathscr{B}$.

(d) All blocks of the form $\{(x_i,z),(x_j,z'),(x_k,z'')\}$, where $x_i,x_j,x_k \in X \setminus Y$, $z,z',z'' \in Z$. $\{z,z',z''\}$ is a block of $\mathscr{D}$, and $i+j+k=0$ in $\mathbb{Z}_m$. [Recall that points of $X \setminus Y$ are indexed by elements of $\mathbb{Z}_m$.]

Now we have to show that any two points lie in a unique block. There are several cases:

- Two points of Y lie in a unique block of type (a).
- A point of Y and a point of $(X \setminus Y) \times Z$ lie in a unique block of type (b).
- Two points of $(X \setminus Y) \times Z$ with the same Z-coordinate lie in a unique block of type either (b) or (c).
- Two points of $(X \setminus Y) \times Z$ with different Z-coordinates lie in a unique block of type (d). [Note that any two of i,j,k uniquely determine the third.]

So we do have a STS.

For the last part, assume that $u \neq 0$. Then u and v are odd, so $m = v-u$ is even. Choose a block of $\mathscr{B}$ of the form $\{y,x,x'\}$, where $y \in Y$ and $x,x' \in X \setminus Y$. Then choose the indexing of $X \setminus Y$ by $\mathbb{Z}_m$ such that $x = x_0$ and $x' = x_{m/2}$. Let $\{z,z',z''\}$ be a block of $\mathscr{D}$. Then the seven points

$$y,(x_0,z),(x_0,z'),(x_0,z''),(x_{m/2},z),(x_{m/2},z'),(x_{m/2},z'')$$

and the seven blocks

$$\{y,(x_0,z),(x_{m/2},z)\},\{y,(x_0,z'),(x_{m/2},z')\},\{y,(x_0,z''),(x_{m/2},z'')\}$$
$$\{(x_0,z),(x_{m/2},z'),(x_{m/2},z'')\},\{(x_{m/2},z),(x_0,z'),(x_{m/2},z'')\},$$
$$\{(x_{m/2},z),(x_{m/2},z'),(x_0,z'')\},\{(x_0,z),(x_0,z'),(x_0,z'')\}$$

form a subsystem. [Note that $0+0+0 = 0+(m/2)+(m/2) = 0$ in $\mathbb{Z}_m$.]

Example Earlier, we were unable to construct STS of orders 25 or 37. With Theorem 10.6, we can now construct these, using

$$25 = 1+3(9-1), \qquad 37 = 1+3(13-1).$$

(This is shorthand for saying: there exists a STS(9) containing a STS(1) subsystem, and also a STS(3), so by Theorem 10.6 there is a STS$(1+3(9-1))$, and similarly for the other one.

We now turn to the proof of Kirkman's Theorem. I will write the proof as an argument by contradiction. That is, let $A(n)$ be the statement that a STS(n) exists. We say that n is *admissible* if n is congruent to 1 or 3 mod 6. I will assume that n is the smallest admissible number for which an STS(n) does not exist, and deduce a contradiction. It is necessary to have another induction going on at the same time. Let $B(n)$ be the statement that there exists a STS(n) containing a subsystem of order 7. I will prove that if n is congruent to 9 mod 12 and $n \geq 15$, then $B(n)$ holds.

First let us consider $A(n)$. Suppose that $A(m)$ is true for all admissible numbers $m < n$.

- If n is congruent to 3 mod 6, then the direct construction in Theorem 10.2 gives an STS(n). So we only have to deal with numbers congruent to 1 mod 6.

- If n is congruent to 7 mod 12, then $n = 12k+7 = 1+2(6k+3)$. By assumption, $A(6k+3)$ holds; then $A(12k+7)$ follows from Proposition 10.3. So we only have to deal with numbers congruent to 1 mod 12.

- We separate these according to their congruence mod 36.

  - If n is congruent to 1 mod 36, then $n = 36k+1 = 1+3(12k+1-1)$, and $A(12k+1)$ is true, so $A(36k+1)$ is true.
  - If n is congruent to 25 mod 36, then $n = 36k+25 = 1+3(12k+9-1)$, and $A(12k+9)$ is true, so $A(36k+25)$ is true.
  - If n is congruent to 13 mod 36, then $n = 36k+13 = 7+3(12k+9-7)$. so we need a STS$(12k+9)$ with a subsystem of order 7, that is, we need to know that $B(12k+9)$ is true, and $A(36k+13)$ will follow. We are going to prove this for $k \geq 1$. For $k = 0$, we gave a direct construction of STS(13) on pp. 99–100. (Note that we already know that an STS of order $12k+9$ exists by Theorem 10.2, but that one doesn't have the required subsystem.)

So we have to prove $B(12k+9)$ for $k \geq 1$. Again we split into congruences mod 36.

- If n is congruent to 9 mod 36, then $n = 36k+9 = 3+(9-3)(6k+1)$. So $A(6k+1)$ together with Theorem 10.6 give the result.

- If n is congruent to 21 mod 36, then $n = 36k+21 = 3+(9-3)(6k+3)$. So $A(6k+3)$ together with Theorem 10.6 gives the result.

- If n is congruent to 33 mod 36, then $n = 36k+33 = 3+(12k+13-3)3$, So $A(12k+13)$ together with Theorem 10.6 give the result.

Although phrased as a "minimal counterexample' argument, this proof is essentially constructive. For example, how to construct a STS(625)? We have

- 625 is congruent to 13 mod 36, so we write $625 = 7+3(213-7)$; we need a STS(213) with a subsystem of order 7.
- 213 is congruent to 33 mod 36, so we write $213 = 3+3(73-3)$; we need a STS(73).
- 73 is congruent to 1 mod 36, so we write $73 = 1+3(25-1)$; we need a STS(25).
- We already saw how to construct this: write $25 = 1+3(9-1)$, so we need a STS(9), which of course we know (it is given by the direct construction at the start of Theorem 10.2).

In fact, we could construct STS(625) more easily by using the fact that $625 = 25^2 = 0+25(25-0)$ and the existence of STS(25). But a general proof cannot rely on lucky accidents like this!

Exercises

1. Suppose that an STS(v) of order v on a set X has a subsystem Y of order u, with $u < v$. Show that $v \geq 2u+1$. [*Hint:* Let x be a point not in Y. Show that the triples containing x and a point of Y are all distinct.]

2. Prove directly that if an STS(v) and an STS(w) exist, then an STS(vw) exists. Show further that, if $v,w \geq 3$, then we can find a STS(vw) containing a subsystem of order 9.

3. Let $\mathbb{Z}_2$ denote the integers mod 2. Let X be the set of all **non-zero** vectors in the n-dimensional vector space $(\mathbb{Z}_2)^n$. Let $\mathscr{B}$ be the set of all triples $\{x,y,z\}$ of vectors of X satisfying $x+y+z = 0$.

(a) Prove that $(X,\mathscr{B})$ is a Steiner triple system of order 2^n-1.

(b) Identify the Fano plane as a Steiner triple system of this form.

4. Let $X = \{1,\ldots,n\}$ and suppose that $\mathscr{B}$ is a collection of 3-element subsets of X with the property that any two members of $\mathscr{B}$ intersect in **at most** one element.

(a) Let r_x denote the number of members of $\mathscr{B}$ containing the element $x \in \{1,\ldots,n\}$. Show that $r_x \leq \lfloor (n-1)/2 \rfloor$.

(b) By counting pairs (x,B), with $x \in X$ and $B \in \mathscr{B}$, show that

$$|\mathscr{B}| = \frac{\sum_{x\in\{1,\ldots,n\}} r_x}{3}.$$

(c) Deduce that

$$|\mathscr{B}| \leq \left\lfloor \frac{n}{3} \left\lfloor \frac{n-1}{2} \right\rfloor \right\rfloor.$$

(d) Hence or otherwise show that, if $n = 6$, then $|\mathscr{B}| \leq 4$.

(e) Find an example of four 3-element subsets of $\{1,\ldots,6\}$ which satisfy the hypothesis $|B_i \cap B_j| \leq 1$ for $B_i, B_j \in \mathscr{B}$, $i \neq j$.

Appendix A

Solutions to odd-numbered exercises

Chapter 1

1. $1001 = 7 \cdot 11 \cdot 13$. So the numerator has to contain numbers divisible by all three primes. So $n \geq 13$. A little trial and error shows that

$$1001 = \frac{14 \cdot 13 \cdot 12 \cdot 11}{4 \cdot 3 \cdot 2 \cdot 1} = \binom{14}{4}.$$

3. We show first that $\sum_{k=0}^{n} k(k-1)\binom{n}{k} = n(n-1)2^{n-2}$. This can be shown in two ways:

(a) Take the Binomial Theorem $\sum_{k=0}^{n} \binom{n}{k} x^k = (1+x)^n$, differentiate twice, and put $x = 1$.

(b) Use the fact that $k(k-1)\binom{n}{k} = n(n-1)\binom{n-2}{k-2}$ for $k \geq 2$. (The terms for $k = 0, 1$ are zero.) Now sum over k.

Thus

$$\sum_{k=0}^{n} k^2 \binom{n}{k} = \sum_{k=0}^{n} k(k-1)\binom{n}{k} + \sum_{k=0}^{n} k\binom{n}{k} = n(n-1)2^{n-2} + n2^{n-1} = n(n+1)2^{n-2}.$$

5. Here it is for n congruent to 0 mod 8. In this case, we have

$$(1+\mathrm{i})^n = 2^{n/2}.$$

Equating real and imaginary parts gives

$$\begin{aligned}\sum_{j=0}^{\lfloor n/2\rfloor}(-1)^j\binom{n}{2j} &= 2^{n/2},\\ \sum_{j=0}^{\lfloor (n-1)/2\rfloor}(-1)^j\binom{n}{2j+1} &= 0.\end{aligned}$$

So, if S_0,S_1,S_2,S_3 denote the sums of binomial coefficients for k congruent to $0,1,2,3$ mod 4 respectively, we have

$$S_0-S_2=2^{n/2}, \qquad S_1-S_3=0.$$

Since we already know that

$$S_0+S_2=S_1+S_3=2^{n-1},$$

we conclude that

$$S_0=2^{n-2}+2^{(n-2)/2}, \quad S_1=2^{n-2}, \quad S_2=2^{n-2}-2^{(n-2)/2}, \quad S_3=2^{n-2}.$$

As a check, when $n=8$, we have

$$S_0=1+70+1=72, \quad S_1=8+56=64, \quad S_2=28+28=56, \quad S_3=56+8=64.$$

7. (a) $\dfrac{n-k}{k+1}\binom{n}{k}=\dfrac{(n-k)\cdot n(n-1)\cdots(n-k+1)}{(k+1)\cdot k(k-1)\cdots 1}=\binom{n}{k+1}$.

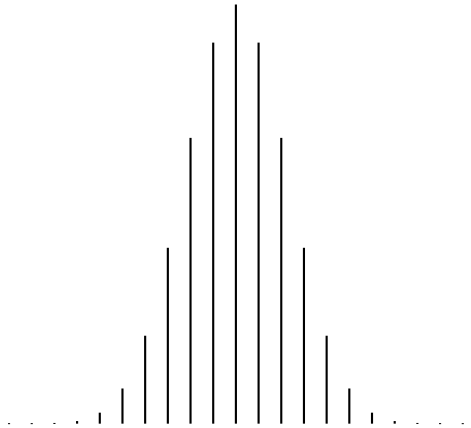
(b) The ratio of $\binom{n}{k+1}$ to $\binom{n}{k}$ is $(n-k)/(k+1)$. This ratio is $>,=,<1$ according as $n-k$ is $>,=,<k+1$, that is, as n is $>,=,<2k+1$.

(c) By part (b), the binomial coefficients increase until the point where $n=2k+1$ (if this occurs), at which point they remain constant for one step and then decrease. This happens if n is odd. If n is even, then there is no value of k for which $n=2k+1$, so the binomial coefficients increase until $k=n/2$ and then decrease.

(d) This follows immediately from (c).

(e) Suppose that $n=2m$. Then $\binom{2m}{m}$ is the largest of the $2m+1$ binomial coefficients $\binom{2m}{0}$, …, $\binom{2m}{2m}$. Now the sum of these binomial coefficients is 2^{2m} (see Property 1 in section 1.3 of the notes). The largest binomial coefficient is smaller than the sum, and larger than the average. This gives the stated result.

Here is a chart of the binomial coefficients $\binom{20}{k}$, for $k=0,\ldots,20$.



We have

$$\frac{2^{20}}{21} = 49932.19\ldots, \qquad \binom{20}{10} = 184756, \qquad 2^{20} = 1048576.$$

Chapter 2

1. (a) $7^5 = 16807$; 7650 (see below) ; $3^3 \cdot 4^2 + 4^3 \cdot 3^2 = 1008$.

The argument for the second number goes like this. There are three choices for which of $1,2,3$ can occur; so we do the calculation for the case that 1 and 2 but not 3 occur and multiply by 3.

Suppose that 1 and 2 occur in k positions. There are $\binom{5}{k}$ ways to choose these positions; $2^k - 2$ ways to fill them with 1s and 2s (we are not allowed to use all 1s or all 2s) and 4^{5-k} ways to fill the remaining positions with $4,5,6,7$. So the total is

$$(2^5 - 2) + 5(2^4 - 2)4 + 10(2^3 - 2)4^2 + 10(2^2 - 2)4^3 = 2550.$$

In (c) there are two terms in the sum because the sequence might begin with an even number or an odd number. In the first case, there are three even numbers which can be chosen from $\{2,4,6\}$ in 3^3 ways, and two odd numbers which can be chosen from $\{1,3,5,7\}$ in 4^2 ways. The other term is similar.

(b) $(7)_5 = 2520$; 1440 (see below); $(3)_3 \cdot (4)_2 + (4)_3 \cdot (3)_2 = 216$.

For the first and third parts, simply replace the formula n^k by $(n)_k$ everywhere.

For the second part, there are 3 choices for which of $1,2,3$ to use, and $5 \cdot 4 = 20$ choices for their positions; the remaining entries are filled from the other four numbers, in $(4)_3$ ways. So there are 1440 such sequences.

3. (a) m^n; (b) $(m)_n$; (c) $n!$ (we must have $m = n$ in this case); (d) see Chapter 6.

5. We have

$$W(n) = 1 + n + n(n-1) + n(n-1)(n-2) + \cdots + n!.$$

Every term except the first two contains the product of two consecutive integers and so is even. So the parity of $W(n)$ is the same as that of $1+n$, that is, even if n is odd and *vice versa*.

Chapter 3

1. Divide up the permutations π according to the value of the smallest number k for which π maps $\{1,\ldots,k\}$ into itself. Note that k takes values $1,2,\ldots,n$, and that $k = n$ if and only if π is indecomposable. (On the other hand, $k = 1$ if and only if π fixes 1.) Then π induces an indecomposable permutation on the set $\{1,\ldots,k\}$, and an arbitrary permutation on the set $\{k+1,\ldots,n\}$. So there are $g(k)(n-k)!$ permutations with a given value of k. Summing over k, we get all permutations, so the total is $n!$.

Consider the product $F(x)(1-G(x))$. The constant term is $1 \cdot 1 = 1$. For $n > 1$, the coefficient of x^n is obtained by taking the term in x^{n-k} from $F(x)$ and the term in x^k from $1-G(x)$ for $k = 1,\ldots,n$, together with one more term which is the term in x^n from $F(x)$ and the constant term from $1-G(x)$. So the coefficient of x^n in the product is

$$\sum_{k=1}^{n} -g(k)(n-k)! + n! = 0.$$

So $F(x)(1-G(x)) = 1$.

Chapter 4

1. Clearly $s(1) = 1$.

An expression with sum n could simply be n. Otherwise, if the last term is $n-k$ (where $k = 1,\ldots,n-1$), then the terms before the last sum to k, and form an arbitrary expression summing to k. So we have

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

Now for $n > 1$, we see that

$$s(n-1) = 1 + \sum_{k=1}^{n-2} s(k),$$

and so

$$s(n) = 1 + \sum_{k=1}^{n-2} s(k) + s(n-1) = 2s(n-1).$$

This recurrence with the initial condition $s(1) = 1$ has solution $s(n) = 2^{n-1}$, as is easily shown by induction.

3. (a) We have seen that the answer is the nth Fibonacci number $F(n)$.

(b) Suppose that I hand over k coins of value 2 pence. Then we must have $0 \le k \le \lfloor n/2 \rfloor$, and I must also include $n-2k$ coins of value 1 penny. So the total number of ways of paying is the number of choices of k, which is $1 + \lfloor n/2 \rfloor$.

5. It is clear that a_n is a power of 2 for all n. So put $a_n = 2^{b_n}$. Then we find that $b_0 = 1$ and $b_n = 2b_{n-1}$ for all n. So $b_n = 2^n$, and we conclude that $a_n = 2^{2^n}$.

7. Begin by recalling the recurrence relation for the derangement numbers:

$$d_0 = 1, d_1 = 0, \qquad d_n = (n-1)(d_{n-1} + d_{n-2}) \text{ for } n \ge 2.$$

(a) Induction on n. For $n = 1$, $d_1 = 0 = 1 + (-1)^1$. so the result holds. Now assume that $d_{n-1} = (n-1)d_{n-2} + (-1)^{n-1}$. Then

$$\begin{aligned} d_n &= (n-1)d_{n-1} + (n-1)d_{n-2} \\ &= (n-1)d_{n-1} + d_{n-1} - (-1)^{n-1} \\ &= nd_{n-1} + (-1)^n, \end{aligned}$$

so it holds for n. Thus the formula is proved for all n.

(b) Induction on n. The formula is clear for $n = 0$. Suppose that it holds for $n-1$, that is, $d_{n-1} = (n-1)! \sum_{k=0}^{n-1} (-1)^k/k!$. Then

$$\begin{aligned} d_n &= nd_{n-1} + (-1)^n \\ &= n! \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} + \frac{(-1)^n n!}{n!} \\ &= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \end{aligned}$$

as required.

(c) We have

$$(\mathrm{e}^{-x})(1-x)^{-1} = \left(\sum_{n \ge 0} \frac{(-1)^n x^n}{n!} \right) \left(\sum_{n \ge 0} x^n \right).$$

The coefficient of x^n in this product is

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

by (b); but this is the same as the coefficient of x^n in $D(x)$.

Chapter 5

1. (i) We have $S(2,2) = 1$ and

$$S(n,2) = S(n-1,1) + 2S(n-1,2) = 1 + 2S(n-1,2)$$

for $n \geq 3$. By induction we now get $S(n,2) = 2^{n-1} - 1$ for all $n \geq 2$. (The induction starts at $n = 2$. Assuming $S(n-1,2) = 2^{n-2} - 1$, we have $S(n,2) = 1 + 2(2^{n-2} - 1) = 2^{n-1} - 1$.)

We have $S(2,1) = 1$ and

$$S(n,n-1) = S(n-1,n-2) + (n-1)S(n-1,n-1) = S(n-1,n-2) + (n-1).$$

Again the required result follows by induction, the details of which are left to you.

(ii) The set $\{1,\ldots,n\}$ has 2^n subsets. Two of these (the empty set and the whole set) cannot occur as parts in a partition with two parts. Each of the other $2^n - 2$ sets occurs with its complement in a unique partition. So the number of partitions is $(2^n - 2)/2 = 2^{n-1} - 1$.

A partition with $n-1$ parts must have one part of size 2 and all the rest of size 1. There are $\binom{n}{2}$ ways to choose the part of size 2; all the other points lie in parts of size 1.

A partition with $n-2$ parts either has two parts of size 2 or one of size 3, with all remaining parts of size 1. So we have

$$\begin{aligned} S(n,n-2) &= \binom{n}{2}\binom{n-2}{2}\Big/2 + \binom{n}{3} \\ &= \frac{n(n-1)(n-2)(3n-1)}{24}. \end{aligned}$$

(The division by 2 in the first term is because the two parts of size 2 can be chosen in either order yielding the same partition.)

3. (a) See the calculation of the table of values of Stirling numbers of the first kind on p. 54 of the text. We see that $s(6,3) = -225$, so the number of permutations is 225 (the minus sign indicates that they are all odd permutations).

The possible cycle lengths are three positive numbers summing to 6; these are $[1,1,4]$, or $[1,2,3]$, or $[2,2,2]$.

There are $\binom{6}{4} = 15$ ways to choose the four points to form the 4-cycle, and $3! = 6$ possible 4-cycles on this set. So 90 permutations with cycle lengths $[1,1,4]$.

There are $\binom{6}{3} = 20$ choices of three points for a 3-cycle, and two 3-cycles on this set. Then $\binom{3}{2}$ choices of two points for the 2-cycle, and the rest is determined uniquely. So 120 permutations with cycle lengths $[1,2,3]$.

There are $\binom{6}{2}\binom{4}{2}\binom{2}{2} = 90$ choices of three 2-cycles; but they could be chosen in any order, so we have to divide by $3! = 6$, giving 15 permutations with cycle lengths $[2,2,2]$.

Total $90 + 120 + 15 = 225$.

Chapter 6

1. The percentage of people satisfied with none of the candidates is

$$100 - 65 - 57 - 58 + 28 + 30 + 27 - 12 = -7,$$

which is impossible. So the data is incorrect.

Chapter 7

1. Let $n = 2k$. Out of each complementary pair of sets of size k, $\mathscr{F}$ contains exactly one. The resulting family satisfies

$$|\mathscr{F}| = \frac{1}{2}\binom{2k}{k} = \binom{2k-1}{k-1},$$

since it contains one out of each complementary pair of k-sets by construction. Moreover, $\mathscr{F}$ is an intersecting family. For take $A, B \in \mathscr{F}$; then A and B cannot be disjoint, since disjoint k-sets are complementary and we took just one out of each complementary pair.

There are $\frac{1}{2}\binom{2k}{k} = \binom{2k-1}{k-1}$ complementary pairs of k-sets, and so 2 raised to this power number of choices of one set from each complementary pair. So this is the number of maximum intersecting families of this form.

Clearly there are only $2k$ sets of the form $\mathscr{F}_k(i)$, since $i = 1, \ldots, n = 2k$. This is much much smaller than the number of families just constructed. Even for $n = 6$, we constructed $2^{10} = 1024$ families of which only six are of the form $\mathscr{F}_n(i)$.

The Erdős–Ko–Rado Theorem says that every intersecting family of k-subsets of an n-set of maximum size $\binom{n-1}{k-1}$ has the form $\mathscr{F}_k(i)$ if $n > 2k$. The above construction shows that this is not true if $n = 2k$.

3. By trial and error, or otherwise, the polynomial $x^3 + x + 1$ is irreducible over $\mathbb{Z}_2$, so we adjoin a root α of this polynomial. Elements of the resulting field have the form $a + b\alpha + c\alpha^2$, where $a, b, c \in \mathbb{Z}_2$; so there are eight elements, namely

$$0, 1, \alpha, \alpha+1, \alpha^2, \alpha^2+1, \alpha^2+\alpha, \alpha^2+1.$$

Addition is done by adding the coefficients a, b, c mod 2, so that, for example,

$$(\alpha^2+1) + (\alpha^2+\alpha) = \alpha + 1.$$

To multiply, we use the fact that $\alpha^3 = \alpha + 1$ (since α is a root of $x^3 + x + 1 = 0$), and so $\alpha^4 = \alpha^2 + \alpha$. So for example

$$\begin{aligned}(\alpha^2+1)\cdot(\alpha^2+\alpha) &= \alpha^4 + \alpha^3 + \alpha^2 + \alpha \\ &= \alpha^2 + \alpha + \alpha + 1 + \alpha^2 + \alpha \\ &= \alpha + 1.\end{aligned}$$

5. Any further set must contain at least one of 1 and 2 (if not, it would be contained in $\{3,4,5\}$) and can't contain both (or it would contain $\{1,2\}$). Similarly, such a set must contain at least one of $3, 4, 5$ but cannot contain them all. So to get such a set we must include one of $\{1\}$ and $\{2\}$, and one of the six subsets $\{3\}$, $\{4\}$, $\{5\}$, $\{3,4\}$, $\{3,5\}$, $\{4,5\}$, and take their union. This gives $2 \cdot 6 = 12$ possible sets.

However, we cannot take all of these sets. If we take $\{1,3\}$, then $\{1,3,4\}$ and $\{1,3,5\}$ are not permitted. In fact, we can take at most six of these sets: for if we arrange the six sets containing 1 in pairs

$$(\{1,3\},\{1,3,4\}), (\{1,4\},\{1,4,5\}), (\{1,5\},\{1,3,5\}),$$

and similarly for the sets containing 2, we can choose at most one of each pair. So we can't have more than $2 + 6 = 8$ sets altogether.

However, we can obtain a Sperner family with 8 sets, by combining 1 with the 1-element subsets of $\{3,4,5\}$, and 2 with the 2-element subsets:

$$\{1,2\}, \{3,4,5\}, \{1,3\}, \{1,4\}, \{1,5\}, \{2,3,4\}, \{2,3,5\}, \{2,4,5\}.$$

7. Follow the hint.

On one hand, if we choose the point i, there are r_i choices of A_j containing it, and $(r_i - 1)$ choices of A_k (which must be different from A_j), so $r_i(r_i - 1)$ such triples beginning with i. To count all the triples, we have to sum this expression over all values of i from 1 to n.

On the other hand, if we choose A_j and A_k first, there are b choices for A_j and $(b-1)$ choices for A_k; then $|A_j \cap A_k| = 1$, so there is just one choice for i belonging to both these sets. So the number of triples is $b(b-1)$.

Equating these two expressions gives the result.

In the family $\mathscr{S}$ of Question 6, each point lies in three sets of $\mathscr{S}$. So each term in the sum is $3 \cdot 2 = 6$, and the sum is $7 \cdot 6 = 42$. On the other hand, $b = 7$, so the right-hand side is also $7 \cdot 6 = 42$.

Chapter 8

1. Take the Fano plane to have the sets $123, 145, 167, 246, 257, 347, 356$. Clearly $(1,4,6,2,7,3,5)$ is an SDR.

There are in all 24 SDRs for the Fano plane. Our arguments for this will be based to some extent on 'symmetry'. Thus, for example, we can choose any element of 123 to be its representative, so 'by symmetry' we may choose 1. Again by symmetry, we can choose either of 4 and 5 as the representative of the second set, and either of 6 and 7 as the representative of the third. Suppose that we choose 4 and 6. Removing these representatives from the last four sets gives

$$2, 257, 37, 35.$$

We must use 2 as representative of the fourth set. Then it is easy to see that there are just two choices for the other three, namely $(5,7,3)$ or $(7,3,5)$.

So altogether there are $3 \cdot 2 \cdot 2 \cdot 2 = 24$ SDRs.

3.

(a) Clearly $\{\{1\},\{2\},\{3\}\}$ has just one SDR.

(b) $\{\{1,2\},\{1,2\},\{3\}\}$ has two SDRs, since both 1 and 2 and be representatives for the first two sets.

(c) $\{\{1,2\},\{2,3\},\{1,2,3\}\}$ has three SDRs. Check that there are three choices of representatives for the first two sets; then the unused element can be the representative of the last set.

(d) $\{\{1,2\},\{1,2,3\},\{1,2,3\}\}$ has four SDRs. For there are two choices for the representative of the first set; if 1 is chosen, then the remaining elements of the other two sets are $\{\{2,3\},\{2,3\}$, with two SDRs, and similarly if 2 is chosen.

(e) For $\{\{1,2,3\},\{1,2,3\},\{1,2,3\}\}$, any permutation of $(1,2,3)$ will be a valid SDR.

Suppose that we have a family of three subsets of $\{1,2,3\}$ with five SDRs, say all except $(1,2,3)$. Then since $(1,3,2)$, $(2,1,3)$ and $(3,2,1)$ are all SDRs, we see that each of $1,2,3$ can appear as representative of each set, so each of the three sets must be $\{1,2,3\}$. But then $(1,2,3)$ would be a SDR, after all.

Chapter 9

1. (a) Since A and B are orthogonal, each pair (i,j) for $i,j = 0,\ldots,n-1$ occurs just once. These pairs represent in base n all the numbers from $00 = 0$ to $(n-1)(n-1) = (n-1)n+(n-1) = n^2-1$, each once.

(b) In each row or column, each of the ns digits $0,\ldots,n-1$ occurs once, and each of the units digits $0,\ldots,n-1$ occurs once. So the row or column sum is

$$n(0+\cdots+(n-1))+(0+\cdots+(n-1)) = (n+1)n(n-1)/2,$$

as required.

3. The addition table of $\mathbb{Z}_n$ is the Latin square $L(1)$ defined in this chapter. If n is odd, then $L(2)$ is a Latin square orthogonal to it.

Suppose that n is even, and suppose (for a contradiction) that M is a Latin square orthogonal to $L(1)$. We begin by observing that the sum of all the elements in $\mathbb{Z}_n$ is $n/2$ if n is even. (For the sum in $\mathbb{Z}$ is $n(n-1)/2$, and n is even, so n divides $n^2/2$.)

Look at the positions (x_i,y_i) in which a given symbol occurs in M. Then

- Each row occurs once as x_i, so $\sum x_i = n/2$.
- Each column occurs as y_i, so $\sum y_i = n/2$.
- The (x_i,y_i) entry of $L(1)$ is x_i+y_i. Since $L(1)$ is orthogonal to M, each symbol occurs once as x_i+y_i, so $\sum(x_i+y_i) = n/2$.

But this is a contradiction, since $n/2+n/2 = 0 \neq n/2$.

5. Suppose that m is not a prime; say $m = ab$, where $1 < a,b < m$. Then

$$(a+1)b = ab+b \equiv b = 1b \bmod m,$$

so the element b occurs twice in column b, in rows 1 and $a+1$. Thus $M(m)$ is not a Latin square.

Conversely, suppose that m is prime. In this case we show that $M(m)$ is a Latin square with entries $1,\ldots,m-1$.

(a) First note that all entries belong to this set. For if the (i,j) entry were not in the set $\{1,\ldots,m-1\}$, then it would necessarily be zero; so m would divide ij, contrary to the fact that m is a prime which does not divide either i or j.

(b) Suppose that the same entry x occurs twice in a row, say in positions (i,j) and (i,k), with $j<k$. Then $ij \equiv ik \equiv x \bmod m$; so $i(k-j) \equiv 0 \bmod m$. This is impossible for the same reason as in (a).

(c) A very similar argument to (b) shows that an entry cannot occur twice in a column.

The last part of the question is more difficult. There is a theorem of algebra saying that, if p is prime, there exists a *primitive root* mod p, that is, an element g whose powers give all the non-zero elements of $\mathbb{Z}_p$. Thus the multiplicative group of $\mathbb{Z}_p$ is isomorphic to the additive group of $\mathbb{Z}_{p-1}$. Now by the solution to Question 3 above, we see that the Cayley table of this group does not have an orthogonal mate if p is odd.

Chapter 10

1. Follow the hint. We know that the number of triples of the STS containing the point x is $(v-1)/2$ (Theorem 10.1(a)). For each $y \in Y$, there is a unique triple containing x and y. No two of these triples are equal, since if the same triple contained $\{x,y_1\}$ and $\{x,y_2\}$ for $y_1,y_2 \in Y$, then this triple would contain two points of Y and hence would be contained in Y, contradicting the fact that it contains x. So $|Y| = u \leq (v-1)/2$, whence $v \geq 2u+1$.

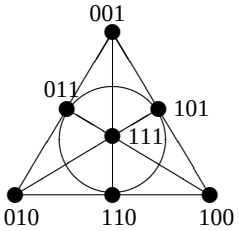
3. (a) We have to show that, given any two distinct non-zero vectors x and y, the unique solution of $x+y+z=0$ (namely $z=-(x+y)$) is non-zero and is not equal to either x or y. It will then follow that x and y lie in a unique triple in $\mathscr{B}$.

Note that $-x=x$ for any vector in $(\mathbb{Z}_2)^n$, so we can write $z=x+y$. Now:

- If $z=0$ then $x+y=0$, so $y=-x=x$, contrary to assumption.
- If $z=x$, then $y=0$, contrary to assumption.
- If $z=y$, then $x=0$, contrary to assumption.

Finally, the order of the STS is the number of non-zero vectors, which is $2^n - 1$.

(b) All seven equations $x+y+z=0$ can be checked from the picture:



Appendix B

Miscellaneous problems

1. Show that the coefficient of x^n in $(x+x^2)^k$ is equal to $\binom{n-k}{k}$.

   Hence show that the coefficient of x^n in $(1-x-x^2)^{-1}$ is equal to $\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-k}{k}$.

   Deduce that

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

2. Prove that $\binom{2^a+b}{2b+1}$ is odd if and only if $b = 2^a - 1$. [This exercise is due to Thomas Müller.]

3. How many words can be made using some or all (possibly none) of the letters of the word MAMMAL?

4.

   (a) How many permutations of $\{1,\ldots,9\}$ are there?

   (b) How many of them consist of a single cycle?

   (c) How many of them have exactly three cycles, none of which is of length 1?

5.

   (a) In how many ways can 25 sweets be distributed to a class of 12 children?

   (b) How many ways are there if each child is to have at least one sweet?

   (c) How many ways are there if each child is to have at least two sweets?

6. Solve the recurrence relation and initial conditions

$$a_0 = 2, a_1 = 4, a_2 = 7, \qquad a_n = 4a_{n-1} - 5a_{n-2} + 2a_{n-3} \text{ for } n \geq 3.$$

7. Let $B(n)$ be the nth Bell number, the number of partitions of $\{1,\ldots,n\}$. Prove directly that $B(n) \leq n!$ for all n.

8. Let F_n be the nth Fibonacci number. Prove by induction that

$$F_n^2 - F_{n-1}F_{n+1} = (-1)^n$$

for $n \geq 1$.

9. Let $S(n,k)$ be the Stirling number of the second kind (the number of partitions of $\{1,\ldots,n\}$ into k parts, and let

$$F_k(x) = \sum_{n \geq k} S(n,k)x^n.$$

(a) Prove that $F_1(x) = 1/(1-x)$.

(b) Write down a recurrence relation for $S(n,k)$.

(c) Use it to show that

$$F_k(x) = \frac{x}{1-kx}F_{k-1}(x)$$

for $k \geq 1$.

(d) Hence show by induction on k that

$$F_k(x) = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}$$

for $k \geq 1$.

10. Show that an permutation is even if and only if it has an even number of cycles of even length (with no restriction on cycles of odd length).

11. Let $A_1,\ldots,A_n$ be subsets of a set X. For $J \subseteq \{1,\ldots,n\}$, let

$$A_J = \bigcap_{i \in J} A_i$$

be the set of elements which lie in the sets A_i for $i \in J$ (and possibly in some other sets as well). Let B_J be the set of elements which lie in the sets A_i for $i \in J$, and

do not lie in the sets A_k for $k \notin J$. Use the Principle of Inclusion and Exclusion to show that

$$|B_J| = \sum_{K \supseteq J} (-1)^{|K \setminus J|} A_K.$$

12. Let B be the matrix obtained from Pascal's Triangle by moving the rows right so that the left-hand side is vertical. So the entry in row n and column k is $\binom{n}{k}$. Let B^* be the matrix in which the entry in row n and column k is $(-1)^{n-k}\binom{n}{k}$. Prove that the matrices B and B^* are inverses of each other, by finding two bases for the space of all polynomials such that B and B^* are the transition matrices. [Hint: The Binomial Theorem.]

13.

(a) Show that there does not exist a Latin square orthogonal to the square

1	2	3	4	5
2	1	4	5	3
3	4	5	1	2
4	5	2	3	1
5	3	1	2	4

[*Hint:* Suppose that B is orthogonal to the given square and has entry 1 in position $(1,1)$. Where can the other 1s in B occur?]

(b) Write down two orthogonal Latin squares of order 5.

14. Let $F = \mathbb{Z}_3$, the integers mod 3. Let $V = F^n$ be the vector space of all n-tuples of elements of F. Let

$$\mathscr{B} = \{\{x,y,z\} : x,y,z \in \mathscr{B}, x,y,z \text{ distinct }, x+y+z = 0\}.$$

Show that $(V, \mathscr{B})$ is a Steiner triple system of order 3^n.

15. Let $\mathscr{F}$ be an intersecting family of subsets of $X = \{1,2,\ldots,n\}$.

(a) Show that $|\mathscr{F}| \leq 2^{n-1}$.

(b) Show that, if $|\mathscr{F}| = 2^{n-1}$, then for any subset A of X, either $A \in \mathscr{F}$, or $X \setminus A \in \mathscr{F}$.

(c) Show that, if $|\mathscr{F}| = 2^{n-1}$, $A \in \mathscr{F}$, and B is a subset of X containing A, then $B \in \mathscr{F}$.

16. Let $\mathscr{F}$ be an intersecting family of 2-element subsets of $\{1,\ldots,n\}$. Show that *either*

(a) there is an element $x \in \{1,\ldots,n\}$ contained in every set in $\mathscr{F}$; *or*

(b) $\mathscr{F} = \{\{x,y\},\{y,z\},\{x,z\}\}$ for some $x,y,z \in \{1,\ldots,n\}$.

17. Let

$$\mathscr{F} = \{123,456,789,147,258,369,159,267,348\},$$

where, for example, 123 means $\{1,2,3\}$.

(a) Prove directly that $\mathscr{F}$ satisfies Hall's condition.

(b) Find a 3×9 Latin rectangle whose columns are the sets of $\mathscr{F}$.

Index