MAS334 COMBINATORICS

1. Counting sets

We begin with a very simple point, which is easy to get slightly wrong (Google "off-by-one error").

Definition 1.1. In this course, we will rarely be using real numbers. We will therefore use interval notation to refer to intervals of integers:

$$[n,m] = \{k \in \mathbb{Z} \mid n \leq k \leq m\} \qquad (n,m] = \{k \in \mathbb{Z} \mid n < k \leq m\}$$
$$[n,m) = \{k \in \mathbb{Z} \mid n \leq k < m\} \qquad (n,m) = \{k \in \mathbb{Z} \mid n < k < m\}.$$

For example, we have

$$[3,7] = \{3,4,5,6,7\} \qquad (3,7] = \{4,5,6,7\}$$
$$[3,7) = \{3,4,5,6\} \qquad (3,7) = \{4,5,6\}.$$

The sizes of these sets are

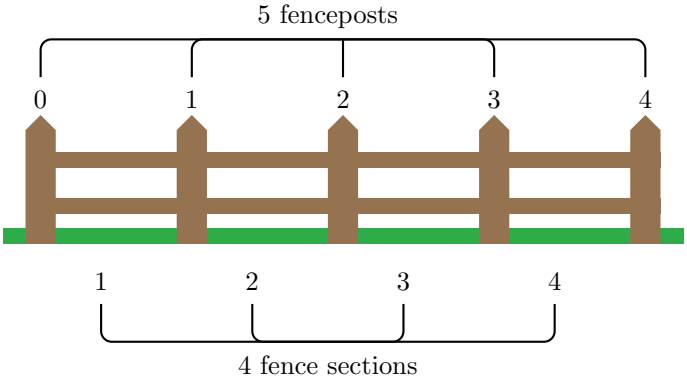
$$|[n,m]| = m-n+1 \qquad |(n,m]| = m-n$$
$$|[n,m)| = m-n \qquad |(n,m)| = m-n-1.$$

(The first three of these are valid for $n \leq m$, but the last is only valid for $n < m$.) For example, we have

$$|[3,7]| = 5 = 7-3+1 \qquad |(3,7]| = 4 = 7-3$$
$$|[3,7)| = 4 = 7-3 \qquad |(3,7)| = 3 = 7-3-1.$$

It is a common mistake to say that $|[n,m]| = m-n$ or $|(n,m)| = m-n$, but the above examples show that this is not correct.

Remark 1.2. Here is a related observation: if we have a fence consisting of n sections supported by fenceposts, then the number of posts is one more than the number of sections. Each section has a post at the right hand end, and there is one more post at the left hand end of the whole fence. (Google "fencepost error".)



Definition 1.3. A *binary sequence* of length n is a sequence $a = (a_1, \ldots, a_n)$ with $a_i \in \{0,1\}$. We write B_n for the set of binary sequences of length n. We also write B_{nk} for the subset of binary sequences of length n in which there are k ones.

Example 1.4.

- The sequence $a = (0,1,0,1,1,0)$ is a binary sequence of length 6, so $a \in B_6$. We will typically use abbreviated notation and write $a = 010110$ instead of $a = (0,1,0,1,1,0)$. As there are 3 ones in a, we can also say that $a \in B_{63}$.

- The full list of elements of B_3 is
$$B_3 = \{000, 001, 010, 011, 100, 101, 110, 111\}.$$
(We have written these in dictionary order, which is good practice. It is much easier to deal with these kind of constructions if we list things in a systematic and consistent order.)
- The full list of elements of B_{53} is
$$B_{53} = \{00111, 01011, 01101, 01110, 10011, 10101, 10110, 11001, 11010, 11100\}.$$

In the above example we saw that $|B_3| = 8 = 2^3$. Of course, this can be generalised.

Proposition 1.5. $|B_n| = 2^n$.

Proof. To choose an element $a = (a_1, \dots, a_n) \in B_n$ we have 2 choices for a_1, 2 choices for a_2 and so on, making $2 \times 2 \times \cdots \times 2 = 2^n$ choices for the sequence as a whole. □

Definition 1.6. Let A be a finite set. We let PA denote the set of all subsets of A. We also let P_kA denote the set of all subsets of size k in A.

Example 1.7. Take $A = \{a, b, c\}$. Then
$$PA = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$
We also have
$$P_2A = \{\{a, b\}, \{a, c\}, \{b, c\}\}.$$
We might also use more abbreviated notation:
$$PA = \{\emptyset, a, b, c, ab, ac, bc, abc\}.$$

Proposition 1.8. *If $|A| = n$ then $|PA| = 2^n$.*

Proof. Interactive demo

List the elements of A as $x_1, \dots, x_n$. To choose a subset of A, we first choose whether to include x_1, then choose whether to include x_2 and so on. We have two choices for each x_i, and thus 2^n choices altogether.

Here is another way to say essentially the same thing. Given a binary sequence $a = (a_1, \dots, a_n)$, we define
$$U_a = \{x_i \mid a_i = 1\}.$$
For example, in the case $n = 6$, we have
$$U_{111000} = \{x_1, x_2, x_3\} \qquad U_{100011} = \{x_1, x_5, x_6\}.$$
(In the left hand example, we have ones in positions 1, 2 and 3, so the set is $\{x_1, x_2, x_3\}$. In the right hand example, we have ones in positions 1, 5 and 6, so the set is $\{x_1, x_5, x_6\}$.) This construction gives a one-to-one correspondence between subsets of A and binary sequences, so $|PA| = |B_n| = 2^n$. □

Definition 1.9. For a finite set A, we define F_kA to be the set of sequences $a = (a_1, \dots, a_k)$ such that the entries a_i are distinct elements of A.

Example 1.10. If $A = \{a, b, c\}$ then
$$F_2A = \{ab, ac, ba, bc, ca, cb\}$$
$$F_3A = \{abc, acb, bac, bca, cab, cba\}.$$

Proposition 1.11. *If $|A| = n$ and $0 \leq k \leq n$ then*
$$|F_kA| = \prod_{i=0}^{k-1}(n - i) = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$
In particular, we have $|F_nA| = n!$, but F_kA is empty for $k > n$.

Proof. Suppose we want to choose a sequence $a = (a_1, \dots, a_k) \in F_k A$. Then a_1 can be any element of A, so there are n choices. Then a_2 can be any element of A other than a_1, so there are $n-1$ choices. Then a_3 can be any element other than a_1 and a_2, so there are $n-2$ choices, and so on. At the last stage, a_k can be any element of A except for $a_1, \dots, a_{k-1}$, so there are $n-(k-1) = n-k+1$ choices. Thus, the overall number of choices is

$$|F_k A| = n(n-1)\cdots(n-k+1) = \prod_{i=0}^{k-1}(n-i).$$

Note also that

$$\begin{aligned} n! &= n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 \\ &= n \times (n-1) \times \cdots \times (n-k+1) \times (n-k) \times (n-k-1) \cdots \times 2 \times 1 \\ (n-k)! &= (n-k) \times (n-k-1) \cdots \times 2 \times 1 \\ n!/(n-k)! &= n \times (n-1) \times \cdots \times (n-k+1) = |F_k A|. \end{aligned}$$

In particular, we have $|F_n A| = n!/0! = n!$. In the other hand, if $k > n$, it is clear that we cannot have a list of k distinct elements in A, because A has only n elements; so $F_k A = \emptyset$. □

Definition 1.12. For integers n, k with $0 \leq k \leq n$ we define

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

For $k < 0$ or $k > n$ we define $\binom{n}{k} = 0$. In this course, we will consider $\binom{n}{k}$ to be undefined for $n < 0$.

Corollary 1.13. *If $|A| = n$, then $|P_k A| = \binom{n}{k}$.*

Proof. If $k > n$ it is clear that $P_k A$ is empty so $|P_k A| = 0 = \binom{n}{k}$. Suppose instead that $0 \leq k \leq n$. Every list $a = (a_1, \dots, a_k) \in F_k A$ gives a subset $U_a = \{a_1, \dots, a_k\} \in P_k A$, and every subset of size k arises in this way. However, we can reorder the list a in $k!$ different ways, and they all give the same subset. Thus, we have

$$|P_k A| = \frac{|F_k A|}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

□

Corollary 1.14. *We also have $|B_{nk}| = \binom{n}{k}$.*

Proof. To specify an element of B_{nk}, we just need to specify the k positions in $[1, n]$ where the ones appear. There are $\binom{n}{k}$ subsets of size k in $[1, n]$, so $|B_{nk}| = \binom{n}{k}$. □

Problem 1.15. *Suppose that 6 people compete in an Olympic pie-eating competition. In how many ways can the medals be awarded? If the BBC decides to interview three of the finalists, chosen at random, in how many ways can they do that? What if there were 100 finalists?*

Solution. Interactive demo

Let A be the set of competitors, so $|A| = 6$. For the first question, we need an ordered list of three distinct medal winners (gold, then silver, then bronze), so the number of possibilities is $|F_3 A| = 6 \times 5 \times 4 = 120$. In more detail, there are 6 choices for who gets the gold. When we have awarded the gold, there are 5 choices left for who gets silver, then 4 choices for who gets bronze. Thus, the total number of ways in which the medals can be awarded is $6 \times 5 \times 4 = 120$.

For the second question, we need an unordered set of three interviewees, so the number of possibilities is $|P_3 A| = \binom{6}{3} = 20$. In more detail, there are 120 possible choices for the list of people who get interviewed, in the order in which they get interviewed. But we do not care about the order, we only care about the set of interviewees. So we need to divide by the number of possible orders, which is $3! = 6$. Thus, the number of ways to choose a set of three interviewees is $120/6 = 20$.

If there were 100 finalists, then the number of ways of awarding the medals would be $100 \times 99 \times 98 = 970200 \simeq 9.7 \times 10^5$, and the number of ways of choosing the interviewees would be $(100 \times 99 \times 98)/6 = 161700 \simeq 1.6 \times 10^5$.

Problem 1.16. *In the National Lottery, six balls are drawn from a set of 59 balls. How many possible outcomes are there?*

Solution. We need to count the subsets of size 6 in a set of size 59; the answer is

$$\binom{59}{6} = \frac{59 \times 58 \times 57 \times 56 \times 55 \times 54}{6!} = 45057474 \simeq 4.5 \times 10^7.$$

The other familiar place where we see binomial coefficients is in the binomial expansion formula:

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

We next recall how this works.

Example 1.17. Interactive demo

Consider the case $n = 4$. We have

$$\begin{aligned}
(1+x)^4 =& (1+x)(1+x)(1+x)(1+x) \\
=& 1111 + 111x + 11x1 + 11xx + 1x11 + 1x1x + 1xx1 + 1xxx + \\
& x111 + x11x + x1x1 + x1xx + xx11 + xx1x + xxx1 + xxxx \\
=& 1111 + \\
& 111x + 11x1 + 1x11 + x111 + \\
& 11xx + 1x1x + 1xx1 + x11x + x1x1 + xx11 + \\
& 1xxx + x1xx + xx1x + xxx1 + \\
& xxxx \\
=& 1 + 4x + 6x^2 + 4x^3 + x^4 = \sum_{k=0}^{4} \binom{4}{k} x^k.
\end{aligned}$$

In the first step we have just expanded everything out in the obvious way, writing the terms in dictionary order. Each term is a product of four factors, each of which is either 1 or x. To generate all the terms, we have to make 4 choices of whether to have a 1 or an x, giving $2^4 = 16$ terms altogether. In the second step, we just regroup the terms according to how many x's appear. There is one term with no x's, 4 terms with one x, 6 terms with two x's, 4 terms with three x's and one term with four x's. In general, to generate a term with k x's, we just need to choose k slots from 4 in which the x's appear, and put ones in the other slots. Thus, there are $\binom{4}{k}$ terms with k x's, and each of these contributes x^k to the expansion. Thus, we have $(1+x)^4 = \sum_k \binom{4}{k} x^k$.

As an exercise in notation, we can write this slightly differently. Let A be a subset of $\{1, \dots, n\}$. Let t_A be the term in the expansion where we take x from the factors corresponding to $i \in A$, and 1 from the factors corresponding to $i \notin A$. The number of x's is then equal to $|A|$, so the product is $x^{|A|}$. We get a term for every possible subset $A \subseteq \{1, \dots, n\}$, so we get

$$(1+x)^n = \sum_A t_A = \sum_A x^{|A|}.$$

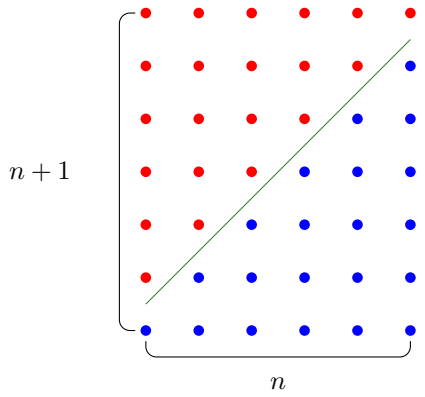
The number of x^k's in this sum is the number of subsets A such that $|A| = k$, or in other words $\binom{n}{k}$. We therefore have $(1+x)^n = \sum_k \binom{n}{k} x^k$ as before.

Proposition 1.18. *For any $n \geq 0$, we have $1 + 2 + \cdots + n = \binom{n+1}{2}$.*

Proof. Put $S_n = 1 + 2 + \ldots + (n-1) + n$. We can rewrite this with the terms in reverse order as $S_n = n + (n-1) + \cdots + 2 + 1$. Adding these two equations together, we get

$$2S_n = (1+n) + (2 + (n-1)) + \cdots + ((n-1) + 2) + (n+1).$$

The right hand side consists of n terms, each of which is equal to $n+1$, so the total is $(n+1)n$. It follows that $S_n = (n+1)n/2 = \binom{n+1}{2}$ as claimed. This proof can be illustrated as shown below: there are S_n red dots above the diagonal line and S_n blue dots below it, showing that $2S_n = (n+1)n$.



Video

Alternatively, we can give a proof by induction. For $n = 0$ the claim is that $0 = \binom{1}{2}$, which is clear. For $n = 1$ the claim is that $1 = \binom{2}{2}$, which is also clear. For $n > 1$, we can assume as an induction hypothesis that

$$S_{n-1} = 1 + 2 + \cdots + (n-1) = \binom{n}{2} = n(n-1)/2 = \tfrac{1}{2}n^2 - \tfrac{1}{2}n.$$

Adding n to both sides, we get

$$S_n = (1 + 2 + \cdots + (n-1)) + n = \tfrac{1}{2}n^2 - \tfrac{1}{2}n + n = \tfrac{1}{2}(n^2 + n) = \binom{n+1}{2},$$

as required. □

Proposition 1.19. *For $n, k \in \mathbb{N}$ with $(n,k) \neq (0,0)$ we have* $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

Note that we have explicitly excluded the case $n = k = 0$. If $k = 0$ and $n > 0$ then the claim is that $1 = 1 + 0$ which is true. If $k > n$ then the claim is that $0 = 0 + 0$ which is true. If $k = n > 0$ then the claim is that $1 = 0 + 1$ which is true. This just leaves the interesting case where $0 < k < n$. We will give two different proofs for this case.

Bijective proof. Interactive demo

The binomial coefficient $\binom{n}{k}$ is the number of subsets $A \subseteq [1,n]$ with $|A| = k$. To choose such a subset, we first decide whether we want n to be an element of A. If we decide that n should not be an element of k, then we just choose A to be a subset of size k in $[1, n-1]$, and there are $\binom{n-1}{k}$ possibilities for this. If we decide that we do want n to be an element of A, then we need to choose a further $k-1$ elements from $[1, n-1]$ to make up the rest of A, and there are $\binom{n-1}{k-1}$ possibilities for this. Thus, we have $\binom{n-1}{k} + \binom{n-1}{k-1}$ possibilities for A, and this must agree with the number $\binom{n}{k}$ that we obtained more directly. □

Algebraic proof. Recall that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. On the top we have

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

We can also write the n here as $(n-k)+k$, giving $n! = (n-k) \times (n-1)! + k \times (n-1)!$. By substituting this into the definition of $\binom{n}{k}$, we get

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

In the first term, we can rewrite $k!$ as $k \times (k-1)!$, and in the second term, we can rewrite $(n-k)!$ as $(n-k)(n-k-1)!$. (These are valid because we are assuming that $0 < k < n$, so $k, n-k > 0$.) This gives

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

Video

□

Proposition 1.20. *For $0 \leq k \leq n$ we have $\binom{n}{k} = \binom{n}{n-k}$.*

Bijective proof. Interactive demo

The binomial coefficient $\binom{n}{k}$ is the number of subsets $A \subseteq [1,n]$ size k in $[1,n]$. To choose such a subset, we can just choose a subset $B \subseteq [1,n]$ of size $n-k$ in $[1,n]$, and take $A = B^c$. This gives a one-to-one correspondence between subsets of size k and subsets of size $n-k$, so $\binom{n}{k} = \binom{n}{n-k}$. □

Algebraic proof.

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

□

Definition 1.21. A subset $A \subseteq [1,n]$ is *gappy* if there are no adjacent elements. In more detail, the condition is that there should not exist $a \in [1, n-1]$ such that $a \in A$ and $a+1 \in A$. Similarly, we say that a binary sequence is *gappy* if it has no adjacent ones. We write G_{nk} for the set of gappy subsets $A \subseteq [1,n]$ with $|A| = k$.

Example 1.22. The set $A = \{1,5,7,11\} \subseteq [1,20]$ is gappy. The set $B = \{1,5,6,11\}$ is not gappy, because it contains the adjacent elements 5 and 6. The full list of elements of G_{73} is

$$G_{73} = \{135, 136, 137, 146, 147, 157, 246, 247, 257, 357\},$$

so $|G_{73}| = 10$.

Proposition 1.23. *If $n \geq 2k-1$ then $|G_{nk}| = \binom{n-k+1}{k}$, but if $n < 2k-1$ then $G_{nk} = \emptyset$ so $|G_{nk}| = 0$.*

Proof. We have discussed before that subsets of $[1,n]$ correspond to binary sequences, and it is clear that gappy subsets correspond to gappy sequences, so we will work with binary sequences from now on. We will also assume that $k > 0$, leaving the trivial case $k = 0$ to the reader. Suppose that we have a gappy sequence $a \in G_{nk}$. The first one in a might appear in the very first position, so it need not be preceded by a zero. However, there are $k-1$ more ones, and by the gappy condition, each of them must have a zero immediately before it. The ones and these adjacent zeros take $2k-1$ slots altogether, so we must have $n \geq 2k-1$. This shows that $G_{nk} = \emptyset$ if $n < 2k-1$; we will assume that $n \geq 2k-1$ from now on. If we delete these zeros, we get a binary sequence of length $n-k+1$ containing k ones, or in other words, and element of $B_{n-k+1,k}$. On the other hand, if we are given an element of $B_{n-k+1,k}$, then we can get an element of G_{nk} by inserting zeros to the left of all the ones, except for the first one. Thus, we have a one-to-one correspondence between G_{nk} and $B_{n-k+1,k}$, showing that $|G_{nk}| = |B_{n-k+1,k}| = \binom{n-k+1}{k}$.

Here are some examples of elements of G_{73}, and the corresponding elements of B_{53}:

B_{53}	G_{73}
1 0 1 0 1	1 0 0 1 0 0 1
0 1 1 1 0	0 1 0 1 0 1 0
1 1 0 0 1	1 0 1 0 0 0 1

Video

□

Problem 1.24. *Suppose that a doctor's surgery has a single row of* 12 *chairs, and there are* 5 *patients waiting. In how many ways can they be seated such that no two are next to each other?*

Solution. The number is $|G_{12,5}| = \binom{12-5+1}{5} = \binom{8}{5} = 56$.

Problem 1.25. *In a draw for the National Lottery (as in Problem 1.16), what is the probability that there is an adjacent pair of numbers?*

Solution. We are selecting an element of $P_6[1, 59]$ at random, and we want to know the probability that it is not gappy. The total size of $P_6[1, 59]$ is $\binom{59}{6} \simeq 4.5 \times 10^7$. The number of gappy sets is

$$|G_6[1, 59]| = \binom{59 - 6 + 1}{6} = \binom{54}{6} = 28827165 \simeq 2.6 \times 10^7.$$

Thus, the number of non-gappy sets is $\binom{59}{6} - \binom{54}{6} \simeq 1.9 \times 10^7$, and the proportion of non-gappy sets is

$$\frac{\binom{59}{6} - \binom{54}{6}}{\binom{59}{6}} \simeq \frac{1.9 \times 10^7}{4.5 \times 10^7} \simeq 0.42.$$

Thus, approximately 42% of draws will have an adjacent pair of numbers.

2. Counting solutions

Proposition 2.1. *Consider an equation $x_1 + \cdots + x_k = n$, where $x_1, \ldots, x_k$ are required to be strictly positive integers. Then*

$$\textit{number of solutions} = \binom{n-1}{k-1} = \binom{\textit{right hand side} - 1}{\textit{number of variables} - 1}.$$

Proof. Consider a subset $A \subseteq [0, n-1]$ with $|A| = k-1$. We can list the elements as $0 < a_1 < a_2 < \cdots < a_{k-1} < n$ say. We then put

$$\begin{aligned} x_1 &= a_1 > 0 \\ x_2 &= a_2 - a_1 > 0 \\ x_3 &= a_3 - a_2 > 0 \\ \cdots &\quad \cdots \\ x_{k-1} &= a_{k-1} - a_{k-2} > 0 \\ x_k &= n - a_{k-1} > 0. \end{aligned}$$

If we add these equations together, then the a's will all cancel, and we get $x_1 + \cdots + x_k = n$, so we have a solution to the original equation. Conversely, if we have a solution $x_1 + \cdots + x_n = k$ (with $x_i > 0$) then we have a corresponding subset

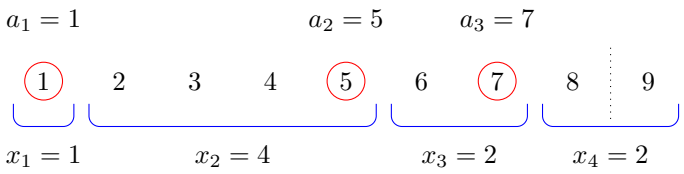
$$A = \{x_1, x_1 + x_2, x_1 + x_2 + x_3, \ldots, x_1 + \cdots + x_{k-1}\}$$

of size $k-1$ in $[1, n-1]$. This gives a one-to-one correspondence between the set of solutions and $P_{k-1}[1, n-1]$, so the number of solutions is $\binom{n-1}{k-1}$.

Video

□

Example 2.2. We can illustrate the above proof as follows. Take $k = 4$ and $n = 9$, so we are considering the equation $x_1 + x_2 + x_3 + x_4 = 9$. The proof gives a bijection between the solution set and the set $P_3[1, 8]$ of subsets of size 3 in $[1, 8]$. One such subset is $\{1, 5, 7\}$; it corresponds to the solution $1 + 4 + 2 + 2 = 9$, as shown below.



Example 2.3. Consider the equation $x_1 + x_2 + x_3 + x_4 = 6$ (with $x_i > 0$). The proposition tells us that the number of solutions is $\binom{5}{3} = 10$. They can be listed (in dictionary order) as follows.

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

Proposition 2.4. *Consider an equation $y_1 + \cdots + y_k = m$, where $y_1, \ldots, y_k$ are required to be nonnegative integers. Then*

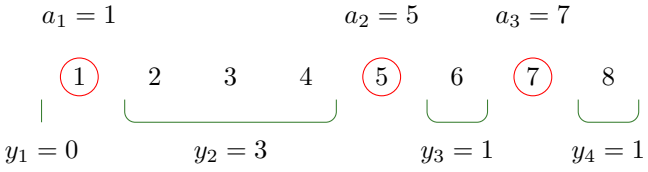
$$\textit{number of solutions} = \binom{m+k-1}{k-1} = \binom{\textit{right hand side} + \textit{number of variables} - 1}{\textit{number of variables} - 1}.$$

Proof. If we put $x_i = y_i + 1$, then the variables x_i are strictly positive integers, and must satisfy $x_1 + \cdots + x_k = m + k$. By Proposition 2.1, the number of solutions to this new equation is $\binom{m+k-1}{k-1}$, so this is also the number of solutions to the original equation. □

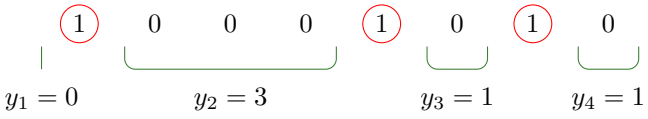
Example 2.5. The above argument shows that nonnegative solutions to $y_1 + \ldots + y_k = m$ biject with subsets $\{a_1, \ldots, a_k\} \subseteq [1, m + k - 1]$ of size k. Algebraically, the correspondence is

$$\begin{aligned} y_1 &= a_1 - 1 \geq 0 \\ y_2 &= a_2 - a_1 - 1 \geq 0 \\ y_3 &= a_3 - a_2 - 1 \geq 0 \\ \cdots &\quad \cdots \\ y_{k-1} &= a_{k-1} - a_{k-2} - 1 \geq 0 \\ y_k &= m + k - 1 - a_{k-1} \geq 0. \end{aligned}$$

For a pictorial example, consider the equation $y_1 + y_2 + y_3 + y_4 = 5$, so $k = 4$ and $m = 5$ and $m + k - 1 = 8$. The set $\{1, 5, 7\} \in P_3[1, 8]$ corresponds to the solution $0 + 3 + 1 + 1 = 5$, as illustrated below:



We can also draw this slightly differently, by writing the binary sequence 10001010 corresponding to the set $\{1, 5, 7\}$:



In this representation, the numbers y_i are just the lengths of the blocks of zeros between the ones (including the blocks at the left and right hand ends).

Remark 2.6. We can now give another approach to the problem of counting gappy sets. Suppose we want a gappy set $A = \{a_1, \dots, a_k\}$ of size k in $[1, n]$. Let x_0 be the size of the gap before a_1, and let x_k be the size of the gap after a_k. These are both allowed to be zero. However, the gap between a_i and a_{i+1} is required to have size at least one, so we can express it as $x_i + 1$, where $x_i \geq 0$. As A has size k in $[1, n]$, see that the total size of the gaps is $n - k$. This gives the equation

$$x_0 + (x_1 + 1) + \cdots + (x_{k-1} + 1) + x_k = n - k.$$

On the left hand side, we have $k - 1$ extra ones, so we can rearrange to get

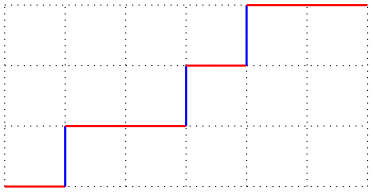
$$x_0 + \cdots + x_k = n - 2k + 1.$$

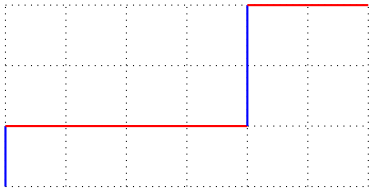
Here we have $k + 1$ variables and $n - 2k + 1$ on the right hand side, so the number of solutions is

$$\binom{\text{right hand side } + \text{ number of variables } - 1}{\text{number of variables } - 1} = \binom{(n - 2k + 1) + k}{k} = \binom{n - k + 1}{k}.$$

This agrees with the number of gappy sets, as we found in Proposition 1.23.

Problem 2.7. *Consider an $n \times m$ grid. Suppose that we want to go from the bottom left to the top right by taking a sequence of steps, each step going one space to the right or one space upwards. For example, two such routes across a 6×3 grid are shown below.*





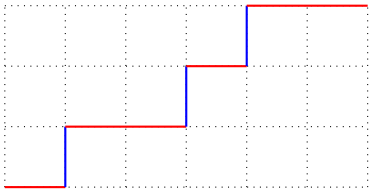
How many different routes are possible?

Solution. Interactive demo

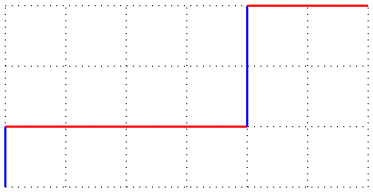
A route from the bottom left to the top right must consist of $n + m$ steps, of which n must be horizontal and m vertical. To choose such a path, we just need to choose which of the steps are horizontal. The number of ways of making that choice is $\binom{n+m}{n}$. Thus, the number of paths is $\binom{n+m}{n}$.

Remark 2.8. Interactive demo

We now have a new way to think about Proposition 2.4. Consider for example the equation $x_1 + x_2 + x_3 + x_4 = 6$. The proposition tells us that the number of solutions is $\binom{6+4-1}{4-1} = \binom{9}{3} = 84$. Given any solution, we can construct a grid path like this: we start at $(0, 0)$, then take x_1 horizontal steps, then a vertical step, then x_2 horizontal steps, then a vertical step, then x_3 horizontal steps, then a vertical step, then x_4 horizontal steps. Altogether this gives $x_1 + x_2 + x_3 + x_4 = 6$ horizontal steps and 3 vertical steps, so we have a grid path from $(0, 0)$ to $(6, 3)$. For example, the path on the left below has horizontal segments of length 1, 2, 1 and 2, so it corresponds to the equation $1 + 2 + 1 + 2 = 6$. The path on the right starts by going upwards, which we count as having an initial horizontal segment of length 0. We then have a horizontal segment of length 4, then a vertical segment of length 2. We count this as two vertical segments of length one, with a horizontal segment of length 0 in between. Finally, we have a horizontal segment of length 2. Thus, the corresponding solution is $0 + 4 + 0 + 2 = 6$.



$1 + 2 + 1 + 2 = 6$



$0 + 4 + 0 + 2 = 6$

With this construction, we get one grid path for every solution to the equation, and vice-versa. Thus, the number of solutions is the same as the number of grid paths. Any such path consists of 9 steps, of which 3 must be vertical. Thus, the number of grid paths is $\binom{9}{3} = 84$, and the number of solutions to our equation is also 84. All this can be generalised in a straightforward way: if we have a nonnegative solution to the equation $x_1 + \cdots + x_k = n$, then we can use it to make a grid path consisting of horizontal segments of lengths $x_1, \dots, x_k$, and a single vertical step between these, making $k-1$ vertical steps and $x_1 + \cdots + x_k = n$ horizontal steps altogether. To specify such a path, we take $n+k-1$ steps and choose which $k-1$ of them should be vertical. The number of solutions is the number of possible ways to make this choice, which is $\binom{n+k-1}{k-1}$. This agrees with Proposition 2.4.

Proposition 2.9. *For $0 < k \leq n$ we have*

$$\binom{n}{k} = \binom{k-1}{k-1} + \binom{k}{k-1} + \cdots + \binom{n-1}{k-1} = \sum_{m=k}^{n} \binom{m-1}{k-1}.$$

Bijective proof. Interactive demo

The left hand side is the number of subsets $A \subseteq [1, n]$ of size k. We will show that the right hand side can also be interpreted in the same way. To choose A, we can start by choosing the largest element of A, say m. Then we need to choose $k-1$ additional elements, which must all be less than m. This will only be possible if $m \geq k$, so we can assume that $k \leq m \leq n$. Once we have chosen m, the remaining $k-1$ elements must be taken from $[1, m-1]$, and there are $\binom{m-1}{k-1}$ ways to do this. The total number of possible choices is therefore $\sum_{m=k}^{n} \binom{m-1}{k-1}$, as required. □

Inductive proof. We will argue by induction on n. The base case is when $n = 1$. As $0 < k \leq n$, we must also have $k = 1$ in this case. The claim is then that $\binom{1}{1} = \sum_{m=1}^{1} \binom{0}{m-1} = \binom{0}{0}$, and this is true because $\binom{1}{1} = \binom{0}{0} = 1$.

Now suppose that $n > 1$ and $0 < k \leq n$. We can assume as an induction hypothesis that

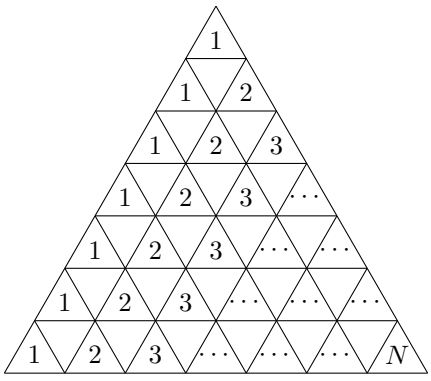
$$\binom{n-1}{k} = \binom{k-1}{k-1} + \cdots + \binom{n-2}{k-1} = \sum_{m=k}^{n-1} \binom{m-1}{k-1}.$$

Adding $\binom{n-1}{k-1}$ to both sides gives

$$\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{k-1}{k-1} + \cdots + \binom{n-2}{k-1} + \binom{n-1}{k-1} = \sum_{m=k}^{n} \binom{m-1}{k-1}.$$

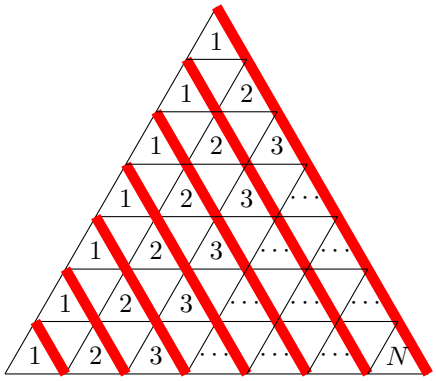
However, Proposition 1.19 tells us that the left hand side is the same as $\binom{n}{k}$, so we see that $\binom{n}{k} = \sum_{m=k}^{n} \binom{m-1}{k-1}$ as claimed. □

Proposition 2.10. *In a triangle as shown, the sum of all the entries is $\binom{N+2}{3}$.*



Proof. Interactive demo

Let T be the sum of all the numbers in the triangle. We can divide the triangle into stripes as follows:



The sum of the terms in the p'th stripe is $1+2+\cdots+p$, which is the same as $\binom{p+1}{2}$ by Proposition 1.18. Thus, the sum of all the numbers in the triangle is $T = \sum_{p=1}^{N} \binom{p+1}{2}$. On the other hand, we can take $n = N+2$ and $k = 3$ in Proposition 2.9 to get

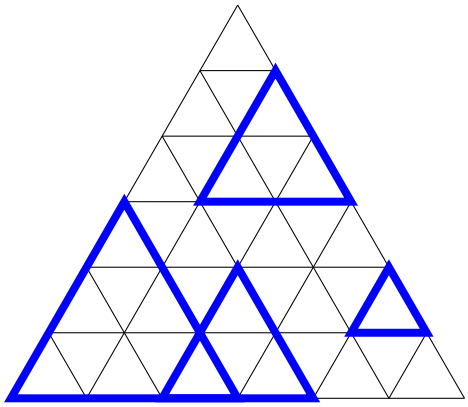
$$\binom{N+2}{3} = \sum_{q=3}^{N+2} \binom{q-1}{2}.$$

If we put $q = p+2$, this becomes

$$\binom{N+2}{3} = \sum_{p=1}^{N} \binom{p+1}{2} = T,$$

as claimed. □

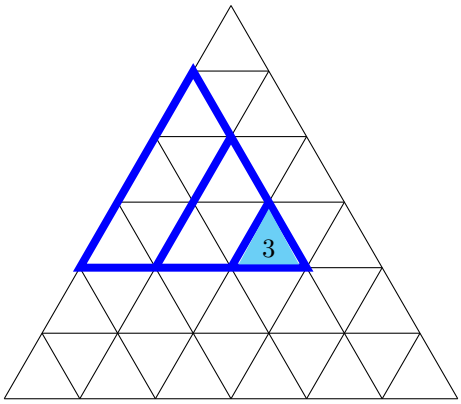
Problem 2.11. *Consider the same triangle again.*



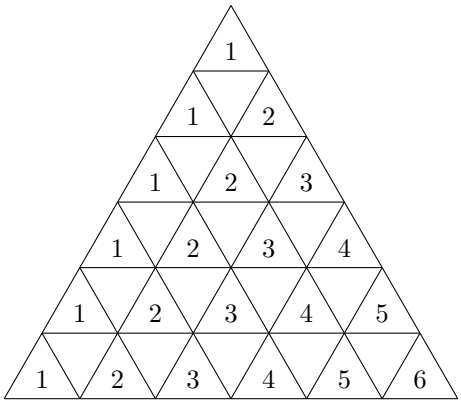
We have marked four different upward-pointing subtriangles. How many such subtriangles are there in total?

Solution. Interactive demo

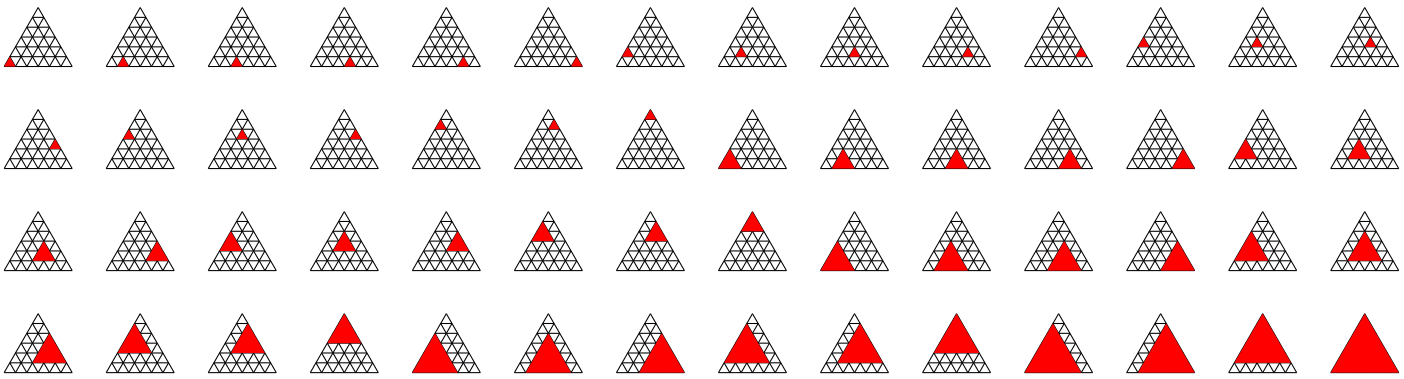
Consider the following picture:



The shaded triangle appears as the bottom right corner of three different subtriangles, one of size 1, one of size 2 and one of size 3, which are also shown in the picture. Because of this, we have marked the shaded triangle with a 3. In the same way, for each upward triangle T of size one, we can count all the subtriangles that have T as the bottom right corner, and mark T with that number. We get the following picture:



The total number of all subtriangles is the sum of the numbers in this picture. This is just the same as the sum considered in Proposition 2.10 (with $N = 6$), so the total number of subtriangles is $\binom{6+2}{3} = \binom{8}{3} = 56$.



More generally, if we start with a triangle of size N, the total number of upward-pointing subtriangles is $\binom{N+2}{3}$.

Definition 2.12. The *Fibonacci numbers* f_n are defined by $f_0 = 1$ and $f_1 = 1$ and $f_n = f_{n-2} + f_{n-1}$ for all $n \geq 0$. For example, we have

$$f_2 = f_0 + f_1 = 1 + 1 = 2 \qquad f_3 = f_1 + f_2 = 1 + 2 = 3$$
$$f_4 = f_2 + f_3 = 2 + 3 = 5 \qquad f_5 = f_3 + f_4 = 3 + 5 = 8$$
$$f_6 = f_4 + f_5 = 5 + 8 = 13 \qquad f_7 = f_5 + f_6 = 8 + 13 = 21.$$

Proposition 2.13. *For all $n \geq 0$ we have*

$$f_n = \binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots = \sum_{k \geq 0} \binom{n-k}{k}.$$

(Recall here that $\binom{m}{k}$ is defined to be zero if $k > m$, so the terms in the sum are eventually zero.)

Proof. Video

Put $g_n = \sum_{k \geq 0} \binom{n-k}{k}$, so the claim is that $g_n = f_n$. We will prove this by induction. For the first few cases, we have

$$g_0 = \binom{0}{0} + \binom{-1}{1} + \binom{-2}{2} + \cdots = 1 + 0 + 0 + \cdots = 1 = f_0$$
$$g_1 = \binom{1}{0} + \binom{0}{1} + \binom{-1}{2} + \cdots = 1 + 0 + 0 + \cdots = 1 = f_1$$
$$g_2 = \binom{2}{0} + \binom{1}{1} + \binom{0}{2} + \cdots = 1 + 1 + 0 + \cdots = 2 = f_2.$$

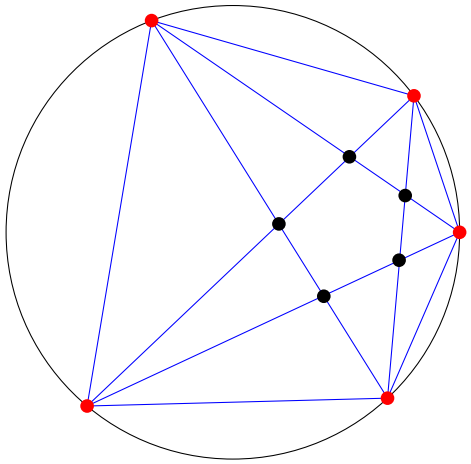
Now suppose that $n > 2$, and consider $g_n = \sum_{k \geq 0} \binom{n-k}{k}$. Proposition 1.19 tells us that $\binom{n-k}{k} = \binom{n-k-1}{k} + \binom{n-k-1}{k-1}$. This gives

$$g_n = \sum_{k \geq 0} \binom{n-k-1}{k} + \sum_{k \geq 0} \binom{n-k-1}{k-1}.$$

The first sum here directly matches the definition of g_{n-1}. In the second sum, we note that the term for $k = 0$ is $\binom{n-1}{-1} = 0$, so we can start from $k = 1$ instead of $k = 0$. We can then rewrite the sum in terms of the variable $j = k - 1$, so that $j \geq 0$ and $k = j + 1$ and $n - k - 1 = n - 2 - j$. The second sum then becomes $\sum_{j \geq 0} \binom{n-2-j}{j}$, which is g_{n-2}. We now see that $g_n = g_{n-1} + g_{n-2}$. We can assume as an inductive hypothesis that $g_{n-1} = f_{n-1}$ and $g_{n-2} = f_{n-2}$, so we have $g_n = f_{n-1} + f_{n-2}$. Using the definition of the Fibonacci numbers, this becomes $g_n = f_n$, as claimed. □

Proposition 2.14. *Suppose we have points $a_1, \ldots, a_n$ in anticlockwise order around the unit circle, and we draw a line l_{ij} from a_i to a_j for each $i \neq j$. Suppose that the points are in general position, so there is no point where more than two of the lines cross. Then the resulting diagram has $\binom{n}{2}$ lines, and $\binom{n}{4}$ interior crossing points, and $1 + \binom{n}{2} + \binom{n}{4}$ regions.*

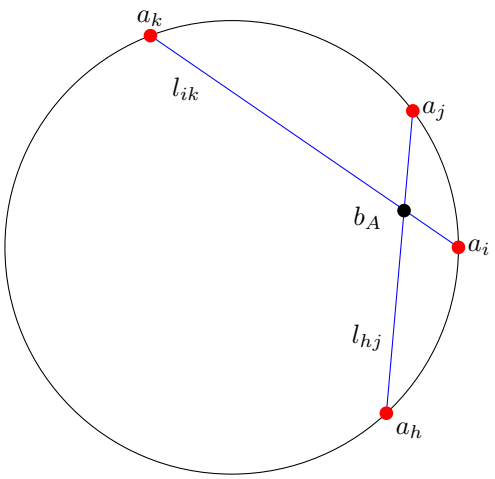
Example 2.15. The following picture shows the case shows the case $n = 5$. The number of lines is 10, which is $\binom{5}{2}$ as expected. The number of interior crossing points (marked in black) is 5, which is $\binom{5}{4}$ as expected. The lines divide the disk into 16 regions, and $16 = 1 + \binom{5}{2} + \binom{5}{4}$ as expected.



Proof of Proposition 2.14. Interactive demo

As $l_{ij} = l_{ji}$, we see that the number of lines is the number of possible subsets $\{i, j\} \subseteq \{1, \dots, n\}$ of size two, which is $\binom{n}{2}$.

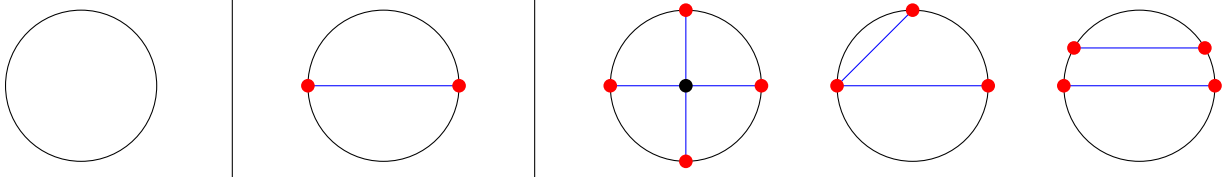
Now suppose we have a subset $A \subseteq \{1, \dots, n\}$ of size 4. We can list the elements in order as h, i, j, k with $h < i < j < k$. As we have numbered the points a_p in order around the circle, we find that the line l_{hj} meets the line l_{ik} at a single point b_A lying inside the circle:



This construction gives a bijection from the set of subsets of size 4 to the set of internal crossing points, so the number of such points is $\binom{n}{4}$ as claimed.

Now suppose we start with an empty disc, and add in the lines l_{ij} one by one. At each stage, we keep track of the number L of lines, the number C of internal crossings and the number R of regions. We also keep track of the number $E = R - L - C$. At the beginning (stage 0) there are no lines or crossings, and the disc is a single undivided region, so $L = C = 0$ and $R = 1$ and $E = 1 - 0 - 0 = 1$. More compactly, we can write $(L, C, R, E) = (0, 0, 1, 1)$. At stage 1, we add a single line, which splits the disk into two regions, but there are still no crossings. We therefore have $L = 1$ and $C = 0$ and $R = 2$ and $E = 2 - 1 - 0 = 1$, or in other words $(L, C, R, E) = (1, 0, 2, 1)$. At stage 2, we add a second line, and there are several different possibilities, depending on how

the two lines are placed relative to each other. However, we find that all three possibilities still have $E = 1$.



$(0, 0, 1, 1)$ Stage 0 | $(1, 0, 2, 1)$ Stage 1 | $(2, 1, 4, 1)$ $(2, 0, 3, 1)$ $(2, 0, 3, 1)$ Three possibilities for stage 2

In fact, we claim that E stays equal to 1 throughout the whole process. Indeed, suppose we add in a new line l_{ij} from a_i to a_j. This may create new crossing points. We list these in order (moving from a_i to a_j) as $x_1, \ldots, x_r$ say, and we also write $x_0 = a_i$ and $x_{r+1} = a_j$. This divides the new line into segments $s_i = [x_{i-1}, x_i]$, for $i = 1, \ldots, r + 1$. Each of these $r + 1$ segments cuts one of the old regions into two new regions, so the number R of regions increases by $r + 1$. At the same time, L increases by 1 and C increases by r, so the combination $E = R - L - C$ is unchanged. (It can also happen that there are no new crossing points; then everything works in essentially the same way, but with $r = 0$.) At stage 0 we have $E = 1$, so at the last stage we still have $E = 1$. However, at the last stage we have added in all the lines, so by our previous discussion we have $L = \binom{n}{2}$ and $C = \binom{n}{4}$. We now see that

$$1 = E = R - L - C = R - \binom{n}{2} - \binom{n}{4},$$

and we can rearrange this to get $R = 1 + \binom{n}{2} + \binom{n}{4}$ as claimed. □

3. Parity

Problem 3.1. *How many solutions are there for the equation $2x + 6y = 11$? What if we insist that x and y must be integers?*

Solution. If we work with real numbers then the equation is equivalent to $y = (11 - 2x)/6$. Thus, for every real number t we have a solution $(x, y) = (t, (11 - 2t)/6)$, showing that there are infinitely many solutions.

However, there are no solutions at all if we required that x and y are integers. Indeed, if x and y are integers then $2x + 6y$ is an even integer, and so cannot be equal to 11.

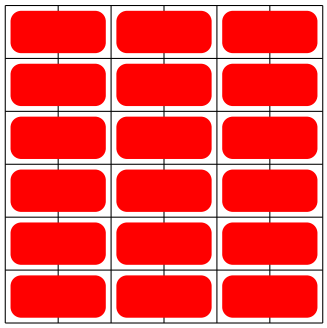
Problem 3.2. *Are there integer solutions for $12x + 18y = 250$?*

Solution. By comparison with the previous example, we might be tempted to say yes: both the left and right hand sides are even, so there is no problem with parity. However, just because there is no problem with parity, we cannot conclude that there are no other problems. In fact, the left hand side is always divisible by 3, but the right hand side is not, so in fact there are no integer solutions.

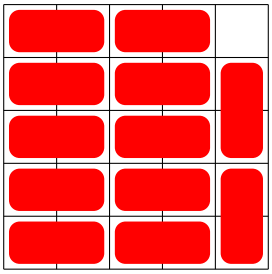
Problem 3.3. *Suppose we have an $n \times n$ chessboard. Can we cover it by non-overlapping dominos? (It matters whether n is even or odd). What if we remove two opposite corners from the board; can we cover the remainder of the board by non-overlapping dominos?*

Solution. Interactive demo

Suppose that n is even, say $n = 2m$. Then we can cover each row of the board with m non-overlapping horizontal dominos, and thus cover the whole board with $2m^2$ non-overlapping dominos. The case where $n = 6$ and $m = 3$ is like this:



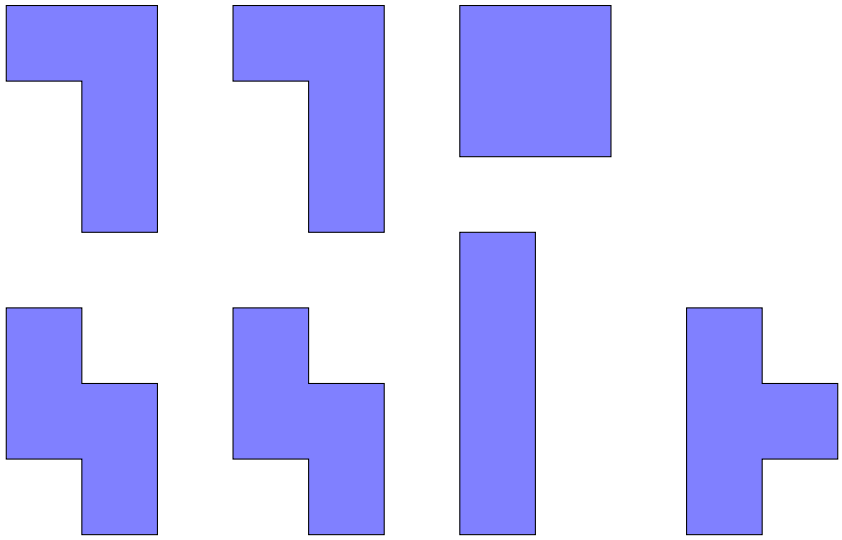
Now suppose that n is odd. We could try to cover the board like this:



This fails, because there is one extra square that we have not covered. However, this does not really prove anything. Our first unimaginative approach has failed, but perhaps that is just because we were not clever enough; perhaps there is a different approach of stupendous complexity and cunning that will cover the whole board? In fact that is not the case, but we need a proper proof to explain why. Fortunately, that is not very difficult. We are assuming that n is odd. There are n^2 squares in total, and this number is also odd. On the other hand, each domino covers two squares, so any set of m non-overlapping dominos covers $2m$ squares, and this number is even. Thus, it is impossible for a set of non-overlapping dominos to cover the whole board.

Now suppose we consider a nicked $n \times n$ board (where $n \geq 2$), with two opposite corners removed. The total number of squares is $n^2 - 2$. If n is odd then $n^2 - 2$ is again odd, so the nicked board still cannot be covered by disjoint dominos. Suppose instead that n is even, say $n = 2m$. Then the nicked board has $4m^2 - 2 = 2(2m^2 - 1)$ squares, and this number is even. We might be tempted to deduce that the board can be covered by disjoint dominos, but that would be too hasty. Suppose that we colour the squares in the usual chessboard pattern. The full $(2m) \times (2m)$ board will then have $2m^2$ white squares and $2m^2$ black squares. The two removed corners can be joined by a diagonal line, and all the squares on that line will have the same colour, so in particular, the two removed corners will have the same colour. Suppose for the sake of example that they are both white. This nicked board will then have $2m^2 - 2$ white squares and $2m^2$ black squares. However, every domino will cover one white square and one black square. Thus, any set of disjoint dominos will cover the same number of black squares as white squares. As the nicked board has more black squares than white squares, it cannot be covered by a disjoint set of dominos.

Problem 3.4. *Consider a puzzle consisting of the following tiles:*

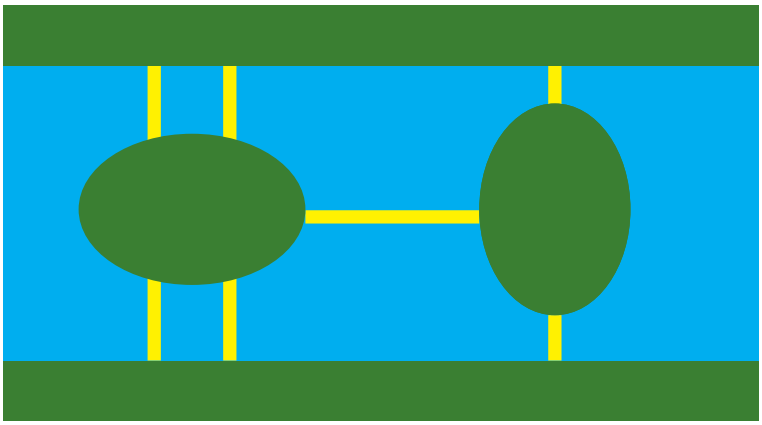


Note that these cover 28 *squares in total. Can they be arranged to cover a* 4×7 *square? (You are allowed to rotate the tiles or flip them over.)*

Solution. Interactive demo

We can imagine placing the tiles on a grid that is coloured like a chessboard. It is not hard to see that however you place the first six tiles, each of them will cover two white squares and two black squares. (One way to see this is to note that each of the first six tiles can be divided into two non-overlapping dominos.) That makes 12 white squares and 12 black squares altogether. However, the last T-shaped tile is different: however you place it, it will cover either three black squares and one white square, or three white squares and one black square. Thus, the full set of seven tiles will cover either 15 black and 13 white, or 13 black and 15 white. However, any 4×7 rectangle will cover 14 squares of each colour. This proves that the seven tiles cannot cover such a rectangle.

Problem 3.5. *When the mathematician Euler was living there, the city of Königsburg used to have a network of bridges like this:*



People were discussing the following problem: is it possible to do a tour of the city which crosses each bridge precisely once? Ideally the tour should start and end in the same place, but we do not insist on that.

Solution. Interactive demo

Euler used a parity argument to show that no such tour is possible. Indeed, there are four sections of land in the city: the north bank (N), the south bank (S), the western island (W) and the eastern island (E). We will call these *nodes*. Nodes N, S and E are each connected to three bridges, and W is connected to five

bridges: in each case, the number is odd. Now suppose we have a tour of the city that crosses each bridge precisely once. Let P be a node that is neither the beginning nor the end of the tour. Every time we reach P on one bridge, we must leave on a different bridge. Thus, if we visit P a total of n times, then we will have crossed $2n$ of the bridges that touch P. However, P has an odd number of bridges, so we cannot have crossed all of them, which is a contradiction.

Remark 3.6. Video

The above solution can be generalised as follows. Suppose we have a network of nodes $A_1, \dots, A_n$, with some bridges between them. Let d_i be the number of bridges with one end at A_i. Suppose that there is a tour which starts at A_p and ends at A_q. If $i \neq p, q$ then the same logic as before shows that d_i must be even. However, we expect d_p to be odd: we cross one bridge when we leave A_p as the first step of the tour, then any further visits to A_p will involve crossing one bridge to arrive and another bridge to leave, giving an odd number of bridges altogether. Similarly, we expect d_q to be odd: throughout most of the tour we use bridges in pairs, then we have one additional bridge for the last step of the tour when we arrive at A_q and do not leave again. However: there is one exception to this picture: we could have a circular tour, which starts and ends at the same node A_p. Then we have a bridge for the first step, a bridge for the last step, and pairs of bridges in between, giving an even number altogether. In summary:

(a) If there is a circular tour crossing every bridge precisely once, then the numbers d_i are all even with no exceptions.
(b) If there is a non-circular tour crossing every bridge precisely once, then the numbers d_i are all even with precisely two exceptions.

Thus, if three of the d_i's are odd then there cannot be a tour of the required type.

4. The pigeonhole principle

Remark 4.1. Let A be a finite set, with $|A| = n$ say. Suppose we have a sequence $a_1, \dots, a_m$ of elements of A, with $m > n$. As A only has n elements, it is clearly impossible for all the elements a_i to be different. Thus, we can find $i < j$ with $a_i = a_j$.

Proposition 4.2 (Pigeonhole principle)**.** *Let A be a finite set, with $|A| = m$ say. Suppose we have a list of subsets $P_1, \dots, P_n$ (called* pigeonholes*) such that every element of A lies in one of these subsets. Suppose that $m > n$ (so the number of elements is greater than the number of pigeonholes). Then there exists i such that $|P_i| > 1$.*

Proof. List the elements of A as $a_1, \dots, a_m$. Each element a_i lies in some pigeonhole, so we can choose $k_i \in [1, n]$ such that $a_i \in P_{k_i}$. We now have a sequence $k_1, \dots, k_m$ in $[1, n]$. As $m > n$, these numbers cannot all be different, so we can choose $i < j$ such that $k_i = k_j = p$ say. This means that $a_i \in P_p$ and $a_j \in P_p$, so $|P_p| > 1$. □

Problem 4.3. *Show that there are two people in the world with the same number of hairs.*

Solution. Interactive demo

Let N be the number of people in the world, so $N \simeq 7.5 \times 10^9$. Let M be the maximum number of hairs that anyone has. A typical person has about 10^5 hairs, so it would be safe to assume that $M < 10^6$, and certainly M is very much smaller than N. Let h_i be the number of hairs on the i'th person, so $h_i \in [0, M]$. The sequence $h_1, \dots, h_N$ is much longer than the size of the set $[0, M]$, so there must exist $i < j$ such that $h_i = h_j$ as claimed.

As another way to say essentially the same thing, we can divide all people into pigeonholes $H_0, \dots, H_M$, where H_r is the set of people with r hairs. The number of pigeonholes is smaller than the number of people, so there must be some pigeonhole that contains two people. If H_r contains two people, then those two people have the same number of hairs.

Proposition 4.4. *Suppose that people $p_1, \dots, p_n$ have a meeting, and some of them shake each others hand. (No one shakes their own hand, and no pair of people shake hands more than once.) Then there are two people who shake the same number of hands.*

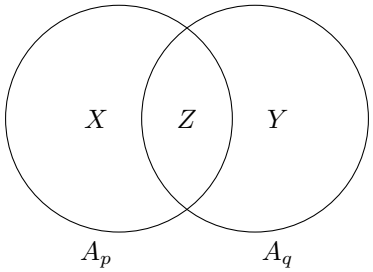
Proof. Interactive demo

Let d_i be the number of hands shaken by person p_i. There are $n-1$ people whose hands p_i could shake, so $0 \leq p_i \leq n-1$. In other words, if we put $N = \{0, 1, \dots, n-1\}$, then $p_i \in N$. Note also that because we start at zero, we have $|N| = n$. Now suppose, for a contradiction, that the numbers d_i are all different. We now have n different numbers $d_1, \dots, d_n$ all living in the set N of size n, so the numbers d_i must fill the whole set N. In particular, we must have $n-1 = d_a$ for some a, and $0 = d_b$ for some b. Because $d_a = n-1$ we see that person p_a shakes everyone else's hand. Because $d_b = 0$, we see that p_b shakes no-one else's hand. This is a contradiction, because p_a shakes p_b's hand. Thus, the numbers d_i cannot all be different, after all. Thus, we have $d_i = d_j$ for some $i \neq j$, so p_i and p_j shake the same number of hands. □

Problem 4.5. *Let U be a set of 8 integers, with all the members of U between 1 and 32 (inclusive). Show that there are two distinct disjoint subsets of U with the same sum.*

Solution. Interactive demo

As $|U| = 8$, the number of subsets of U is $2^8 = 256$. Let these subsets be $A_1, A_2, \dots, A_{256}$, and let s_k be the sum of the elements of A_k. Note that s_k is the sum of 8 distinct terms, each of which is between 1 and 32. The largest possible sum of this form is $32+31+30+29+28+27+26+25 = 228$. (We can just do this addition by hand, or use the arithmetic progression rule: we have 8 equally spaced terms between 25 and 32, so the average is $\frac{25+32}{2}$ and the total is $8 \times \frac{25+32}{2} = 228$.) We thus have 256 numbers $s_1, \dots, s_{256}$ lying in the set $S = \{0, \dots, 228\}$. As $|S| = 229 < 256$, there is not enough space for the numbers s_k to all be different. We can therefore find indices $p \neq q$ such that $s_p = s_q$. In other words, the sets A_p and A_q are different, but they have the same sum. However, we have not yet solved the problem, because we were supposed to find *disjoint* subsets. However, this is easily fixed. We put $X = A_p \setminus A_q$ and $Y = A_q \setminus A_p$ and $Z = A_p \cap A_q$, so we have a Venn diagram as follows:



It is then clear that X and Y are distinct, disjoint subsets with the same sum. (In a bit more detail, we can let x be the sum of all elements of X, and similarly for y and z. We then have $s_p = x + z$ and $s_q = y + z$. We chose p and q such that $s_p = s_q$, and it follows that $x = y$ as required.)

Remark 4.6. The above discussion is carefully phrased to work around the following finicky point: it is possible for two sets X and Y to be disjoint but not distinct, if they both happen to be empty. However, in the proof we have $A_p \neq A_q$ so at least one of X and Y must be nonempty, and they have the same sum, so both must be nonempty.

Problem 4.7. *Consider a sequence $x_1, \dots, x_n$ of integers. Show that there is a consecutive subsequence $x_p, x_{p+1}, \dots, x_q$ (for some $p \leq q$) such that $x_p + \dots + x_q$ is divisible by n.*

Proof. Interactive demo

For $0 \leq i \leq n$ we put

$$y_i = \left(\sum_{j=1}^{i} x_i \right) \pmod{n} \in \{0, 1, \dots, n-1\},$$

so $y_0 = 0$ and $y_1 = x_1 \pmod{n}$ and $y_2 = (x_1 + x_2) \pmod{n}$ and so on. This gives $n+1$ numbers $y_0, \dots, y_n$, all lying in the set $\{0, \dots, n-1\}$, which has size n. This means that the numbers y_i cannot all be different,

so we can find indices $m < q$ such that $y_m = y_q$. Now note that

$$\begin{aligned} y_m &= x_1 + \cdots + x_m \pmod n \\ y_q &= x_1 + \cdots + x_m + x_{m+1} + \cdots + x_q \pmod n \\ y_q - y_m &= x_{m+1} + \cdots + x_q \pmod n. \end{aligned}$$

We chose m and q so that $y_m = y_q$, and it follows that $x_{m+1} + \cdots + x_q = 0 \pmod n$. In other words, if we take $p = m + 1 \leq q$, we find that $x_p + \cdots + x_q$ is divisible by n. □

Problem 4.8. *Suppose that I deposit money in my piggybank every day for* 30 *days. On* 15 *of the days I deposit* £1*, and on each of the other* 15 *days I deposit* £2*. Given an integer k with $1 \leq k \leq 15$, show that there is a sequence of consecutive days during which I deposit precisely k pounds.*

Remark 4.9. For small values of k, we can give a direct argument.

- For $k = 1$: There are 15 days on which I deposit £1, and any one of those days counts as a sequence of length one during which I deposit precisely £1.
- For $k = 2$: There are 15 days on which I deposit £2, and any one of those days counts as a sequence of length one during which I deposit precisely £2.
- For $k = 3$: At some point I must swap over between depositing £1 and £2. Thus, for some i, I deposit £1 on day i, and £2 on day $i + 1$, or *vice versa*. Either way, $\{i, i + 1\}$ is a sequence of consecutive days over which I deposit precisely £3.
- For $k = 4$: As there are 15 days on which I deposit £2, I can certainly find a day i when I deposit £2 that is not the beginning or end of the month, so $1 < i < 30$. If I deposit £2 on day $i - 1$, then $\{i - 1, i\}$ is a sequence of consecutive days over which I deposit precisely £4. If I deposit £2 on day $i + 1$, then $\{i, i + 1\}$ is a sequence as required. If neither of these possibilities occur, then I must have deposited £1 on days $i - 1$ and $i + 1$, so $\{i - 1, i, i + 1\}$ is a sequence as required.

This is clearly becoming increasingly unwieldy, so it is better to take a less direct approach.

Solution. Interactive demo

Let x_i be the amount deposited on day i, so $x_i \in \{1, 2\}$ for $i = 1, \ldots, 30$. Let y_i be the total deposited up to and including day i, so $y_i = x_1 + \cdots + x_i$. We also put $y_0 = 0$. Note that y_{30} is the total amount deposited over the whole month, which is $15 \times 1 + 15 \times 2 = 45$. Because some money is deposited every day, we have

$$0 = y_0 < y_1 < y_2 < \cdots < y_{30} = 45.$$

In particular, the numbers y_i are all different, so the set $Y = \{y_0, \ldots, y_{30}\}$ has size 31. Now suppose we are given k with $1 \leq k \leq 15$. We put $Z = \{y_0 + k, \ldots, y_{30} + k\}$, and note that this also has size 31. Because $y_i \leq 45$ and $k \leq 14$ we have $y_i + k \leq 60$. We now see that Y and Z are both subsets of the set $N = \{0, 1, \ldots, 60\}$, which has size 61. If Y and Z were disjoint, we would have $|Y \cup Z| = |Y| + |Z| = 62$, which is impossible, as $Y \cup Z$ is a subset of N. It follows that Y and Z are not disjoint, so we can choose $m \in Y \cap Z$. As $m \in Y$, we have $m = y_q$ for some q. As $m \in Z$, we have $m = y_p + k$ for some p, so $y_q = y_p + k$. This means that we must have $p < q$, and

$$x_{p+1} + \cdots + x_q = (x_1 + \cdots + x_q) - (x_1 + \cdots + x_p) = y_q - y_p = k.$$

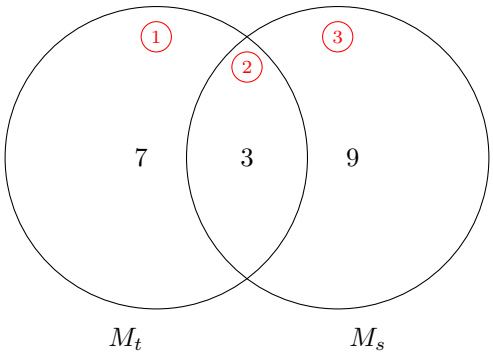
Thus, over the sequence of days $\{p + 1, \cdots, q\}$, I deposit precisely k pounds.

5. The inclusion-exclusion principle

Problem 5.1. *Consider a sports club in which people can play tennis or squash. Some members play both sports, some play only one, and some just drink at the bar. Suppose that* 10 *members play tennis (and maybe squash as well),* 12 *members play squash (and maybe tennis as well), and* 3 *members play both sports. How many members play at least one sport?*

Solution. Interactive demo

Let M_t be the set of members who play tennis, and let M_s be the set of members who play squash. We are given that $|M_t| = 10$ and $|M_s| = 12$ and $|M_s \cap M_t| = 3$, and we need to find $|M_s \cup M_t|$. For this, we need to fill in the numbers in the following Venn diagram:



Region 2 is $M_t \cap M_s$, which has 3 elements, so we write 3 there. Regions 1 and 2 together make up M_t, which has 10 elements, so we need $10 - 3 = 7$ elements in region 1 to make the total correct. Regions 2 and 3 together make up M_s, which has 12 elements, so we need $12 - 3 = 9$ elements in region 3 to make the total correct. Now $M_t \cup M_s$ consists of regions 1, 2 and 3, so it has $7 + 3 + 9 = 19$ elements in total. In other words, there are 19 members who play at least one sport.

Another way to describe the solution is as follows: we could just take $|M_t| + |M_s|$, but that would count the people who play both sports twice, once as members of M_t, and once more as members of M_s. To compensate for this, we need to subtract the number of people who play both sports, which is $|M_t \cap M_s|$. This gives

$$|M_t \cup M_s| = |M_t| + |M_s| - |M_t \cap M_s| = 10 + 12 - 3 = 19$$

as before.

Problem 5.2. *Now suppose instead that we have a club that offers tennis (t), squash (s) and badminton (b). We are given the following data:*

10 members play t	*5 members play t and s*
15 members play s	*4 members play t and b*
12 members play b	*3 members play s and b*
2 members play t, s and b	*There are 40 members altogether.*

How many members play no sport at all?

Solution. Interactive demo

This time we will give a more algebraic explanation. We write M for the set of all members, and $m = |M|$. We write M_t for the set of members who play t, and $m_t = |M_t|$ for the number of players who play t, and similarly for M_s, M_{tb} and so on. The initial data is then as follows:

$$\begin{aligned} m_t &= 10 & m_{ts} &= 5 \\ m_s &= 15 & m_{tb} &= 4 \\ m_b &= 12 & m_{sb} &= 3 \\ m_{tsb} &= 2 & m &= 40. \end{aligned}$$

Now let M_t^* be the set of members who play t and nothing else, and let M_{ts}^* be the set of members who play t and s and nothing else, and so on. To complete the pattern, we put $M_{tsb}^* = M_{tsb}$, and we write M^* for the set of members who play no sport at all, so our problem is to find $|M^*|$. The people who play t can be divided into four groups:

- Those that play t and nothing else
- Those that play t and s but not b
- Those that play t and b but not s
- Those that play t and s and b.

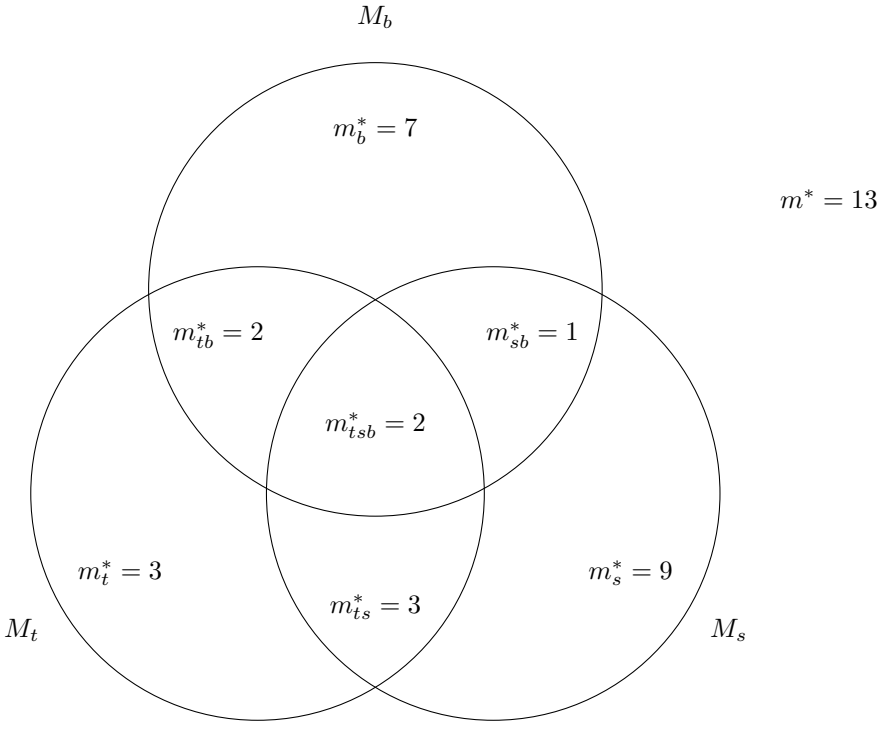
From this we get $M_s = M_s^* \cup M_{ts}^* \cup M_{tb}^* \cup M_{tsb}^*$, and these sets do not overlap, so we get $m_s = m_s^* + m_{ts}^* + m_{tb}^* + m_{tsb}^*$. By a similar analysis, we get the following equations:

$$\begin{aligned}
m_{tsb}^* &= m_{tsb} = 2 \\
m_{ts}^* + m_{tsb}^* &= m_{ts} = 5 \\
m_{tb}^* + m_{tsb}^* &= m_{tb} = 4 \\
m_{sb}^* + m_{tsb}^* &= m_{sb} = 3 \\
m_t^* + m_{st}^* + m_{tb}^* + m_{tsb}^* &= m_t = 10 \\
m_s^* + m_{st}^* + m_{sb}^* + m_{tsb}^* &= m_s = 15 \\
m_b^* + m_{tb}^* + m_{sb}^* + m_{tsb}^* &= m_b = 12 \\
m^* + m_t^* + m_s^* + m_b^* + m_{ts}^* + m_{tb}^* + m_{sb}^* + m_{tsb}^* &= m = 40.
\end{aligned}$$

These equations are easily solved to give

$$\begin{aligned}
m_{tsb}^* &= m_{tsb} = 2 \\
m_{ts}^* &= m_{ts} - m_{tsb} = 5 - 2 = 3 \\
m_{tb}^* &= m_{tb} - m_{tsb} = 4 - 2 = 2 \\
m_{sb}^* &= m_{sb} - m_{tsb} = 3 - 2 = 1 \\
m_t^* &= m_t - m_{ts} - m_{tb} + m_{tsb} = 10 - 5 - 4 + 2 = 3 \\
m_s^* &= m_s - m_{ts} - m_{sb} + m_{tsb} = 15 - 5 - 3 + 2 = 9 \\
m_b^* &= m_b - m_{tb} - m_{sb} + m_{tsb} = 12 - 4 - 3 + 2 = 7 \\
m^* &= m - m_t - m_s - m_b + m_{ts} + m_{tb} + m_{sb} - m_{tsb} = 40 - 10 - 15 - 12 + 5 + 4 + 3 - 2 = 13.
\end{aligned}$$

In particular, we have $m^* = 13$, so there are 13 members who play no sport; this answers the original question.



We now want to discuss the *Inclusion-Exclusion Principle* (IEP), which generalises the last two problems. Suppose we have a finite set B, together with a family of subsets $B_a \subseteq B$ for each a in some set A of labels.

(For example, in Problem 5.2 we have a set $B = M$, together with subsets M_s, M_t and M_b, indexed by the set $A = \{s, t, b\}$ of available sports.) We put

$$B' = \{b \in B \mid b \text{ lies in at least one of the sets } B_a\}$$
$$B^* = B \setminus B' = \{b \in B \mid b \text{ lies in none of the sets } B_a\}.$$

In the common case where $A = \{1, \dots, n\}$, this can be written as

$$B' = B_1 \cup B_2 \cup \cdots \cup B_n$$
$$B^* = B \setminus (B_1 \cup B_2 \cup \cdots \cup B_n).$$

The IEP tells us about $|B'|$ and $|B^*|$. To formulate it, we use the following notation. Given a subset $I \subseteq A$, we put $B_I = \bigcap_{i \in I} B_i$. This means that

$$B_{\{i\}} = B_i$$
$$B_{\{i,j\}} = B_i \cap B_j$$
$$B_{\{i,j,k\}} = B_i \cap B_j \cap B_k$$

and so on. For the case $I = \emptyset$, we interpret this as $B_\emptyset = B$. We will often abbreviate the notation, by writing B_{ijk} for $B_{\{i,j,k\}}$ and so on.

Theorem 5.3. *For B and B_i as above, we have*

$$|B^*| = \sum_{I \subseteq A} (-1)^{|I|} |B_I|$$
$$|B'| = \sum_{I \neq \emptyset} (-1)^{|I|+1} |B_I|.$$

In the case where $A = \{1, \dots, n\}$, this can be written as

$$|B^*| = |B| - |B_1| - \cdots - |B_n| + |B_{12}| + \cdots + |B_{n-1,n}| - \cdots \pm |B_{12\cdots n}|$$
$$|B'| = |B_1| + \cdots + |B_n| - |B_{12}| - \cdots - |B_{n-1,n}| + \cdots \mp |B_{12\cdots n}|$$

There is a single video covering the statement and proof of the IEP, together with two lemmas required for the proof:

Video

.

The equation for $|B^*|$ is called the *negative form* of the IEP, and the equation for $|B'|$ is called the *positive form*. Because $B^* = B \setminus B'$ we have $|B^*| = |B| - |B'|$, which makes it easy to see that the two forms are equivalent. For $n = 2$ and $n = 3$ the equations are as follows:

$$|B_1 \cup B_2| = |B_1| + |B_2| - |B_{12}|$$
$$|B \setminus (B_1 \cup B_2)| = |B| - |B_1| - |B_2| + |B_{12}|$$
$$|B_1 \cup B_2 \cup B_3| = |B_1| + |B_2| + |B_3| - |B_{12}| - |B_{13}| - |B_{23}| + |B_{123}|$$
$$|B \setminus (B_1 \cup B_2)| = |B| - |B_1| - |B_2| - |B_3| + |B_{12}| + |B_{13}| + |B_{23}| - |B_{123}|.$$

Problem 5.1 is just an example of the positive IEP with $n = 2$. Problem 5.2 is an example of the negative IEP with $n = 3$.

We will prove the IEP after some preliminary discussion.

Lemma 5.4. *Let I be a finite set, and consider the sum $s = \sum_{J \subseteq I} (-1)^{|J|}$. Then $s = 1$ if I is empty, and $s = 0$ if I is not empty.*

Proof. If I is empty, then the only term in the sum is for $J = \emptyset$, and that term is $(-1)^0 = 1$, so $s = 1$. Suppose instead that $I \neq \emptyset$, and put $n = |I| > 0$. Then there are $\binom{n}{k}$ possible choices of J with $|J| = k$, and this gives $s = \sum_k \binom{n}{k} (-1)^k$. This is just the binomial expansion of $(1-1)^n = 0^n = 0$, so $s = 0$. Alternatively, we can choose an element $a \in I$, and put $I' = I \setminus \{a\}$. For every $J' \subseteq I'$ we have a term $(-1)^{|J'|}$ in s for

$J = J'$, and another term $(-1)^{1+|J'|}$ for $J = J' \cup \{a\}$, and these terms cancel out. All the terms cancel in pairs in this way, so we are left with $s = 0$. □

Definition 5.5. In the context of the IEP, for an element $b \in B$, we put

$$A\langle b\rangle = \{a \in A \mid b \in B_a\} \subseteq A.$$

For example, consider a member x of the club in Problem 5.2. Then $A\langle x\rangle$ is just the set of sports that x plays. For example, if x plays tennis and badminton but not squash, then $A\langle x\rangle = \{t, b\}$.

- Consider the set $I = \{t, s, b\}$ and the corresponding set $M_I = M_{tsb} = M_t \cap M_s \cap M_b$. Member x does not lie in this set M_I, because x does not play s. Here $I \not\subseteq A\langle x\rangle = \{t, b\}$.
- Consider instead the set $I = \{t, s\}$ and the corresponding set $M_I = M_{ts} = M_t \cap M_s$. Member x does not lie in this set M_I, because x does not play s. Here $I \not\subseteq A\langle x\rangle = \{t, b\}$.
- Now consider the set $I = \{t, b\}$ and the corresponding set $M_I = M_{tb} = M_t \cap M_b$. Member x does lie in this set M_I, because x does not play both t and b. Here $I \subseteq A\langle x\rangle = \{t, b\}$.
- Similarly, if $I = \{b\}$ then x does lie in the set $M_I = M_b$, and again $I \subseteq A\langle x\rangle$.
- For an arbitrary subset $I \subseteq A = \{t, s, b\}$ we find that $x \in M_I$ iff x plays all the sports in I iff $I \subseteq K\langle x\rangle$.

We record the obvious generalisation as a lemma:

Lemma 5.6. *Suppose we have a family of subsets $(B_a)_{a\in A}$ as before, and an element $b \in B$, and a subset $I \subseteq A$. Then $b \in B_I$ iff $I \subseteq A\langle b\rangle$.*

Proof. By definition $B_I = \bigcap_{i\in I} B_i$, so $b \in B_I$ iff $b \in B_i$ for all $i \in I$. However, we have $b \in B_i$ iff $i \in A\langle b\rangle$, by the definition of $A\langle b\rangle$. Thus, we can say that $b \in B_I$ iff for all $i \in I$, we have $i \in A\langle b\rangle$. This is clearly equivalent to the condition $I \subseteq A\langle b\rangle$. □

Lemma 5.7. *$A\langle b\rangle$ is empty iff $b \in B^*$.*

Proof. We have $i \in A\langle b\rangle$ iff $b \in B_i$. Thus $A\langle b\rangle$ is empty iff the condition $b \in B_i$ is false for all i, which means that b lies in none of the sets B_i, which means that $b \in B^*$. □

Proof of Theorem 5.3. Put $u = \sum_{I\subseteq A}(-1)^{|I|}|B_I|$. We need to prove that this is the same as $|B^*|$. We have $|B_I| = \sum_{b\in B_I} 1$, so we can rewrite the definition of u as

$$u = \sum_{I\subseteq A}\sum_{b\in B_I}(-1)^{|I|}.$$

Lemma 5.6 tells us that $b \in B_I$ iff $I \subseteq A\langle b\rangle$, so we can regroup this sum as

$$u = \sum_{b\in B}\sum_{I\subseteq A\langle b\rangle}(-1)^{|I|}.$$

Now Lemma 5.4 tells us that $\sum_{I\subseteq A\langle b\rangle}(-1)^{|I|}$ is zero if $A\langle b\rangle \neq \emptyset$, but is 1 if $A\langle b\rangle = \emptyset$. Using Lemma 5.7, we therefore see that $\sum_{I\subseteq A\langle b\rangle}(-1)^{|I|}$ is zero if $b \notin B^*$, but is 1 if $b \in B^*$. We now have

$$u = \sum_{b\in B^*} 1 = |B^*|,$$

as required. This proves the negative form of the IEP. For the positive form, we note that

$$|B'| = |B| - |B^*| = |B| - \sum_{I\subseteq A}(-1)^{|I|}|B_I|.$$

The term for $I = \emptyset$ in the sum cancels out the extra term of $|B|$ outside the sum. We can also bring the minus sign inside the sum to get

$$|B'| = \sum_{I\neq\emptyset}(-1)^{|I|+1}|B_I|.$$

□

Definition 5.8. Let S_n be the set of all permutations of the set $N = \{1, \dots, n\}$, so $|S_n| = n!$. A *derangement* of $\{1, \dots, n\}$ is a permutation $\sigma \in S_n$ with the property that for all i we have $\sigma(i) \neq i$. We write D_n for the set of derangements, so $D_n \subseteq S_n$. We also write $p_n = |D_n|/|S_n|$ (which is the probability that a randomly chosen permutation is a derangement).

Example 5.9. This picture lists all 24 possible permutations of the set $N = \{1, 2, 3, 4\}$. For example, the top right box contains 1432, which refers to the permutation sending 1, 2 3 and 4 to 1, 4, 3 and 2 respectively. (In disjoint cycle notation, this would be (2 4).) The numbers 1 and 3 are sent to themselves, so they are underlined. As some numbers are sent to themselves, this is not a derangement. However, in the bottom right box we have the permutation 4321. This does not send anything to itself, so no numbers are underlined, and we have a derangement. All the derangements are circled; there are 9 of them. Thus, the fraction of derangements is $p_4 = 9/24 = 3/8 = 0.375$.

1234	1243	1324	1342	1423	1432
2134	2143	2314	2341	2413	2431
3124	3142	3214	3241	3412	3421
4123	4132	4213	4231	4312	4321

Example 5.10. Suppose that n people arrive at a party, each wearing a hat. At the end of the party, no one can remember which hat they brought, so they pick one up at random. This means that guest i picks up the hat belonging to guest $\sigma(i)$, for some randomly chosen permutation σ. This permutation is a derangement iff no one gets the right hat. Thus, the probability that no one gets the right hat is p_n. How does this change as n increases? For any given guest, there are more hats to choose from, so the probability of getting the right hat goes down. On the other hand, as there are more guests, there are more chances for at least one guest to get the right hat. It is not obvious how these competing effects balance out, but the answer is given by our next result.

Proposition 5.11. *$p_n = \sum_{k=0}^{n} (-1)^k/k!$, and this converges to $e^{-1} \simeq 0.368$ as $n \to \infty$.*

Proof. Interactive demo Interactive demo

Put

$$N = \{1, \dots, n\}$$
$$P = S_n = \{ \text{ all permutations of } N\}$$
$$P_i = \{\sigma \in P \mid \sigma(i) = i\} = \{ \text{ permutations that fix } i\}.$$

Note that a permutation is a derangement iff it lies in none of the sets P_i, so $D_n = P^*$ in the usual notation of the IEP. We therefore have

$$|D_n| = |P^*| = \sum_{I \subseteq N} (-1)^{|I|} |P_I|$$
$$p_n = n!^{-1} |D_n| = \sum_{I \subseteq N} (-1)^{|I|} |P_I|/n!.$$

Here P_I is the set of permutations σ that fix all the elements of I, but are free to permute the remaining elements of $N \setminus I$ in any way. If $|I| = k$ we have $|N \setminus I| = n - k$ so there are $(n-k)!$ possible permutations of $N \setminus I$. This means that $|P_I| = (n-k)!$. On the other hand, there are $\binom{n}{k}$ possible choices of I with $|I| = k$. Putting this together, we get

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

However, we also have

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

so our previous expression simplifies to $p_n = \sum_{k=0}^{n}(-1)^k/k!$ as claimed. As n tends to infinity, this converges to $\sum_{k=0}^{\infty}(-1)^k/k!$, which is e^{-1} by the standard Taylor series for e^x. □

Problem 5.12. *Of the numbers $0, 1, \ldots, 41$, how many are coprime with 42?*

Solution. Interactive demo

Put $D = \{0, \ldots, 41\}$, and let U be the subset of numbers that are coprime with 42. We need to find $|U|$. Put $P = \{2, 3, 7\}$, which is the set of primes that divide 42. For any $p \in P$, put

$$D_p = \{i \in D \mid i \text{ is divisible by } p\}.$$

In the standard notation for the IEP, we have $U = D^*$, and so $|U| = \sum_{I\subseteq P}(-1)^{|I|}|D_I|$. We therefore need to understand $|D_I|$. Let q_I be the product of the primes in I (to be interpreted as $q_I = 1$ in the case $I = \emptyset$). We note that q_I divides 42, and D_I is the set of multiples of q_I in D, so $D_I = \{k\,q_I \mid 0 \le k < 42/q_I\}$ and $|D_I| = 42/q_I$. We now have

$$|U| = 42\sum_{I\subseteq P}\frac{(-1)^I}{q_I} = 42\left(1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{7} + \frac{1}{2\times 3} + \frac{1}{2\times 7} + \frac{1}{3\times 7} - \frac{1}{2\times 3\times 7}\right).$$

It is not hard to see that this factors as

$$|U| = 42\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{7}\right) = 12.$$

Alternatively, we could say that the proportion of coprime numbers is $|U|/|D| = (1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{7}) = 2/7$.

The following more general statement can be proved in the same way:

Proposition 5.13. *Consider an integer $m > 1$, and let P be the set of primes that divide m. Put $D = \{0, 1, \ldots, m-1\}$, and let x be the proportion of numbers in D that are coprime with m. Then*

$$x = \prod_{p\in P}(1 - p^{-1}). \quad \square$$

6. Matching problems

Video

A large part of this course will be about matching problems. A typical example is as follows. We have a set J of jobs, and a set P of people. Each person is qualified to do some subset of the jobs. Ideally, we would like to give every person a job that they are qualified to do, in such a way that every person has exactly one job, and every job is filled. Of course this is only possible if the number of jobs is the same as the number of people. If there are more people than jobs, we can still hope to fill every job, leaving some people unemployed. This still might not be possible, if we have a large number of difficult jobs, and not many highly skilled people. If a perfect matching is not possible, we might try to find partial matchings that fill as many jobs as possible. We might also ask how many different partial matchings are possible.

As well as the basic problem mentioned above, there are a number of possible variants. We could have jobs that need a team of people rather than just a single person. We could have a subset of enthusiastic people who really want a job, and we could try to organise the matching so that all of them are employed. We could try to account for the fact that some people are more qualified than others, rather than just distinguishing qualified people from unqualified people.

We can also apply the same mathematical ideas in different contexts. Instead of allocating people to jobs, we could allocate A-level students to university places, or junior doctors to hospitals, or processes to processors in a multiprocessor computer. We could also match romantic partners to each other; some of the earliest mathematical literature on matching was written in terms of this problem. Moreover, there are many purely abstract mathematical applications of the same theory.

Definition 6.1. A *matching problem* consists of finite sets A and B, together with a subset $E \subseteq A \times B$. A *row* in $A \times B$ is a set of the form $\{a\} \times B$, and a *column* is a set of the form $A \times \{b\}$. A *partial matching* for E is a subset $M \subseteq E$ that contains at most one element in each row, and at most one element in each column. We say that M is *row-full* if it meets every row, and *column-full* if it meets every column.

Example 6.2. A could be a set of people, B could be a set of jobs, and E could be the set of pairs (a, b) such that person a is qualified for job b. Each person a gives a row $\{a\} \times B$, and the intersection $(\{a\} \times B) \cap E$ tells us the set of jobs that person a is qualified to do. Each job b gives a column $A \times \{b\}$, and the intersection $(A \times \{b\}) \cap E$ tells us the set of people that are qualified to do job b. Given a partial matching $M \subseteq E$, we can allocate job b to person a for every pair $(a, b) \in M$. Part of the definition of a partial matching says that $M \subseteq E$; this ensures that we only allocate people to jobs that they are qualified to do. Another part of the definition says that every row contains at most one element of M; this ensures that no one has more than one job. The last part of the definition says that every column contains at most one element of M; this ensures that we do not give the same job to more than one person. The matching M is row-full iff everyone gets a job, and it is column-full iff every job is filled.

Example 6.3. A and B could be two disjoint sets of people, so that everyone in A wants to marry someone from B and *vice-versa*. Then E could be the set of pairs (a, b) such that a and b would be content to marry each other. A partial matching then gives a set of disjoint compatible couples. The matching is row-full if everyone in A has a partner, and column-full if everyone in B has a partner.

Example 6.4. Consider again a sports club as in Problem 5.2. Suppose that the committee is supposed to have a Tennis Officer, a Squash Officer and a Badminton Officer. These are required to be three different people, who must be players of the relevant sport. To choose these officers, we can consider an appropriate matching problem. We take P to be the set of members and S to be the set of sports. We then put

$$E = \{(p, x) \in P \times S \mid \text{ person } p \text{ plays sport } x \ \}.$$

Suppose we have a partial matching $M \subseteq E$ that is column-full. Then M must have the form $\{(p, t), (q, s), (r, b)\}$ where p is a tennis player, q is a squash player and r is a badminton player, and p, q and r are all different. Thus, we could make p the Tennis Officer, q the Squash Officer and r the Badminton Officer.

Remark 6.5. We can reformulate the description of a matching problem as follows.

- For each $a \in A$ we put $R_a = \{b \in B \mid (a, b) \in E\}$, and call this the *$a$'th row set* for E.
- For each $b \in B$ we put $C_b = \{a \in A \mid (a, b) \in E\}$, and call this the *$b$'th column set* for E.

We note that

$$E = \{(a, b) \mid b \in R_a\} = \{(a, b) \mid a \in C_b\}$$
$$R_a = \{b \mid (a, b) \in E\} = \{b \mid a \in C_b\}$$
$$C_b = \{a \mid (a, b) \in E\} = \{a \mid b \in R_a\}.$$

Thus, if we know the row sets we can determine the column sets and the set E. Similarly, if we know the column sets then we can determine the row sets and the set E. In the job allocation context of Example 6.2, we just have

$$R_p = \{ \text{ jobs that person } p \text{ is qualified to do } \} \subseteq J$$
$$C_j = \{ \text{ people who are qualified to do job } j \ \} \subseteq P.$$

Example 6.6. Interactive demo

We can give a more specific example of a job allocation problem as follows. We have a set P of people called Ann, Bob, Cath, Dave and Ella, abbreviated a,b,c,d,e. We have a set J of jobs: librarian, musician, nurse, optician, pilot, abbreviated l,m,n,o,p. A typical element of $P \times J$ is the pair (Bob, nurse) $= (b, n)$; we will usually just write this as bn for brevity. We now need to specify the set

$$E = \{(x, y) \mid \text{ person } x \text{ is qualified for job } y \ \}.$$

We take

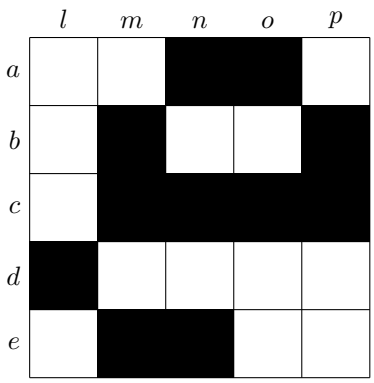
$$E = \{al, am, ap, bl, bn, bo, cl, dm, dn, do, dp, ea, eo, ep\}.$$

For example, the pair ap is an element of E, indicating that Ann is qualified to be a pilot. The pair cm is not an element of E, indicating that Cath is not qualified to be a musician.

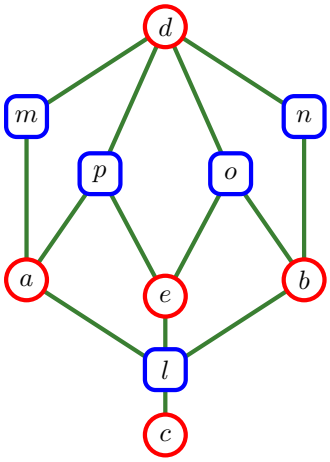
It is generally more convenient to indicate this information graphically, rather than listing the elements of E explicitly. This can be done in several different ways. In the left hand picture below, the positions in E are marked by white squares, but the positions not in E are marked by black squares. For example, the top left position is al, which is an element of E, indicating that Ann is qualified to be a librarian, so the to left square is white. The middle square is cn, which is not in E, indicating that Cath is not qualified to be a nurse, so the middle square is black. We call this picture the *chessboard diagram* for the matching problem.

The middle picture is essentially the same as the left hand one, except that we have 1's (instead of white squares) for the positions in E, and 0's (instead of black squares) for the positions that are not in E. We call this the *incidence matrix* for the matching problem.

In the right hand picture, we have a red circle for each person, a blue square for each job, and a green line for each element of E. For example, there is a green line between a and m, reflecting the fact that $am \in E$, or that Ann is qualified to be a musician. There is no green line between b and p, reflecting the fact that $bp \notin E$, or that Bob is not qualified to be a pilot. We call this picture the *incidence graph*.



	l	m	n	o	p
a	1	1	0	0	1
b	1	0	1	1	0
c	1	0	0	0	0
d	0	1	1	1	1
e	1	0	0	1	1



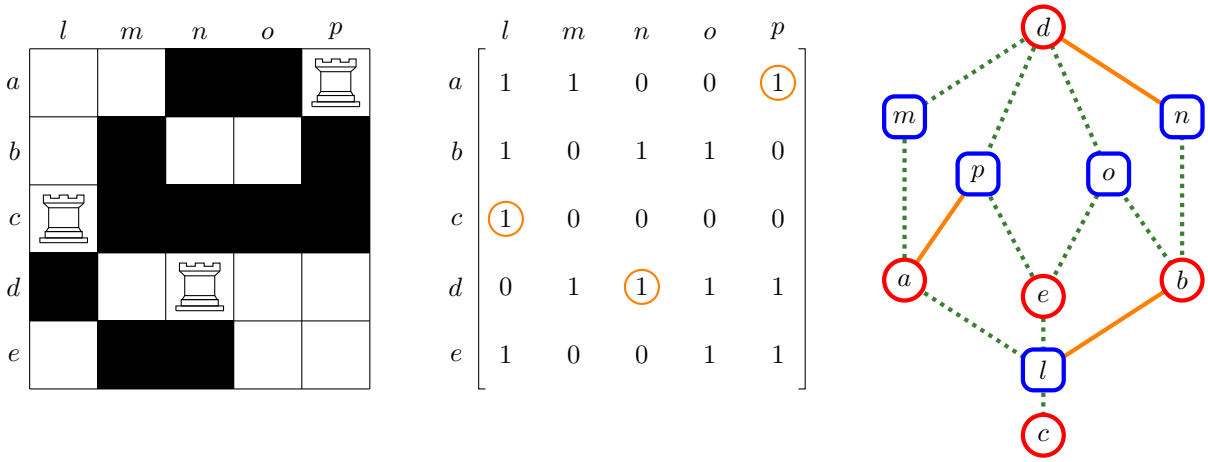
As yet another way to represent the same information, we can list the row sets and the column sets:

$$R_a = \{l, m, p\} \qquad C_l = \{a, b, c, e\}$$
$$R_b = \{l, n, o\} \qquad C_m = \{a, d\}$$
$$R_c = \{l\} \qquad C_n = \{b, d\}$$
$$R_d = \{m, n, o, p\} \qquad C_o = \{b, d, e\}$$
$$R_e = \{l, o, p\} \qquad C_p = \{a, d, e\}.$$

The fact that $R_b = \{l, n, o\}$ means that Bob is qualified to be a librarian, nurse or optician, but not to do any other job. We can read this off by looking for white squares in the second row of the chessboard diagram, or by looking for 1's in the second row of the incidence matrix, or by looking for nodes in the incidence graph that are connected by an edge to b. The fact that $C_p = \{a, d, e\}$ means that Ann, Dave and Ella (and nobody else) are qualified to be pilots. We can read this off by looking for white squares in the last column of the chessboard diagram, or by looking for 1's in the last column of the incidence matrix, or by looking for nodes in the incidence graph that are connected by an edge to p.

Example 6.7. We next discuss graphical representation of partial matchings. Recall that a partial matching is a subset $M \subseteq E$ that has at most one element in any row, and at most one element in any column. The condition $M \subseteq E$ means that M corresponds to a subset of the white squares in the chessboard diagram. We can mark this subset by placing a rook on each of the relevant squares. Recall the usual rules of chess: two rooks can attack each other horizontally if they lie in the same row, and they can attack each other vertically if they lie in the same column. Thus, the row and column conditions for a partial matching just say that none of the rooks can attack each other.

For instance, we can consider again the matching problem from Example 6.6. Take $M = \{ap, cl, dn\}$. This can be shown graphically as follows:



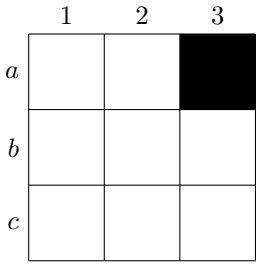
In the chessboard diagram, we have placed rooks in positions ap, cl and dn. The rooks are all in white squares and cannot attack each other, so this is a valid partial matching. In the incidence matrix, we have circled the entries in positions ap, cl and dn. All the circled entries are 1's, and no row contains more than one circle, and no column contains more than one circle. This is another way to express the fact that we have a valid partial matching. In the incidence graph, we have coloured the edges ap, cl and dn orange, and left them solid, while making the other edges dotted. The fact that the rooks cannot attack each other is reflected by the fact that the solid edges are disjoint.

7. Rook polynomials

In this section we will consider some matching problems, which will be represented by their chessboard diagrams. We will try to count the number of partial matchings of various sizes. We will represent partial matchings by placements of non-challenging rooks, as in Example 6.7.

Video

Problem 7.1. *Consider the matching problem E with the following chessboard diagram:*



(so $E = \{a1,\ a2,\ b1,\ b2,\ b3,\ c1,\ c2,\ c3\} \subset \{a, b, c\} \times \{1, 2, 3\}$).

(a) *How many ways are there of placing one rook?*
(b) *How many ways are there of placing two non-challenging rooks?*
(c) *How many ways are there of placing three non-challenging rooks?*

Solution. Interactive demo

(a) There are 8 white squares, and thus 8 ways of placing one rook. We can list them as follows:

$$a1,\ a2,\ b1,\ b2,\ b3,\ c1,\ c2,\ c3.$$

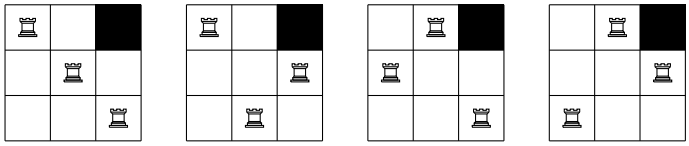
(b) By inspection, there are 14 ways of placing two non-challenging rooks. They can be listed as follows:

$$a1b2,\ a1b3,\ a1c2,\ a1c3,\ a2b1,\ a2b3,\ a2c1,\ a2c3,\ b1c2,\ b1c3,\ b2c1,\ b2c3,\ b3c1,\ b3c2.$$

(c) If we place three non-challenging rooks, then there must be one in each row. There are two choices for where to place the rook in row a. Then there are again two choices for where to place the rook in row b, because it is not allowed to go directly underneath the rook in row a. We have now blocked out two of the squares in row c, leaving only one possible choice for the third rook. This gives $2 \times 2 \times 1 = 4$ ways of placing three non-challenging rooks. We can list them as follows:

$$a1b2c3,\ a1b3c2,\ a2b1c3,\ a2b3c1.$$

Alternatively, we can display the chessboard diagrams:



$a1b2c3$ $a1b3c2$ $a2b1c3$ $a2b3c1$

(d) It is clearly impossible to place more than three non-challenging rooks. By convention, we also say that there is one way of placing no rooks.

Definition 7.2. Let $B \subseteq \{1, \dots, n\}^2$ be a matching problem (typically represented by an $n \times n$ chessboard diagram, with the elements of E coloured white, and the elements not in E coloured black). We write $C_k(B)$ for the set of partial matchings $M \subseteq E$ with $|M| = k$. (Such a partial matching corresponds to a placement of k non-challenging rooks on B.) We write $c_k(B) = |C_k(B)|$, so $c_k(B)$ is the number of partial matchings of size k, or the number of ways of placing k non-challenging rooks on B. We note that when $k > n$ we have $C_k(B) = \emptyset$ and $c_k(B) = 0$, and we put

$$r_B(t) = \sum_{k=0}^{n} c_k(B) t^k.$$

We call this the *rook polynomial* for B.

Remark 7.3. There is only one way of placing no rooks, so $c_0(B) = 1$. The number of ways of placing one rook is just the number of white squares in B, which will write as $|B|$, so $c_1(B) = |B|$. This means that

$$r_B(x) = 1 + |B|x + \cdots .$$

The higher coefficients are harder to calculate. In Problem 7.1, we showed that $c_2(B) = 14$ and $c_3(B) = 4$ and $c_k(B) = 0$ for $k > 3$, so

$$r_B(x) = 1 + 8x + 14x^2 + 4x^3.$$

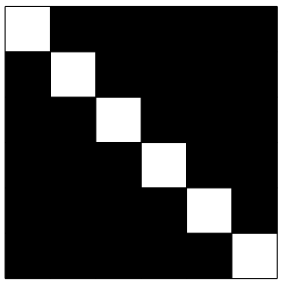
Remark 7.4. It is clear that rotating or reflecting a board does not affect the set of non-challenging rook placements, and so does not affect the rook polynomial.

Example 7.5. Let L_n be the full $1 \times n$ board, with all squares white:

$L_6 =$ 

It is clearly impossible to place more than one rook on L_n without the rooks challenging each other, so we just have $r_{L_n}(x) = 1 + nx$.

Example 7.6. Put $D_n = \{(1,1), (2,2), \dots, (n,n)\}$, so the corresponding chessboard diagram has white squares on the diagonal and black squares everywhere else.

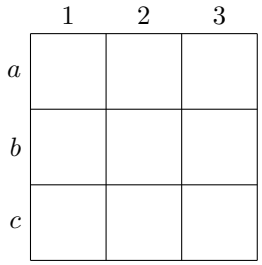
$D_6 =$ 

It is clearly impossible for two rooks placed anywhere on D_n to challenge each other. Thus, $C_k(D_n)$ just consists of all subsets of size k in D_n, which means that $c_k(D_n) = \binom{n}{k}$. This gives

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

Example 7.7. Interactive demo

Consider a full 3×3 board B, with all squares white:



(a) There are 9 squares, and thus 9 ways of placing one rook, so $c_1(B) = 9$. We can list them as follows:

$$C_1(B) = \{a1,\ a2,\ a3,\ b1,\ b2,\ b3,\ c1,\ c2,\ c3\}.$$

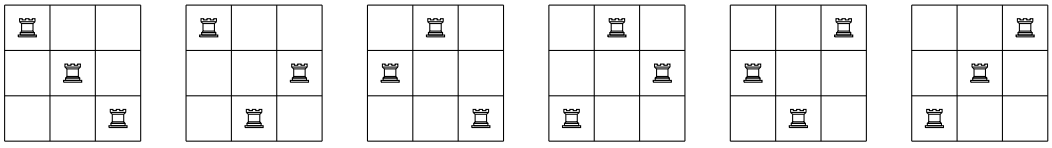
(b) If we place two non-challenging rooks, then they must be in different rows, leaving the third row empty. There are three ways to choose the empty row. Then there are three ways place a rook in the upper non-empty row. This blocks off one of the spaces in the lower non-empty row, leaving only two choices for where to place the second rook. Altogether, this gives $3 \times 3 \times 2 = 18$ ways of placing two non-challenging rooks, so $c_2(B) = 18$. We can list them as follows:

$$C_2(B) = \{a1b2,\ a1b3,\ a1c2,\ a1c3,\ a2b1,\ a2b3,\ a2c1,\ a2c3,\ a3b1,$$
$$a3b2,\ a3c1,\ a3c2,\ b1c2,\ b1c3,\ b2c1,\ b2c3,\ b3c1,\ b3c2\ \}.$$

(c) If we place three non-challenging rooks, then there must be one in each row. There are three choices for where to place the rook in row a. Then there are only two choices for where to place the rook in row b, because it is not allowed to go directly underneath the rook in row a. We have now blocked out two of the squares in row c, leaving only one possible choice for the third rook. This gives $3 \times 2 \times 1 = 6$ ways of placing three non-challenging rooks, so $c_3(B) = 6$. We can list them as follows:

$$C_3(B) = \{a1b2c3,\ a1b3c2,\ a2b1c3,\ a2b3c1,\ a3b1c2,\ a3b2c1\}.$$

Alternatively, we can display the chessboard diagrams:

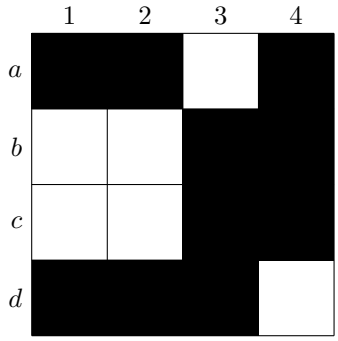


We now see that

$$r_B(x) = 1 + 9x + 18x^2 + 6x^3.$$

Example 7.8. Interactive demo

Now consider the following board B:



The possible rook placements are as follows:

- $C_0(B) = \{\emptyset\}$ so $c_0(B) = 1$.
- $C_1(B) = \{a3,\ b1,\ b2,\ c1,\ c2,\ d4\}$ so $c_1(B) = 6$.
- $C_2(B) = \{a3b1,\ a3b2,\ a3c1,\ a3c2,\ a3d4,\ b1c2,\ b1d4,\ b2c1,\ b2d4,\ c1d4,\ c2d4\}$ so $c_2(B) = 11$.
- $C_3(B) = \{a3b1c2,\ a3b1d4,\ a3b2c1,\ a3b2d4,\ a3c1d4,\ a3c2d4,\ b1c2d4,\ b2c1d4\}$ so $c_3(B) = 8$.
- $C_4(B) = \{a3b1c2d4,\ a3b2c1d4\}$ so $c_2(B) = 2$.

(Later, we will discuss a systematic method to construct these lists.) From this we see that the rook polynomial is

$$r_B(x) = 1 + 6x + 11x^2 + 8x^3 + 2x^4.$$

We next generalise Example 7.7.

Definition 7.9. We write F_{mn} for the full $m \times n$ board (with all squares white). Thus, as a set, we just have

$$F_{mn} = \{1, \dots, m\} \times \{1, \dots, n\} = \{(i, j) \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}.$$

We also write $F_n = F_{nn}$, so F_n is the full $n \times n$ square board.

Proposition 7.10. *Consider $C_n(F_n)$, which is the set of ways of placing n non-challenging rooks on the full $n \times n$ board. Then this set can be identified with the set of permutations of $\{1, \dots, n\}$, so $c_n(F_n) = n!$.*

Proof. Interactive demo

Any element of $C_n(B)$ is a placement of n non-challenging rooks. As the rooks do not challenge each other, we have at most one rook per row. As we have n rooks and only n rows, there must be a rook in every row. Let $\sigma(i)$ be the horizontal position of the rook in row i, so σ is a function from the set $N = \{1, \dots, n\}$ to itself. Similarly, each column must contain precisely one rook. Let $\tau(j)$ be the vertical position of the rook in column j. It is then easy to see that the functions σ and τ are inverse to each other, so σ is a bijection, or in other words a permutation. Conversely, if we are given a permutation σ then we can place a rook in position $(i, \sigma(i))$ for $i = 1, \dots, n$, and this gives us n non-challenging rooks. □

Proposition 7.11. *Consider the full $m \times n$ board, denoted by F_{mn}. Then for $0 \leq k \leq \min(m, n)$ there are precisely $\binom{m}{k}\binom{n}{k}k!$ ways of placing k non-challenging rooks on F_{mn}, so $c_k(F_{mn}) = \binom{m}{k}\binom{n}{k}k!$. Thus, the rook polynomial is*

$$r_{F_{mn}}(x) = \sum_{k=0}^{\min(m,n)} \binom{m}{k}\binom{n}{k}k!x^k.$$

In particular, for a square board of size $n \times n$ we have

$$r_{F_n}(x) = \sum_{k=0}^{n} \binom{n}{k}^2 k!x^k.$$

Example 7.12. For the case $m = n = 3$, the claim is that

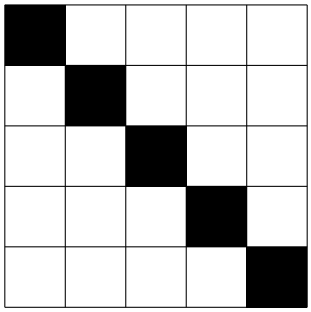
$$\begin{aligned} r_B(x) &= \binom{3}{0}^2 0! + \binom{3}{1}^2 1!x + \binom{3}{2}^2 2!x^2 + \binom{3}{3} 3!x^3 \\ &= 1 + 9x + 18x^2 + 6x^3, \end{aligned}$$

which agrees with Example 7.7.

Proof. Interactive demo

If we have k non-challenging rooks, then they must lie in k different rows. There are $\binom{m}{k}$ ways of choosing the rows in which the rooks will appear. Similarly, the rooks must lie in k different columns. There are $\binom{n}{k}$ ways of choosing the columns in which the rooks will appear. Now suppose we have chosen the rows and columns. If we ignore all the other rows and columns, we are just left with a $k \times k$ board, on which we need to place k non-challenging rooks. There are $k!$ ways to do this, by Proposition 7.10. Thus, we have $\binom{m}{k}\binom{n}{k}k!$ possibilities altogether. □

Problem 7.13. *Let B be an $n \times n$ board in which the main diagonal squares are black and all other squares are white. The case $n = 5$ is shown below.*



How many ways are there of placing n non-challenging rooks on this board?

Solution. Interactive demo

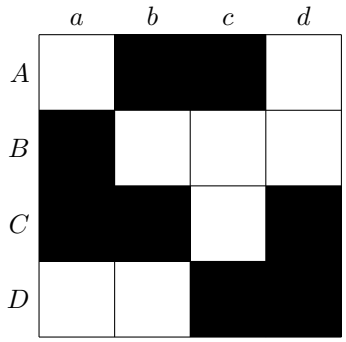
This is a problem that we have already solved, but slightly disguised. For any permutation σ of the set $N = \{1, \ldots, n\}$ we have a rook placement with rooks in positions $(i, \sigma(i))$. However, we want to ensure that no rooks are on the black squares, which are in positions (i, i). Thus, we need to have $\sigma(i) \neq i$ for all i. This is precisely the condition for σ to be a derangement, as introduced in Definition 5.8. Thus, the number of ways of placing n non-challenging rooks is the same as the number of derangements, which is $n! \sum_{k=0}^{n} (-1)^k/k! \simeq n!e^{-1}$ by Proposition 5.11.

Problem 7.14. *Suppose we have people A, B, C and D, and jobs a, b, c and d, with qualifications as follows:*

- *a can be done by A or D;*
- *b can be done by B or D;*
- *c can be done by B or C;*
- *d can be done by A or B.*

Is it possible to solve the job allocation problem, and if so, in how many ways?

Solution. The chessboard diagram for this matching problem is as follows:



The job allocation problem is equivalent to the problem of placing four non-challenging rooks on this board. This means that we must have a rook in each row. The rook in row C must be at Cc, and this means that the rook in row B cannot be at Bc, so it must be at Bb or Bd.

- Suppose we place a rook at Aa. This blocks Da, so the rook in row D must be at Db. Now Bb and Bc are blocked by Db and Cc, so the rook in row B must be at Bd. This gives a full matching $AaBdCcDb$.
- Suppose instead that we place a rook at Ad. This blocks Bd, and Bc is also blocked by Cc, so the rook in row B must be at Bb. Now Db is blocked by Bb, so the rook in row D must be at Da. This gives a full matching $AdBbCcDa$.

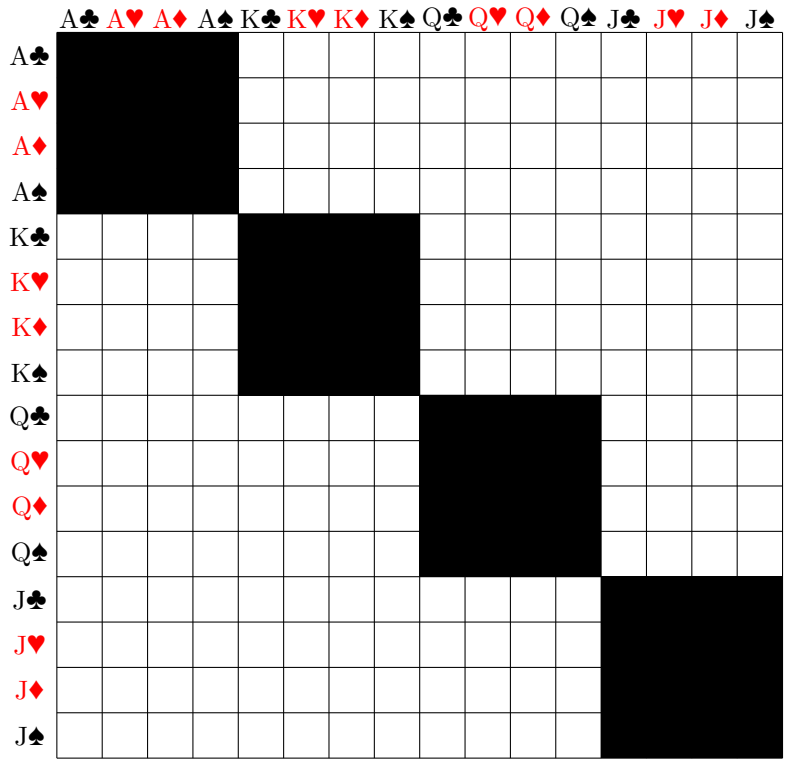
From this we see that there are precisely two full matchings. Equivalently, there are precisely two ways to allocate the jobs subject to the usual rules: each person should have precisely one job, and each job should be done by precisely one person.

We will next give two examples where we can show that an interesting problem is equivalent to a rook placement problem. However, we will not yet solve the resulting rook placement problems. Instead, we will first develop some general techniques, which will make the task easier.

Problem 7.15. *One version of the game of Snap works as follows. There are two players, each of whom has a full pack of 52 cards. At each turn, both players take a card from the top of their pack. If the two cards have the same value (for example, they are both kings or both sevens) then that counts as a snap, and the game ends. Of course, it could happen that the game continues for 52 turns and the players play all of their cards but still no snap has happened. What is the probability of this?*

Solution. Interactive demo

For simplicity, we will first consider the case where the decks just have the aces, kings, queens and jacks, so there are only 16 cards altogether. We will number the cards as $1, \ldots, 16$, with cards $1-4$ being the aces, cards $5-8$ being the kings, $9-12$ being the queens and $13-16$ the jacks. We will analyse the problem using the following board:



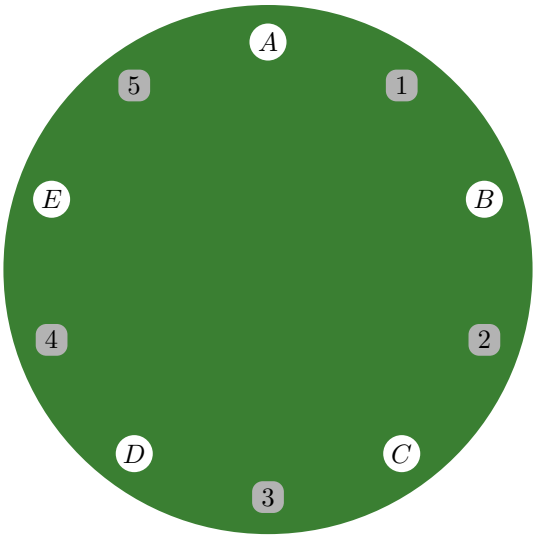
Squares are coloured black if the corresponding row and column markers have the same value; they are coloured white if the row and column markers have different values. For example, the (K♣,K♥) square is black (because the two markers are both kings) but the (Q♥,J♠) square is white.

We will assume for simplicity that the first player's pack is unshuffled, so cards 1 to 16 appear in order, but that the second player's pack is shuffled randomly. (We leave it as an exercise to understand why this simplifying assumption does not really change anything.) Let $\sigma(i)$ be the card in position i in the second player's pack, so σ is a random permutation, and cards i and $\sigma(i)$ get played at the same time. This permutation can be represented in the usual way by placing 16 rooks on the board, with the i'th rook at position $(i, \sigma(i))$. Suppose, for example, that $\sigma(6) = 3$, so cards 6 and 3 get played together. Card 6 is K♥, and card 3 is A♦. These have different values, so this is not a snap. This corresponds to the fact that the square (6, 3) is white. Suppose, for another example, that $\sigma(9) = 11$, so cards 9 and 11 are played together. Card 9 is Q♣ and card 11 is Q♦, so this is a snap. This corresponds to the fact that the square (9,11) is black. Thus, our game has a snap if there is a rook on a black square, but there is no snap if all the rooks are on white squares. Thus, our main task is to calculate the number of ways of placing 16 non-challenging rooks using the white squares only. We will return to this calculation when we have developed a bit more theory.

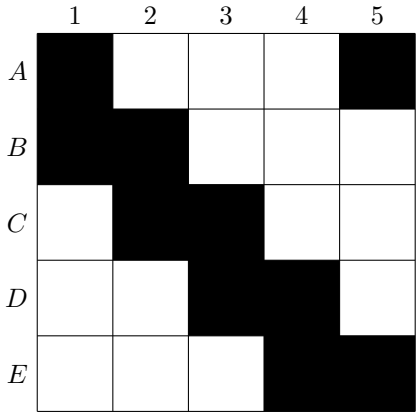
Problem 7.16. *Suppose we have a dinner party with ten guests, who are to be seated at a round table. There are five married couples, each consisting of a man and a woman. We want to arrange the seating so that the men alternate with the women, and no one sits next to their spouse. How many ways are there to do this? (This is sometimes called the ménage problem.)*

Solution. Interactive demo

We label the seats as follows:



The first choice to make is whether the women get the seats with letters or the seats with numbers; this will give an overall factor of two. For the rest of the analysis, we will assume that the women get the letters. Next, there are $5! = 120$ ways to assign the 5 women to seats $A, \ldots, E$. Now suppose we have seated the women, and we want to seat the men. We will refer to the woman in seat A as woman A, and to her husband as man A, and so on. Now man A is not supposed to sit next to his wife, so he can sit in seats 2, 3 or 4 but not 1 or 5. Similarly, man B can sit in seats 3, 4 or 5, and man C can sit in seats 4, 5 or 1, and so on. These rules can be represented by the following chessboard diagram:



Our seating problem is equivalent to the problem of finding a full matching for this board. (A rook in position $B4$ corresponds to putting man B in seat 4, for example.) We will show later that there are 13 such matchings. This means that there are $2 \times 120 \times 13 = 3120$ possible solutions to the original problem.

8. Reduction theorems

We will now introduce two results that allow us to calculate the rook polynomial of a board in terms of rook polynomials of smaller or simpler boards.

Theorem 8.1. *Let B be an $n \times n$ board, in which some squares may be blocked off. Let s be an unblocked square. Let C be the same as B, except that s is blocked off. Let D be the same as B, except that s's row and column are removed. Then for $k > 0$ we have $c_k(B) = c_k(C) + c_{k-1}(D)$, and so*

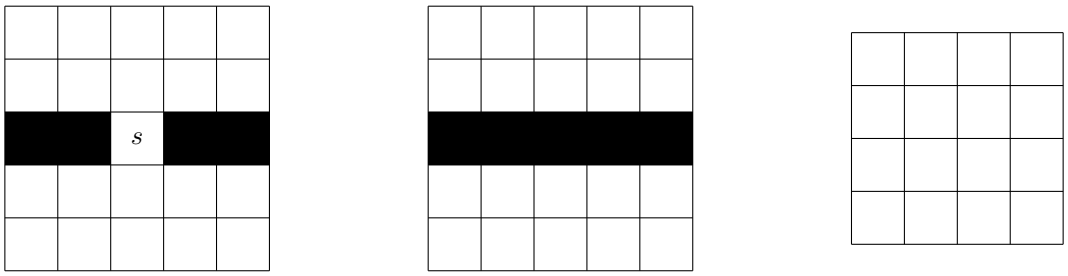
$$r_B(x) = r_C(x) + x\, r_D(x).$$

Proof. Interactive demo

The constant term on both sides is equal to one. Consider instead the coefficients of x^k, where $k > 0$. The coefficient in $r_B(x)$ is $c_k(B)$, which is the number of ways of placing k non-challenging rooks on B. We can write this as $c_k(B) = p + q$, where p is the number of placements that do not have a rook at s, and q is the number of placements that do have a rook at s. To place k rooks on B without using s is the same as to place k rooks on C, so $p = c_k(C)$. To place k rooks on B including s is the same as to place a rook at s, and then place $k-1$ more rooks on B, avoiding s's row and column. This in turn is the same as placing $k-1$ rooks on D, so we have $q = c_{k-1}(D)$. Note also that $c_{k-1}(D)$ is the coefficient of x^{k-1} in $r_D(x)$, which is the same as the coefficient of x^k in $x\,r_D(x)$. The equation $c_k(B) = p + q$ now tells us that $c_k(B) = c_k(C) + c_{k-1}(D)$, or that the coefficients of x^k are the same in $r_B(x)$ and $r_C(x) + x\,r_D(x)$. As this works for all k, we have $r_B(x) = r_C(x) + x\,r_D(x)$ as claimed. □

Remark 8.2. We say that C is the result of *blocking* s, and that D is the result of *stripping* s.

Example 8.3. Boards B, C and D could be as follows:



B C D

As all squares in D are white, Proposition 7.11 gives

$$r_D(x) = \sum_{k=0}^{4} \binom{4}{k}^2 k!x^k = 1 + 16x + 72x^2 + 96x^3 + 24x^4.$$

Next, in C the middle row is fully blacked out, and it is easy to see that this makes it irrelevant, so C is equivalent to a 4×5 board in which all squares are white. We can therefore use Proposition 7.11 again to get

$$r_C(x) = \sum_{k=0}^{4} \binom{4}{k}\binom{5}{k} k!x^k = 1 + 20x + 120x^2 + 240x^3 + 120x^4.$$

The theorem now gives

$$r_B(x) = r_C(x) + x\,r_D(x) = 1 + 21x + 136x^2 + 312x^3 + 216x^4 + 120x^5.$$

Example 8.4. It is interesting to see how Theorem 8.1 works out in some trivial cases. First consider the linear board L_n from Example 7.5, so L_n just consists of n blank squares in a row. Blocking any square gives L_{n-1}, and stripping any square gives the empty board, so Theorem 8.1 gives $r_{L_n}(x) = r_{L_{n-1}}(x) + x\,r_\emptyset(x)$. Even for the empty board, there is a unique way of placing no rooks, so we have $r_\emptyset(x) = 1$, so $r_{L_n}(x) = r_{L_{n-1}}(x) + x$. From this it follows inductively that $r_{L_n}(x) = 1 + nx$ (which is already obvious from the definitions, as we remarked in Example 7.5.)

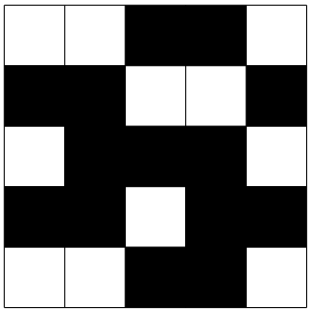
Now consider the diagonal board D_n from Example 7.6. Here blocking any square gives D_{n-1}, and stripping any square also gives D_{n-1}. Thus, Theorem 8.1 gives $r_{D_n}(x) = (1+x)r_{D_{n-1}}(x)$. From this it follows inductively that $r_{D_n}(x) = (1+x)^n$ (which is already obvious from the definitions, as we remarked in Example 7.6.)

Definition 8.5. Let B be an $n \times n$ board with some squares blocked off, as before. Suppose that the set of unblocked squares has been split into two parts C and D, so that $B = C \cup D$ and $C \cap D = \emptyset$. We say that C and D are *fully disjoint* if

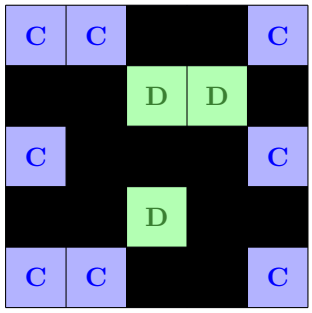
(a) There is no row that contains a square from C and also a square from D; and
(b) There is no column that contains a square from C and also a square from D.

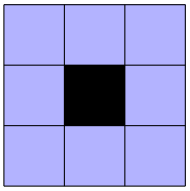
Example 8.6. Interactive demo

Consider the board B as shown on the left below.

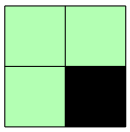


B





C



D

We can divide B into disjoint subboards C and D, as shown by the second picture. Note that:

(a) Rows $1, 3$ and 5 contain only C's, whereas rows 2 and 4 contain only D's.

(b) Columns $1, 2$ and 5 contain only C's, whereas columns 3 and 4 contain only D's.

This shows that C and D are fully disjoint.

Theorem 8.7. *Let B be an $n \times n$ board with some squares blocked off, as before. Suppose that B can be split into two subboards C and D, which are fully disjoint. Note that C and D can be regarded as boards in their own right, so they have rook polynomials $r_C(x)$ and $r_D(x)$. Then*

$$r_B(x) = r_C(x)\, r_D(x).$$

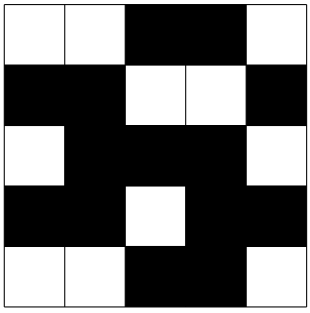
(We call this method factoring*.)*

Proof. To place k rooks in B, we first need to decide how many of them will be in C, and how many in D. If we place i rooks in C, then there will be $k - i$ in D. There are $c_i(C)$ ways of placing i non-challenging rooks in C, and $c_{k-i}(D)$ ways of placing $k - i$ non-challenging rooks in D. Moreover, the rooks in C cannot challenge those in D or *vice-versa*, because of conditions (b) and (c) in the theorem. Thus, we can put the two sets of rooks together, and we always have a set of k non-challenging rooks in B. From this we see that there are $c_i(C) \times c_{k-i}(D)$ ways of placing k non-challenging rooks in B, with precisely i of them in C. By considering all possible values of i, we get $c_k(B) = \sum_{i=0}^{k} c_i(C)c_{k-i}(D)$. This in turn gives

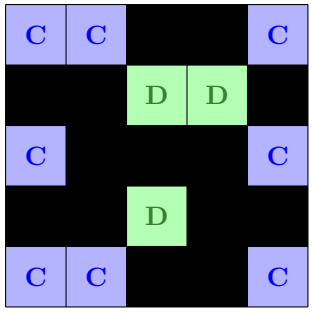
$$\begin{aligned} r_B(x) &= \sum_{k\geq 0} c_k(B)x^k = \sum_{k\geq 0}\sum_{i=0}^{k} c_i(C)c_{k-i}(D)x^k \\ &= \left(\sum_{i\geq 0} c_i(C)x^i\right)\left(\sum_{j\geq 0} c_j(D)x^j\right) = r_C(x)r_D(x). \end{aligned}$$

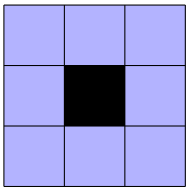
□

Example 8.8. Consider again the board B from example 8.6

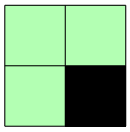


B





C



D

Boards C and D are small enough that we can calculate the rook polynomials by inspection:

$$r_C(x) = 1 + 8x + 14x^2 + 4x^3$$
$$r_D(x) = 1 + 3x + x^2.$$

Theorem 8.7 therefore gives

$$\begin{aligned} r_B(x) = r_C(x)r_D(x) &= (1 + 8x + 14x^2 + 4x^3)(1 + 3x + x^2) \\ &= 1 + 11x + 39x^2 + 54x^3 + 26x^4 + 4x^5. \end{aligned}$$

Problem 8.9. *Consider the following table, in which all the entries are in the range 1-5:*

1	2	3	4	5
2	4	5	3	1
4	3	1	5	2

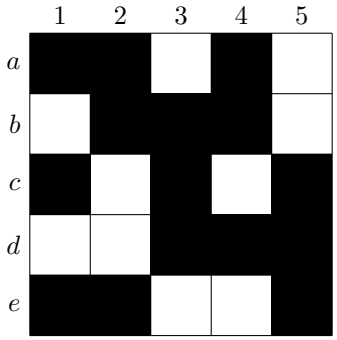
You can check that there are no repeats in any row or column. How many ways are there of adding a fourth row, so that there are still no repeats in any row or column?

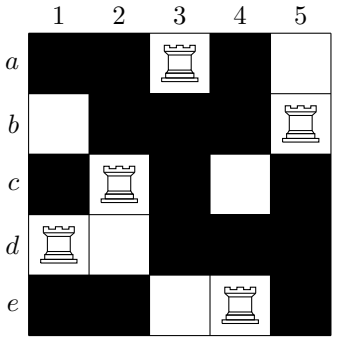
Solution. Interactive demo

We start by extending the table a little:

a	b	c	d	e
1	2	3	4	5
2	4	5	3	1
4	3	1	5	2
$\{3,5\}$	$\{1,5\}$	$\{2,4\}$	$\{1,2\}$	$\{3,4\}$

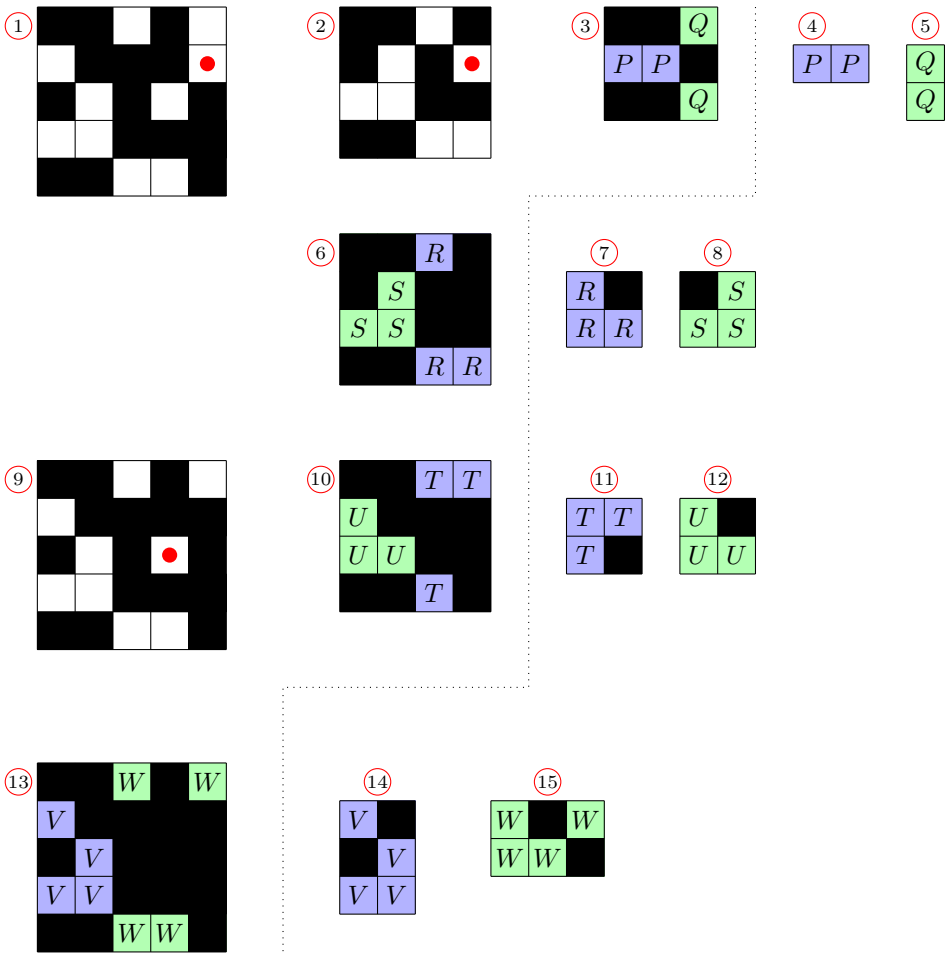
We want to add new row, in such a way that there are no repeats in any row or column. Column a already contains 1, 2 and 4, so the new entry cannot be any of those, so it must be in the set $\{3,5\}$. Column b already contains 2, 4 and 3, so the new entry cannot be any of those, so it must be in the set $\{1,5\}$. In the same way, we see that the new entries in columns c, d and e must be taken from the sets $\{2,4\}$, $\{1,2\}$ and $\{3,4\}$, as indicated in the diagram. This gives us a matching problem, with chessboard diagram as shown on the left below. Adding a new row is equivalent to solving the rook problem for this board. For example, one possible solution is to place rooks at $a3$, $b5$, $c2$, $d1$, $e4$, as shown in the middle picture. This indicates that we can legitimately add a new row to our table, with column a containing 3, column b containing 5, column c containing 2, column d containing 1 and column e containing 4. The result is shown in the right-hand picture.





We want to find the number of ways to add a new row, which is the same as the number of ways of placing 5 non-challenging rooks on the board, which is the coefficient of x^5 in the rook polynomial. As an illustration of the relevant methods, we will in fact calculate the full rook polynomial, by repeatedly using Theorems 8.1 and 8.7.

The following picture shows boards $B_1, \ldots, B_{15}$. We write $r_k(x)$ for the rook polynomial of board B_k. Our problem is to calculate $r_1(x)$.



By inspection, we have

$$r_4(x) = r_5(x) = 1 + 2x$$
$$r_7(x) = r_8(x) = r_{11}(x) = r_{12}(x) = 1 + 3x + x^2$$
$$r_{14}(x) = r_{15}(x) = 1 + 4x + 3x^2.$$

Now consider board B_3. We have divided the empty cells into two groups, marked P and Q respectively. There is no row that contains both P and Q, and there is no column that contains both P and Q, so P and Q are fully disjoint. Thus, Theorem 8.7 is applicable, and we get $r_3(x) = r_4(x)r_5(x)$. We can also factor boards B_6, B_{10} and B_{13} in a similar way. We find that

$$r_3(x) = r_4(x)r_5(x) = (1 + 2x)^2 = 1 + 4x + 4x^2$$
$$r_6(x) = r_7(x)r_8(x) = (1 + 3x + x^2)^2 = 1 + 6x + 11x^2 + 6x^3 + x^4$$
$$r_{10}(x) = r_{11}(x)r_{12}(x) = (1 + 3x + x^2)^2 = 1 + 6x + 11x^2 + 6x^3 + x^4$$
$$r_{13}(x) = r_{14}(x)r_{15}(x) = (1 + 4x + 3x^2)^2 = 1 + 8x + 22x^2 + 24x^3 + 9x^4.$$

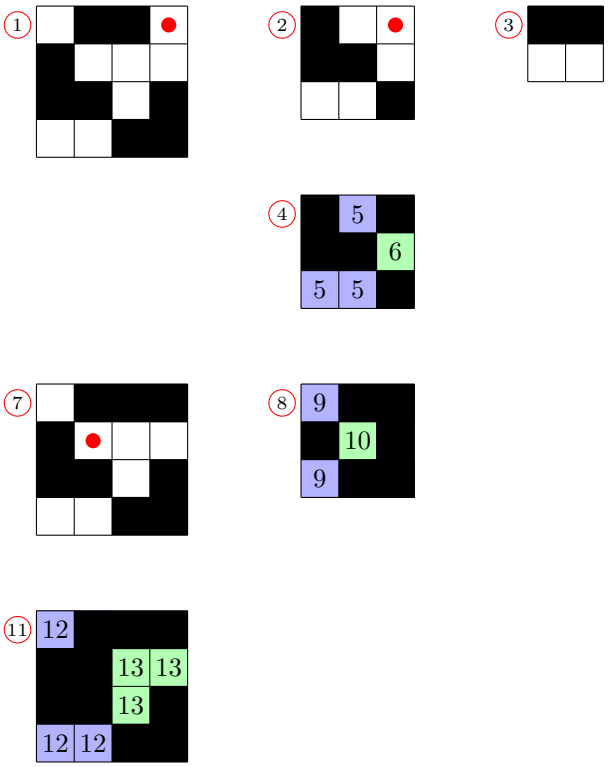
Now consider board B_9, in which we have marked one square with a red dot. Blocking that square gives B_{13}, and stripping it gives B_{10}. Theorem 8.1 therefore tells us that $r_9(x) = r_{13}(x) + x\,r_{10}(x)$. We can also block and strip the marked square in B_2 to get B_6 and B_3, or we can block and strip the marked square in B_1 to get B_9 and B_2. We therefore have

$$\begin{aligned}
r_9(x) &= r_{13}(x) + x\,r_{10}(x) = (1 + 8x + 22x^2 + 24x^3 + 9x^4) + x\,(1 + 6x + 11x^2 + 6x^3 + x^4) \\
&= 1 + 9x + 28x^2 + 35x^3 + 15x^4 + x^5 \\
r_2(x) &= r_6(x) + x\,r_3(x) = (1 + 6x + 11x^2 + 6x^3 + x^4) + x\,(1 + 4x + 4x^2) \\
&= 1 + 7x + 15x^2 + 10x^3 + x^4 \\
r_1(x) &= r_9(x) + x\,r_2(x) = (1 + 9x + 28x^2 + 35x^3 + 15x^4 + x^5) + x\,(1 + 7x + 15x^2 + 10x^3 + x^4) \\
&= 1 + 10x + 35x^2 + 50x^3 + 25x^4 + 11x^4 + 2x^5.
\end{aligned}$$

(If we wanted to save writing, it would not be hard to do this calculation without explicitly drawing any of the boards to the right of the dotted line.) In particular, the number of ways of adding an admissible row to our original table is the number of ways of placing 5 non-challenging rooks on B_1, which is the coefficient of x^5 in $r_1(x)$, which is 2. In fact, the two possible rook placements are $a3b5c2d1e4$ and $a5b1c4d2e3$, so the two possibilities for the extra row are $(3, 5, 2, 1, 4)$ and $(5, 1, 4, 2, 3)$.

Problem 8.10. *Calculate the full rook polynomial for the board in Problem 7.14.*

Solution. We use the same method as in the previous example, but written in a slightly more efficient way.



We have boards $B_1, \ldots, B_{13}$. However, B_5 and B_6 have not been drawn separately, they are just marked as subsets of B_4. Similarly, B_9 and B_{10} are just marked as subsets of B_8, and B_{12} and B_{13} are subsets of B_{11}. By inspection, we have

$$\begin{aligned}
r_3(x) &= r_9(x) = 1 + 2x \\
r_5(x) &= r_{12}(x) = r_{13}(x) = 1 + 3x + x^2 \\
r_6(x) &= r_{10}(x) = 1 + x.
\end{aligned}$$

We next want to use the factoring theorem (Theorem 8.7) to calculate $r_4(x)$, $r_8(x)$ and $r_{11}(x)$. For that result to be applicable, we must check that the relevant subboards are fully disjoint. This is equivalent to the following claim: in boards B_4, B_8 and B_{11}, any two squares in the same row have the same number, and any two squares in the same column have the same number. This is clear by inspection. We therefore have

$$\begin{aligned} r_4(x) &= r_5(x)r_6(x) = (1+3x+x^2)(1+x) = 1+4x+4x^2+x^3 \\ r_8(x) &= r_9(x)r_{10}(x) = (1+2x)(1+x) = 1+3x+2x^2 \\ r_{11}(x) &= r_{12}(x)r_{13}(x) = (1+3x+x^2)^2 = 1+6x+11x^2+6x^3+x^4. \end{aligned}$$

Next, blocking and stripping the marked square in B_7 gives B_{11} and B_8. Similarly, B_2 gives B_4 and B_3, and B_1 gives B_7 and B_2. We therefore have

$$\begin{aligned} r_7(x) &= r_{11}(x) + x\, r_8(x) = 1+6x+11x^2+6x^3+x^4 + x\,(1+3x+2x^2) \\ &= 1+7x+14x^2+8x^3+x^4 \\ r_2(x) &= r_4(x) + x\, r_3(x) = 1+4x+4x^2+x^3 + x\,(1+2x) \\ &= 1+5x+6x^2+x^3 \\ r_1(x) &= r_7(x) + x\, r_2(x) = 1+7x+14x^2+8x^3+x^4 + x\,(1+5x+6x^2+x^3) \\ &= 1+8x+19x^2+14x^3+2x^4. \end{aligned}$$

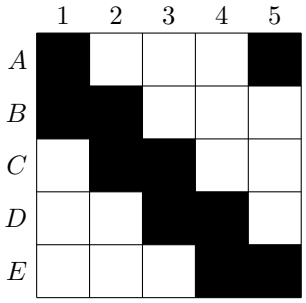
In particular, the number of ways of allocating the jobs in Problem 7.14 is the same as the number of ways of placing 4 non-challenging rooks on B_1, which is the coefficient of x^4 in $r_1(x)$, which is 2. This is the same answer as we found previously in Problem 7.14.

9. Tabular methods

There is a tabular method for finding all the full matchings for a given $n \times n$ board.

Example 9.1. Interactive demo

The picture on the left below is the board from Problem 7.16, and the table on the right finds all the 13 solutions.



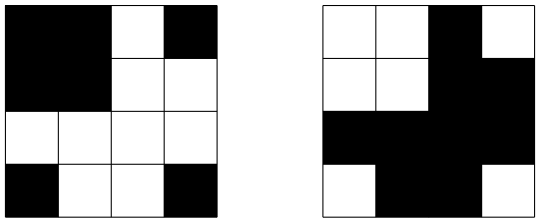
$A3$	$B3$	$C1$	$\underline{D5}$		✗
		$C4$	$D1$		✗
			$\underline{D5}$	$\underline{E1}$	✓
		$\underline{C5}$	$\underline{D1}$		✗
	$B4$	$C1$	$\underline{D5}$	$\underline{E3}$	✓
		$\underline{C5}$	$\underline{D1}$	$\underline{E3}$	✓
	$\underline{B5}$	$C1$			✗
		$\underline{C4}$	$\underline{D1}$	$\underline{E3}$	✓
$A3$	$B4$	$C1$	$D2$		✗
			$\underline{D5}$	$\underline{E2}$	✓
		$\underline{C5}$	$D1$	$\underline{E2}$	✓
			$\underline{D2}$	$\underline{E1}$	✓
	$\underline{B5}$	$C1$	$\underline{D2}$		✗
		$\underline{C4}$	$D1$	$\underline{E2}$	✓
			$\underline{D2}$	$\underline{E1}$	✓
$\underline{A4}$	$B3$	$C1$	$D2$		✗
			$\underline{D5}$	$\underline{E2}$	✓
		$\underline{C5}$	$D1$	$\underline{E2}$	✓
			$\underline{D2}$	$\underline{E1}$	✓
	$\underline{B5}$	$\underline{C1}$	$\underline{D2}$	$\underline{E3}$	✓

We will not explain the method in detail here, as it is much easier to understand by working through the interactive demonstration.

10. Inclusion-exclusion for matching problems

Definition 10.1. Let B be an $n \times n$ board, with each square coloured black or white as before. We then write $\overline{B}$ for the board with the colours exchanged, so the white squares for B are the black squares for $\overline{B}$ and *vice-versa*. We call $\overline{B}$ the *complement* of B.

Example 10.2. B and $\overline{B}$ could be as follows.



B $\qquad$ $\overline{B}$

As before, we write $c_k(B)$ for the number of ways of placing k non-challenging rooks on B, or equivalently the numbers of partial matchings of size k for the corresponding matching problem. Similarly, we write $c_k(\overline{B})$ for the number of ways of placing k non-challenging rooks on $\overline{B}$. We are particularly interested in $c_n(\overline{B})$, which is the number of full matchings for $\overline{B}$. It turns out that if we know $c_0(B), \dots, c_n(B)$ then we can work out $c_n(\overline{B})$:

Theorem 10.3. $c_n(\overline{B}) = \sum_{k=0}^{n} (-1)^k (n-k)! c_k(B)$.

We will prove this after a preliminary result. The Inclusion-Exclusion Principle (from Section 5) will play a central rôle in the proof. We will then use this theorem to solve the snap problem (Problem 7.15) and the ménage problem (Problem 7.16).

Definition 10.4. We write S for the set of ways of placing n non-challenging rooks on the full $n \times n$ board (so $|S| = n!$ by Proposition 7.11). Now let x be a position on the board. We define $S_x \subseteq S$ to be the set of rook placements that include a rook at x. More generally, given a set $X = \{x_1, \dots, x_m\}$ of board positions, we put

$$\begin{aligned} S_X &= \bigcap_i S_{x_i} = S_{x_1} \cap \cdots \cap S_{x_m} \\ &= \{ \text{ non-challenging rook placements with a rook at each position } x_i \}. \end{aligned}$$

Lemma 10.5. *Consider a set $X = \{x_1, \dots, x_m\}$ as above.*

(a) *If there is a row containing two of the positions x_i, or a column containing two of the positions x_i, then $S_X = \emptyset$ and so $|S_X| = 0$.*

(b) *Otherwise, we have $|S_X| = (n-m)!$.*

Proof. Interactive demo

- Suppose that some row contains both x_i and x_j (with $i \neq j$). Then if we place two rooks at x_i and x_j, then they will challenge each other, so the placement does not count as an element of S_X. This means that it is impossible to have any elements of S_X, so $S_X = \emptyset$ and $|S_X| = 0$.
- The same applies if some column contains both x_i and x_j with $i \neq j$.
- Suppose instead that the x_i are all in different rows, say $r_1, \dots, r_m$, and also in different columns, say $c_1, \dots, c_m$. Let A be the set of rows other than $r_1, \dots, r_m$, so $|A| = n - m$. Let B be the set of columns other than $c_1, \dots, c_m$, so $|B| = n - m$. To get a full rook placement in S_X we need to place rooks on X, then place $n-m$ rooks on the board $A \times B$, which is an $(n-m) \times (n-m)$ board with no black squares. Proposition 7.11 tells us that there are $(n-m)!$ ways to do this, so $|S_X| = (n-m)!$.

□

Proof of Theorem 10.3. We need to understand the set $C_n(\overline{B}) \subseteq S$. Consider a rook placement P for the full board, so $P \in S$. We claim that $P \in \bigcup_{x \in B} S_x$ iff $P \notin C_n(\overline{B})$. Indeed, we have $P \in \bigcup_{x \in B} S_x$ iff there exists a square $x \in B$ such that $P \in S_x$. However, we have $P \in S_x$ iff the placement P has a rook at x. Here x is a white square of B so it is a black square of $\overline{B}$, so if the placement has a rook at x, then it does not count as a placement on $\overline{B}$, so $P \notin C_n(\overline{B})$. This argument is reversible, so $P \in \bigcup_{x \in B} S_x$ iff $P \notin C_n(\overline{B})$. By the contrapositive, we have $P \in C_n(\overline{B})$ iff $P \notin \bigcup_{x \in B} S_x$, or in other words $P \in S \setminus \bigcup_x S_x$. The negative form of the IEP now tells us that

$$c_n(\overline{B}) = |C_n(\overline{B})| = \sum_{X \subseteq B} (-1)^{|X|} |S_X|.$$

If X is not a non-challenging rook placement on B, then $|S_X| = 0$ by Lemma 10.5(a), so we can ignore these terms. On the other hand, if X is a placement of k non-challenging rooks on B, then $|S_X| = (n-k)!$ by Lemma 10.5(a). The number of terms of this type is $c_k(B)$, so the sum of all terms of this type is $(-1)^k(n-k)!c_k(B)$. Putting this together, we get $c_n(\overline{B}) = \sum_{k=0}^{n}(-1)^k(n-k)!c_k(B)$ as claimed. □

Remark 10.6. We can apply Theorem 10.3 to $\overline{B}$ and note that $\overline{\overline{B}} = B$; this gives

$$c_n(B) = \sum_{k=0}^{n}(-1)^k(n-k)!c_k(\overline{B}).$$

Example 10.7. As a trivial example, we have $c_n(F_n) = \sum_{k=0}^{n}(-1)^k(n-k)!c_k(\overline{F_n})$. However, all squares in $\overline{F_n}$ are black, so we cannot place any rooks there, so $c_k(\overline{F_n}) = 0$ for $k > 0$. On the other hand, we always have $c_0 = 1$, so our equation becomes

$$c_n(F_n) = (-1)^0 n! c_0(\overline{F_n}) = n!,$$

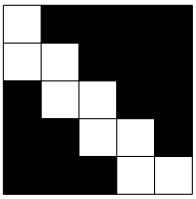
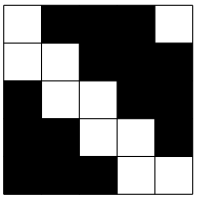
and we knew this already from Proposition 7.10.

Example 10.8. Recall from Example 7.6 that D_n is the $n \times n$ board with only the diagonal squares coloured white, and that $c_k(D_n) = \binom{n}{k} = n!/(k!(n-k)!)$. We thus have

$$c_n(\overline{D_n}) = \sum_{k=0}^{n}(-1)^k(n-k)!c_k(D_n) = n!\sum_{k=0}^{n-1}(-1)^k/k!.$$

On the other hand, Problem 7.13 shows that $C_n(\overline{D_n})$ is the set of derangements of $\{1,\dots,n\}$. We saw in Proposition 5.11 that the number of derangements is $n!\sum_{k=0}^{n-1}(-1)^k/k!$, so everything is consistent.

Definition 10.9. The board Q_n consists of $2n-1$ white squares in an $n \times n$ board arranged in a zigzag pattern as illustrated on the left below. (To see that there are $2n-1$ white squares, note that there are n rows, and each row contains two white squares, except for the first row, which contains only one white square.) We also consider the board Q'_n, which has an extra white square at the top right, making $2n$ white squares in total.

$Q_5 =$  $Q'_5 =$ 

We call a board of type Q_n a *staircase*.

Proposition 10.10. *For $0 \leq k \leq n$ we have $c_k(Q_n) = \binom{2n-k}{k}$. We also have*

$$c_k(Q'_n) = c_k(Q_n) + c_{k-1}(Q_{n-1}) = \binom{2n-k}{k} + \binom{2n-1-k}{k-1}.$$

Proof. Interactive demo

By inspection, two rooks on the staircase challenge each other iff they are adjacent. Thus, the allowable rook placements are just subsets of the staircase that have no adjacent elements; in other words, they must be gappy, as in Definition 1.21. Thus, $c_k(Q_n)$ is the number of gappy subsets of size k in the staircase. Proposition 1.23 therefore gives $c_k(Q_n) = \binom{m-k+1}{k}$, where m is the number of cells in the staircase. This is just $m = 2n-1$, so we get $c_k(Q_n) = \binom{2n-k}{k}$.

Now consider Q'_n. Blocking the top right square gives Q_n, and stripping it gives a reflected copy of Q_{n-1}. Thus, Theorem 8.1 gives $c_k(Q'_n) = c_k(Q_n) + c_{k-1}(Q_{n-1})$, which is $\binom{2n-k}{k} + \binom{2n-1-k}{k-1}$ by the previous paragraph. □

We can now complete our analysis of the ménage problem.

Example 10.11. Using Proposition 10.10, we see that the rook polynomial coefficients of Q_5' are as follows:

$$\begin{aligned}
c_0(Q_5') &= 1 \\
c_1(Q_5') &= \binom{9}{1} + \binom{8}{0} = 10 \\
c_2(Q_5') &= \binom{8}{2} + \binom{7}{1} = 35 \\
c_3(Q_5') &= \binom{7}{3} + \binom{6}{2} = 50 \\
c_4(Q_5') &= \binom{6}{4} + \binom{5}{3} = 25 \\
c_5(Q_5') &= \binom{5}{5} + \binom{4}{4} = 2.
\end{aligned}$$

Now note that the board B considered in Problem 7.16 is just the complement of Q_5'. It follows that

$$\begin{aligned}
c_5(B) &= 5!c_0(Q_5') - 4!c_1(Q_5') + 3!c_2(Q_5') - 2!c_3(Q_5') + 1!c_4(Q_5') - 0!c_5(Q_5') \\
&= 120 \times 1 - 24 \times 10 + 6 \times 35 - 2 \times 50 + 1 \times 25 - 1 \times 0 \\
&= 13.
\end{aligned}$$

Thus, there are precisely 13 full matchings for B. As explained in Problem 7.16, it follows that there are $2 \times 120 \times 13 = 3120$ ways to solve the seating problem described there.

We can also now complete our analysis of the snap problem.

Example 10.12. Let B denote the board shown in Problem 7.15, which is a 16×16 board with four black blocks of size 4×4 on the diagonal. This means that $\overline{B}$ consists of four fully disjoint copies of F_4. Proposition 7.11) tells us that

$$r_{F_4}(x) = \sum_{k=0}^{4} \binom{4}{k}^2 k! x^k = 1 + 16x + 72x^2 + 96x^3 + 24x^4.$$

Moreover, we can use Theorem 8.7 to show that

$$r_{\overline{B}}(x) = r_{F_4}(x)^4.$$

It would be a lot of work to expand this by hand, but it can easily be done with a computer. If $r_{F_4}(x)^4 = \sum_{i=0}^{16} a_i x^i$, we then find that $c_{16}(B) = \sum_i (-1)^i (16-i)! a_i$. Using a computer again, we find that $c_{16}(B) \simeq 2.483 \times 10^{11}$. This is the number of possible games with no snap. On the other hand, the total number of possible games is $16! \simeq 2.092 \times 10^{13}$. Thus, the probability of not getting a snap is

$$c_{16}(B)/52! \simeq \frac{2.483 \times 10^{11}}{2.092 \times 10^{13}} \simeq 0.012.$$

In other words, only about one in a hundred games will end without a snap. This was all for the restricted game where we only use aces, kings, queens and jacks. If we want to use the full pack, we need to expand $r_{F_4}(x)^{13}$ as $\sum_{i=0}^{52} b_i x^i$, and then calculate $c_{52}(B) = \sum_i (-1)^i (52-i)! b_i$, which works out to 1.309×10^{66}. We then divide by $52! \simeq 8.066 \times 10^{67}$ to give a probability of approximately 0.016.

11. Hall's marriage theorem

Video (Up to Lemma 11.5)

Consider a matching problem, with a set A of people, a set B of jobs, and a set $E \subseteq A \times B$ consisting of pairs (a, b) where person a is qualified for job b.

Definition 11.1. We say that the matching problem is *solvable* if there exists a full matching. This means that it is possible to allocate every job to a qualified person, in such a way that no one has more than one job.

Suppose we want to decide whether a given matching problem is solvable. One approach would be to use the methods that we have already seen to count or tabulate the full matchings; that will in particular tell us whether there are any full matchings. However, if we are only interested in the existence question, then some different methods become available, as we will describe in this section.

Definition 11.2. Recall that in Remark 6.5, we defined

$$C_b = \{a \in A \mid (a, b) \in E\} = \{ \text{ people who are qualified to do job } b\}.$$

We will refer to this as the *candidate set* for b. More generally, if $U \subseteq B$ is some subset of the jobs, we put

$$\begin{aligned} C_U &= \bigcup_{b \in U} C_b = \{a \in A \mid (a, b) \in E \text{ for some } b \in U\} \\ &= \{ \text{ people who are qualified for at least one of the jobs in } U\}. \end{aligned}$$

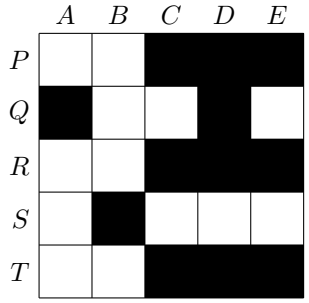
We again call this the *candidate set* for U. We say that

(a) U is *implausible* if $|C_U| < |U|$.
(b) U is *barely plausible* if $|C_U| = |U|$.
(c) U is *very plausible* if $|C_U| > |U|$.
(d) U is *plausible* if $|C_U| \geq |U|$ (so U is either barely plausible or very plausible).

We also say that the whole matching problem is *plausible* if every subset $U \subseteq B$ is plausible.

Remark 11.3. We always have $C_\emptyset = \emptyset$ so $|C_\emptyset| = 0 = |\emptyset|$ so $\emptyset$ is barely plausible.

Example 11.4. Suppose we have a set $A = \{\text{Paula}, \text{Quentin}, \text{Ruth}, \text{Steve}, \text{Tessa}\}$ of people, and a set $B = \{\text{Artist}, \text{Baker}, \text{Courier}, \text{Dentist}, \text{Electrician}\}$ of jobs, with qualifications as follows:



We then have

$$\begin{aligned} C_{\{B,D\}} &= \{ \text{ people qualified to be a baker or a dentist } \} = C_B \cup C_D \\ &= \{P, Q, R, T\} \cup \{S\} = \{P, Q, R, S, T\}. \end{aligned}$$

Proceeding in the same way, we see that

- $C_{\{E\}} = \{Q, S\}$, so $|C_{\{E\}}| = 2 > 1 = |\{E\}|$, so $\{E\}$ is very plausible.
- $C_{\{A,B\}} = \{P, Q, R, S, T\}$, so $|C_{\{A,B\}}| = 5 > 2 = |\{A, B\}|$, so $\{A, B\}$ is very plausible.
- $C_{\{C,D\}} = \{Q, S\}$, so $|C_{\{C,D\}}| = 2 = |\{C, D\}|$, so $\{C, D\}$ is barely plausible.
- $C_{\{C,D,E\}} = \{Q, S\}$, so $|C_{\{C,D,E\}}| = 2 < 3 = |\{C, D, E\}|$, so $\{C, D, E\}$ is implausible.

This last example is a problem. We need to assign jobs C, D and E, and we can ignore people P, R and T, because none of them can do any of these jobs. This just leaves two people Q and S who need to cover three jobs, which is not possible. Thus, our allocation problem is not solvable. The next lemma will generalise this line of argument.

Lemma 11.5. *If the problem is solvable, then it is plausible. Thus, by the contrapositive, if the problem is not plausible then it has no solution.*

Proof. If the problem is solvable, then we can choose a full matching M. Consider a subset $U = \{b_1, \dots, b_m\} \subseteq B$. Then M must allocate each job $b_i \in U$ to some person $a_i \in A$. As M is a matching, we know that the people $a_1, \dots, a_m$ are all different, and that a_i is qualified for job b_i, so $a_i \in C_U$. We thus have distinct elements $a_1, \dots, a_m \in C_U$, so $|C_U| \geq m = |U|$, so the set U is plausible. This holds for all U, so the whole matching problem is plausible. □

The main result of this section is the converse to the above lemma.

Theorem 11.6 (Hall's Marriage Theorem)**.** *If a matching problem is plausible, then it is solvable.*

Remark 11.7. Hall first formulated this result in the context of assigning romantic partners rather than jobs, hence the name.

We will give two different proofs of Hall's theorem. The first proof will rely on the following construction.

Construction 11.8. Video

Suppose we have a matching problem as above, and a partial matching M' that assigns some subset $B' \subseteq B$ of the jobs to some subset $A' \subseteq A$ of the people. (Note that M' gives a bijection from B' to A', so we automatically have $|A'| = |B'|$.) There is then an obvious way to set up a new matching problem for the assignment of the remaining jobs. In more detail, we take $A'' = A \setminus A'$ (the set of people who have not already been given a job), and $B'' = B \setminus B'$ (the set of jobs that still need to be assigned) and $E'' = E \cap (A' \times B')$. This gives a new matching problem, with candidate sets $C''_U = C_U \cap A'' = C_U \setminus A'$ for all $U \subseteq B''$. We call this the *completion problem* for M'. If we can find a solution M'' for the completion problem, then we can combine it with M' to get a solution for the original problem.

First proof of Theorem 11.6. Video

If there are no jobs, then there is nothing to do, and the problem is vacuously solved.

Suppose instead that there is only one job, say $B = \{b\}$. By hypothesis, the set $\{b\}$ is plausible, which means that $|C_b| \geq 1$, so $C_b \neq \emptyset$, so we can choose $a \in C_b$. This means that person a is qualified to do job b, so we can just allocate b to a, and there is nothing more to do.

Now suppose that $|B| = n > 1$. We can assume by induction that any plausible problem with at most $n - 1$ jobs can be solved. By an *intermediate set* we mean a set $U \subseteq B$ with $U \neq \emptyset$ and $U \neq B$. Note that one of the following two cases must hold:

(easy) Every intermediate set is very plausible.
(hard) There is an intermediate set $B' \subseteq B$ such that B' is barely plausible.

First consider the easy case. Choose any job $b_0 \in B$, and put $B' = \{b_0\}$ and $B'' = B \setminus B'$. As $\{b_0\}$ is plausible, we have $|C_{b_0}| \geq 1$, so we can choose $a_0 \in C_{b_0}$ and put $A' = \{a_0\}$ and $A'' = A \setminus A'$. Let M' be the partial matching that assigns b_0 to a_0. We claim that the corresponding completion problem is plausible. In other words, for all $U \subseteq B''$, we claim that $|C''_U| \geq |U|$. If U is empty, this holds by Remark 11.3. Suppose instead that $U \neq \emptyset$. As $U \subseteq B''$, we also have $U \neq B$. As we are in the easy case, it follows that U is very plausible, so $|C_U| > |U|$, or equivalently $|C_U| - 1 \geq |U|$. We also have $C''_U = C_U \setminus \{a_0\}$, so $|C''_U|$ is either $|C_U| - 1$ (if $a_0 \in C_U$) or $|C_U|$ (if $a_0 \notin C_U$). Either way, we have $|C''_U| \geq |C_U| - 1 \geq |U|$, as required. As C'' is a plausible problem with $n - 1$ jobs, our induction hypothesis guarantees that it has a solution, say M''. We can combine this with M' to get a solution for the original problem.

Now suppose instead that we are in the hard case. We can therefore choose an intermediate set B' (so $B' \subset B$ with $\emptyset \neq B' \neq B$) such that B' is barely plausible. Put $B'' = B \setminus B'$. Note that $|B'| < n$ and $|B''| < n$, so our induction hypothesis guarantees that any plausible allocation problem for B' or for B'' is solvable. In particular, we can restrict our original allocation problem to B', and there must exist a solution for this, say M'. This assigns the jobs in B' to some set of people $A' \subseteq A$, which must have $|A'| = |B'|$. We again claim that the completion problem for M' is plausible. However, the proof is a little more complicated than in the easy case. The first point to note is that $A' = C_{B'}$. To see this, note that each person in A' has been assigned a job in B', and M' is assumed to be a valid partial matching so they must be qualified for the job that they have been assigned, so $A' \subseteq C_{B'}$. On the other hand, we have already remarked that $|A'| = |B'|$, and $|B'| = |C_{B'}|$ because B' is assumed to be barely plausible, so $|A'| = |C_{B'}|$. As $A' \subseteq C_{B'}$ with $|A'| = |C_{B'}|$, we see that $A' = C_{B'}$ as claimed. Now consider a subset $U \subseteq B''$. Recall that $C''_U = C_U \setminus A'$, essentially by definition. As $A' = C_{B'}$, this can be rewritten as $C''_U = C_U \setminus C_{B'}$. It is also easy to see that $C_{B' \cup U} = C_{B'} \cup C_U$ and so

$$C_{B' \cup U} \setminus C_{B'} = (C_{B'} \cup C_U) \setminus C_{B'} = C_U \setminus C_{B'} = C''_U.$$

This gives

$$|C''_U| = |C_{B' \cup U} \setminus C_{B'}| = |C_{B' \cup U}| - |C_{B'}| = |C_{B' \cup U}| - |A'|.$$

As the original problem is assumed to be plausible, we have $|C_{B' \cup U}| \geq |B' \cup U|$. Here $U \subseteq B''$, so B' and U are disjoint, so $|B' \cup U| = |B'| + |U| = |A'| + |U|$. Putting this together, we get

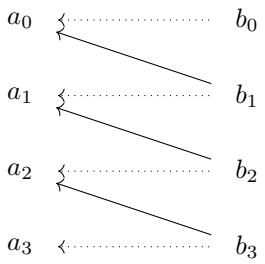
$$|C''_U| \geq (|A'| + |U|) - |A'| = |U|.$$

This shows that the completion problem C'' is plausible, as claimed. We also know that $|B''| < n$, so we can use our induction hypothesis to show that the completion problem has a solution. Just as in the easy case, we can choose a solution to the completion problem and combine it with M' to get a solution for the original problem, as required. □

The above proof is theoretically satisfying, but does not really provide much guidance about how to find a solution. We will therefore give a second proof which is a bit more complicated, but also more constructive. In this second proof, we try to allocate the jobs one by one. Suppose we have already allocated a subset $B' \subseteq B$ of the jobs, and we are trying to allocate one more job, say b_0. It might happen that we get stuck: all the people who are qualified to do b_0 have already been given one of the jobs in B'. To fix this, we need to change the allocation of the jobs in B' in order to free up someone who is qualified to do b_0, and the main problem is to analyse the possibilities for making this kind of adjustment.

Video (Definition 11.9 to the end of Section 11)

Definition 11.9. Suppose we have a matching problem as before, and a partial matching M', which assigns some subset $B' \subseteq B$ of the jobs. Let b_0 be an element of $B \setminus B'$, so job b_0 is not assigned by M'. By an *open zigzag* for (M', b_0), we mean a pattern like this:



M'

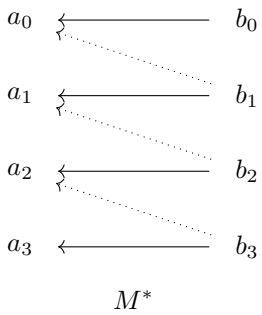
In more detail, an open zigzag is a sequence $(a_0, \ldots, a_r; b_0, \ldots, b_r)$ with properties as follows (the picture above shows $r = 3$).

- The people $a_0, \ldots, a_r$ are all different, and the jobs $b_0, \ldots, b_r$ are all different.
- For $i = 1, \ldots, r$, job b_i is assigned by M' to person a_{i-1} (so $b_i \in B'$, and a_{i-1} is qualified for b_i). This is indicated by the solid arrows in the picture.
- For $i = 0, \ldots, r$, person a_i is qualified for job b_i. This is indicated by the dotted arrows in the picture.
- Person a_r is not assigned a job by M'.

Given such an open zigzag, we define a new matching M^* for the jobs in $B' \cup \{b_0\}$ as follows:

- M^* assigns b_i to a_i for $i = 0, \ldots, r$
- For $b \in B' \setminus \{b_1, \ldots, b_r\}$, the matching M^* assigns b in the same way that M' does.

This can be illustrated as follows:



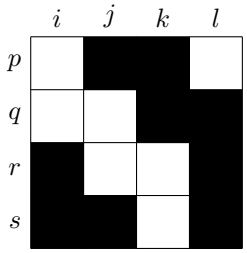
M^*

The process that constructs M^* from M' is called *flipping*.

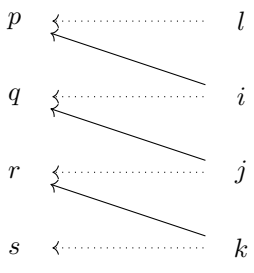
Remark 11.10. The easiest case is the case $r = 0$, where a_0 is qualified for b_0 and has not already been assigned a job by M', so M^* does not change any of the assignments made by M' but simply adds an assignment of b_0 to a_0.

Suppose we can prove that for every M' and b_0 as above, there always exists an open zigzag. We can then flip the zigzag to get a new matching that includes b_0. After doing this repeatedly, we will eventually get a full matching. Thus, the key problem is to prove the existence of open zigzags.

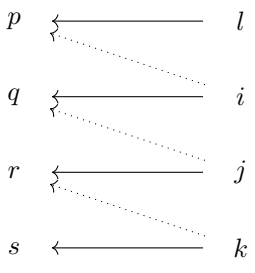
Example 11.11. Consider an allocation problem with people $A = \{p, q, r, s\}$ and jobs $B = \{i, j, k, l\}$ and qualifications as follows:



We have a partial matching M' assigning the first three jobs to the first three people in the obvious way: $i \mapsto p$ and $j \mapsto q$ and $k \mapsto r$. Now we seem to be stuck: the only person who can do l is p, but p is already doing i. We therefore remove p from i and give them l instead, leaving i unfilled. This creates another problem: the only remaining person qualified for i is q, but q is already doing j. We therefore remove q from j and give them i instead, leaving j unfilled. This creates another problem: the only remaining person qualified for j is r, but r is already doing k. We therefore remove r from k and give them j instead, leaving k unfilled. Now we are OK because person s is free and is qualified for k, so we can assign k to s and we have filled all jobs. What we have effectively done is to flip the following zigzag:

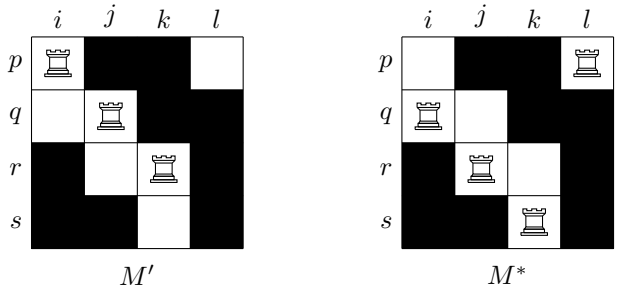


M'



M^*

We could also illustrate M' and M^* using rook placements, as follows:



Definition 11.12. By an *closed zigzag* for (M', b_0), we mean a sequence $x = (a_0, \dots, a_{r-1}; b_0, \dots, b_r)$ such that

- The people $a_0, \dots, a_{r-1}$ are all different, and the jobs $b_0, \dots, b_r$ are all different.
- For $i = 1, \dots, r$, job b_i is assigned by M' to person a_{i-1} (so $b_i \in B'$, and a_{i-1} is qualified for b_i).
- For $i = 0, \dots, r-1$, person a_i is qualified for job b_i.

In other words, a closed zigzag is like an open zigzag, except that the person a_r is not provided. We say that x is *openable* if there exists another person a_r who is qualified for b_r and is not assigned any job by M', so that the sequence $(a_0, \dots, a_r; b_0, \dots, b_r)$ is an open zigzag. We say that x is *extendable* if there exists another person a_r and another job b_{r+1} such that M' assigns b_{r+1} to a_r, but a_r is also qualified for b_r, so that the sequence $(a_0, \dots, a_r; b_0, \dots, b_{r+1})$ is a closed zigzag.

Definition 11.13. We define U to be the set of jobs that occur in some closed zigzag.

Remark 11.14. If a job b occurs in some closed zigzag, then we can discard everything after b, and that gives a closed zigzag that ends with b. Thus, we can also say that U is the set of jobs that appear at the end of some closed zigzag.

Remark 11.15. Suppose that $b \in U$, so there exists a closed zigzag $x = (a_0, \dots, a_{r-1}; b_0, \dots, b_r)$ with $b_r = b$. This usually means that M' assigns b to person a_{r-1}, so in particular $b \in B'$. The only exception is the case where $r = 0$. The sequence $(\emptyset; b_0)$ counts as a closed zigzag, showing that $b_0 \in U$, but $b_0 \notin B'$. We thus have $U = \{b_0\} \cup U'$, where $U' \subseteq B'$.

Second proof of Theorem 11.6. As before, we argue by induction on $|B|$. We choose any job $b_0 \in B$ and put $B' = B \setminus \{b_0\}$. The matching problem is still plausible when restricted to B', so the induction hypothesis gives us a partial matching M' that assigns all the jobs in B'. We just need to assign b_0 as well, which may involve changing some of the assignments for B'. We then define U and U' as in Definition 11.13 and Remark 11.15. Put $m = |U'|$ and note that $|U| = m+1$. Recall that $U' \subseteq B'$, so M' assigns all the jobs in U' to a set of people $V' \subseteq A'$. Because M' is a legitimate matching, it assigns different jobs to different people, so $|V'| = m$. Moreover, the people in V' are qualified for the jobs that they have been assigned, so $V' \subseteq C_U$. However, our matching problem is plausible by assumption, so $|C_U| \geq |U| = m + 1$. As $|C_U| > |V'|$, we can choose $a^* \in C_U$ such that $a^* \notin V'$. Because $a^* \in C_U$, we can choose a job $b^* \in U$ such that a^* is qualified for b^*. Because $b^* \in U$, we can choose a closed zigzag $x = (a_0, \dots, a_{r-1}; b_0, \dots, b_r)$ such that $b_r = b^*$. Note that $b_{i+1} \in U'$ for $i = 0, \dots, r-1$, and M' assigns b_{i+1} to a_i, so $a_i \in V'$. As $a^* \notin V'$, we see that a^* is different from $a_0, \dots, a_{r-1}$. Now suppose for a contradiction that M' assigns some job b' to a^*. Note that M' does not assign b_0, and assigns b_{i+1} to a_i, which is different from a^*; so b' must be different from all of $b_0, \dots, b_r$. It now follows that the sequence

$$x' = (a_0, \dots, a_{r-1}, a^*; b_0, \dots, b_r, b')$$

is again a closed zigzag. As $b' \neq b_0$ and b' appears in a closed zigzag, we have $b' \in U'$. As M' assigns $b' \in U'$ to a^*, it follows that $a^* \in V'$, which contradicts our initial assumption about a^*. It follows that M' does not in fact assign any job to a^*, so that the sequence

$$x^* = (a_0, \dots, a_{r-1}, a^*; b_0, \dots, b_r)$$

is an open zigzag. We can thus flip this zigzag to obtain a new matching M^* that assigns b_0 as well as all of B'. In other words, M^* assigns all the jobs. □

12. Extensions and applications

Suppose we have a job allocation problem as before, but the number of jobs is smaller than the number of people, so that not everyone will get a job. Suppose that there are some enthusiastic people who really want a job, and some other people who do not care so much. Can we arrange a job allocation in which every enthusiastic person gets a job? Obviously this will be easier if every job has many enthusiastic candidates. We can give a precise result as follows.

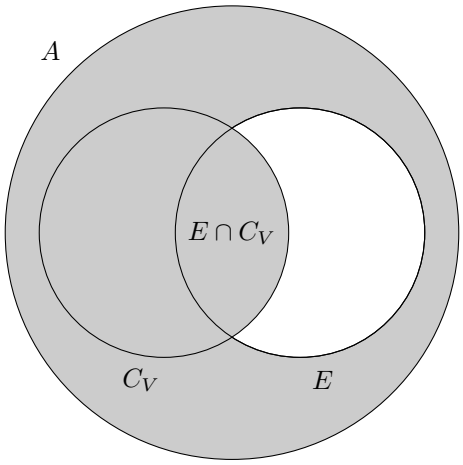
Theorem 12.1. *Video*

Consider a job allocation problem with a set A of people and a set B of jobs such that $|A| \geq |B|$. Suppose we are also given a subset $E \subseteq A$ of enthusiastic people. Suppose that each subset $U \subseteq B$ is plausible, so that $|C_U| \geq |U|$. Suppose also that the set $E \cap C_U$ of enthusiastic candidates has $|E \cap C_U| \geq |U| + |E| - |B|$. Then there is a matching M which allocates all the jobs, in such a way that all the enthusiastic people get a job.

Proof. Imagine a set B' of additional fake jobs that can only be done by unenthusiastic people (watching TV, lying on the beach and so on). The number of additional jobs should be $|B'| = |A| - |B|$, so that the set $B^* = B \cup B'$ has $|B^*| = |A|$. We declare that all the unenthusiastic people are qualified for all of the jobs in B', and that none of the enthusiastic people are qualified. This gives a new job allocation problem, with candidate sets C^*_U for $U \subseteq B^*$ say. If $U \subseteq B$ then the candidates for U are just the same as before. However, if $U \not\subseteq B$ then U contains at least one fake job, so all the unenthusiastic people are candidates, as well as the real candidates for all the real jobs in U. In symbols, we have

$$C^*_U = \begin{cases} C_U & \text{if } U \subseteq B \\ C_{U \cap B} \cup (A \setminus E) & \text{if } U \not\subseteq B. \end{cases}$$

We claim that this new allocation problem is still plausible, or in other words that $|C^*_U| \geq |U|$ for all $U \subseteq B^*$. For $U \subseteq B$ we have $|C^*_U| = |C_U| \geq |U|$ by our original assumptions. Suppose instead that $U \not\subseteq B$, so $U = V \cup V'$ for some $V \subseteq B$ and $V' \subseteq B'$ with $V' \neq \emptyset$. We then have $C^*_U = C_V \cup (A \setminus E)$. Consider the Venn diagram:



We see that C^*_U is the shaded region, and that this is the disjoint union of $E \cap C_V$ with $A \setminus E$. This gives $|C^*_U| = |E \cap C_V| + |A| - |E|$. We also have $|E \cap C_V| \geq |V| + |E| - |B|$ by our original assumptions, so

$$|C^*_U| \geq |V| + |E| - |B| + |A| - |E| = |V| + |A| - |B|.$$

On the other hand, we have $U = V \cup V'$ with $V' \subseteq B'$ and $|B'| = |A| - |B|$ so $|V'| \leq |A| - |B|$ so $|U| = |V| + |V'| \leq |V| + |A| - |B|$. Putting this together, we get $|C^*_U| \geq |U|$ as required. This proves that our new allocation problem is plausible, so Hall's Theorem tells us that there is a solution, say M^*. This allocates all the real jobs and all the fake jobs. As $|B^*| = |A|$, we see that everyone gets a job (either real or fake) for which they are qualified. The enthusiastic people are not qualified for the fake jobs, so they must

all be allocated a real job. Thus, if we just ignore the fake jobs, we have a solution to the original problem in which every enthusiastic person is employed. □

In the above results, we have repeatedly used the column/candidate sets:

$$C_b = \{ \text{ people who are qualified for job } b\}.$$

Recall that we also defined the row sets:

$$R_a = \{ \text{ jobs that person } a \text{ can do } \}.$$

Video (Proposition 12.2 and Corollary 12.3)

Proposition 12.2. *Suppose that there is a constant $d > 0$ such that*

(a) *For every job b we have $|C_b| \geq d$, so every job has at least d candidates.*
(b) *For every person a we have $|R_a| \leq d$, so no person is qualified for more than d jobs.*

Then the matching problem is plausible, and so has a solution by Hall's Theorem.

Proof. Consider a subset $U \subseteq B$; we must show that $|C_U| \geq |U|$. We will consider a set X, which we can describe in three different ways:

$$\begin{aligned} X &= \{(a,b) \in A \times U \mid a \text{ is qualified for } b\} \\ &= \{(a,b) \mid a \in A \text{ and } b \in R_a \cap U\} \\ &= \{(a,b) \mid b \in U \text{ and } a \in C_b\}. \end{aligned}$$

A moment's thought should convince you that these are three ways of saying the same thing. From the third description of X, we get

$$|X| = \sum_{b \in U} |C_b| \geq \sum_{b \in U} d = d|U|.$$

On the other hand, the second description gives

$$|X| = \sum_{a \in A} |R_a \cap U|.$$

If $a \notin C_U$ then a is not qualified for any of the jobs in U so $R_a \cap U = \emptyset$, so the corresponding term is zero. On the other hand, if $a \in C_U$ then we have $|R_a \cap U| \leq |R_a| \leq d$. We therefore have

$$|X| = \sum_{a \in C_U} |R_a \cap U| \leq \sum_{a \in C_U} d = d|C_U|.$$

By putting these inequalities together, we get $d|C_U| \geq d|U|$. As $d > 0$, we can divide by d to get $|C_U| \geq |U|$, as required. □

Corollary 12.3. *Suppose that there is a constant $d > 0$ such that*

(a) *For every job b we have $|C_b| = d$, so every job has precisely d candidates.*
(b) *For every person a we have $|R_a| \leq d$, so no person is qualified for more than d jobs.*

We will say that a person a is talented *if $|R_a| = d$. Then there is a job allocation in which every talented person has a job.*

Proof. Let $T \subseteq A$ be the set of talented people. Proposition 12.2 is still applicable, so we know that every subset $U \subseteq B$ is plausible, i.e. $|C_U| \geq |U|$. By Hall's Theorem, this implies that the job allocation problem is solvable. However, it does not guarantee that the allocation can be arranged so that everyone in T gets a job. For that, we need to check the additional criterion in Theorem 12.1, which is $|T \cap C_U| \geq |U| + |T| - |B|$. Now $|T \cap C_U| = |T| - |T \setminus C_U|$, so the required inequality is equivalent to $|T| - |T \setminus C_U| \geq |U| + |T| - |B|$ or $|T \setminus C_U| \leq |B| - |U| = |B \setminus U|$. For this, we let Y be the set of pairs (a, b) such that a is talented and qualified for b, but a is not qualified for any job in U. This ensures that b cannot be in U, so $b \in B \setminus U$. Once we have chosen b, we can try to choose a; this must in particular be an element of C_b, and $|C_b| \leq d$, so there are at most d choices for a. This analysis gives $|Y| \leq d|B \setminus U|$. Alternatively, we can start by choosing a. This can be any element of $T \setminus C_U$, and then b can be any of the jobs for which a is qualified. As a is

talented, there are precisely d choices for b. This analysis gives $|Y| = d\,|T \setminus C_U|$. Putting these together, we get $d\,|T \setminus C_U| \leq d\,|B \setminus U|$. As $d > 0$ this gives $|T \setminus C_U| \leq |B \setminus U|$ as required. □

We now give another theorem that is mathematically equivalent to Hall's Theorem, but thinly disguised.

Definition 12.4. Suppose we have a list of finite sets $A_1, \dots, A_r$. A *transversal* is a list of elements $a_1, \dots, a_r$ such that $a_i \in A_i$ for all i, and the elements a_i are all different. The list is *plausible* if for every sequence of indices $i_1 < i_2 < \cdots < i_k \leq r$, we have

$$|A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_k}| \geq k.$$

The phrase *distinct set of representatives* means the same as *transversal*.

Example 12.5. Consider the following list:

$$A_1 = \{1,3\} \qquad A_2 = \{2,3\} \qquad A_3 = \{1,3,4,5\} \qquad A_4 = \{2,4,6\} \qquad A_5 = \{1,5\} \qquad A_6 = \{1,2\}.$$

The following choices give a transversal:

$$a_1 = 1 \qquad a_2 = 3 \qquad a_3 = 4 \qquad a_4 = 6 \qquad a_5 = 5 \qquad a_6 = 2.$$

Proposition 12.6. *There exists a transversal iff the list is plausible.*

Proof. Put $A = A_1 \cup \cdots \cup A_r$ and $B = \{1, \dots, r\}$. Define

$$E = \{(a,i) \in A \times B \mid a \in A_i\} \subseteq A \times B;$$

this gives a matching problem with candidate sets $C_i = A_i$. Thus, if $U = \{i_1, \dots, i_k\}$ with $i_1 < \cdots < i_k$, we have $C_U = A_{i_1} \cup \cdots \cup A_{i_k}$. This makes it clear that the list $A_1, \dots, A_r$ is plausible (according to Definition 12.4) iff the matching problem E is plausible (according to Definition 11.2). Also, Hall's Theorem (together with Lemma 11.5) tells us that the matching problem is plausible iff there exists a full matching. If M is a full matching, then it assigns each $i \in B$ to some element $a_i \in A_i$, in such a way that $a_1, \dots, a_r$ are all different, so we have a transversal. This construction gives a bijection between full matchings and transversals, so in particular, a transversal exists iff a full matching exists. Putting all this together, we see that a transversal exists iff the list of sets is plausible. □

Example 12.7. Consider the following list:

$$A_1 = \{1,2,3\} \qquad A_2 = \{2,3\} \qquad A_3 = \{3,5,7\} \qquad A_4 = \{1,2\} \qquad A_5 = \{1,2,3\} \qquad A_6 = \{4,5,6\}.$$

We note that

$$|A_1 \cup A_2 \cup A_4 \cup A_5| = |\{1,2,3\}| = 3.$$

Here we have taken 4 of the sets, and their union only has size 3, which violates the plausibility condition. Thus, there is no transversal.

We now consider a version of the job allocation problem in which jobs can require a team of several workers.

Video (Definition 12.8 to Proposition 12.11)

Definition 12.8. A *team allocation problem* consists of a set A of people, a set B of jobs, a set $E \subseteq A \times B$ of pairs (a,b) where person a is qualified for job b, and numbers $m_b \geq 0$ for each $b \in B$. The problem is to choose a team $T_b \subseteq A$ for each job b, with $|T_b| = m_b$, such that each person in T_b is qualified to do job b, and the sets T_b are disjoint (so that no one has to do more than one job).

Definition 12.9. For any subset $U \subseteq B$, we define $m_U = \sum_{b \in U} m_b$ (so m_U is the total number of people required for all the jobs in U). We say that U is *plausible* if $|C_U| \geq m_U$. We say that the whole team allocation problem is plausible if every subset $U \subseteq B$ is plausible.

Remark 12.10. Suppose we have solved the team allocation problem, by choosing a team T_b for each job b. Then C_U contains $\bigcup_{b \in U} T_b$, and the sets T_b are disjoint and have size m_b, so $|\bigcup_{b \in U} T_b| = \sum_{b \in U} m_b = m_U$, so $|C_U| \geq m_U$. Thus, if the team allocation is solvable, then it is plausible. By the contrapositive, if the problem is not plausible, then there is no solution.

It is not hard to analyse the team allocation problem by converting it into an ordinary allocation problem of the type that we have considered already. We just imagine making a set of badges, labelled by elements of the set

$$B^* = \{(b, i) \mid b \in B,\ 1 \leq i \leq m_b\}.$$

For example, if we need 5 bakers, we make badges marked *Baker 1* up to *Baker 5*. We then put

$$E^* = \{(a, (b, i)) \in A \times B^* \mid (a, b) \in E\},$$

corresponding to the idea that a is qualified to wear badge (b, i) iff a is qualified to do job b. Now the problem of choosing teams is equivalent to the problem of assigning badges. Using this, we can prove the following result.

Proposition 12.11. *If a team allocation problem is plausible, then it is solvable.*

Proof. Consider a plausible team allocation problem (A, B, E, m) as before, and the corresponding badge allocation problem (A, B^*, E^*). As we have explained, it will be enough to prove that the badge allocation problem is solvable. By Hall's theorem, it will be enough to show that the badge allocation problem is plausible. Consider a subset $U \subseteq B^*$; we must show that $|C^*_U| \geq |U|$. Put

$$V = \{b \mid (b, i) \in U \text{ for some } i\}.$$

In other words, U is a set of badges, and V is the set of job titles that appear on at least one of those badges. Now put

$$V^* = \{(b, j) \mid b \in V,\ 1 \leq j \leq m_b\}.$$

In other words, V^* is the set of badges that share a job title with one of the badges in U, so $U \subseteq V^*$, so $|U| \leq |V^*|$. It is clear that $|V^*| = \sum_{b \in V} m_b = m_V$, so the relation $|U| \leq |V^*|$ becomes $|U| \leq m_V$. The team allocation problem is assumed to be plausible, which implies that $m_V \leq |C_V|$. On the other hand, a person is qualified for one of the badges in U iff they are qualified for one of the jobs in V, so $C_V = C^*_U$. Putting this together, we get $|U| \leq |C^*_U|$ as required. □

13. Tournaments

Definition 13.1. Consider a finite set P. A *tournament* on P is a subset $T \subseteq P \times P$ such that

(a) No pair of the form (a, a) lies in T.
(b) For any two players $a \neq b$, either $(a, b) \in T$ or $(b, a) \in T$ but not both.

Remark 13.2. We interpret this as follows. The set P could be a set of players, who play some game against each other in pairs, with each pair playing precisely once. Then T is the set of pairs (a, b) such that a beats b.

Example 13.3. Interactive demo

Here are three different ways to represent the result of a tournament with players $A, \ldots, E$:

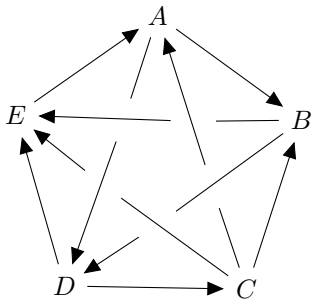
	A	B	C	D	E
A		W	L	W	L
B	L		L	W	W
C	W	W		L	W
D	L	L	W		W
E	W	L	L	L	

AB AC **AD**

AE BC **BD**

BE CD **CE**

DE



The left hand table shows who wins each game. For example, in the (A, D) position (i.e. the row marked A and the column marked D) we see a W, indicating that A wins against D. Correspondingly, we have an L in the (D, A) position, indicating that D loses against A. In the middle, we list all ten possible pairs of players, with the winner marked bold and in red. On the right, we have an arrow between each pair of players, pointing from the winner to the loser.

Remark 13.4. Interactive demo

Here is yet another way of representing the same information. We imagine making a medal for each pair of players, and giving it to the winner of the corresponding match. If we know each player's collection of medals, then we know all the results of the tournament. The tournament in Example 13.3 gives the following medal collections.

A	B	C	D	E
A v B	B v D	A v C	C v D	A v E
A v D	B v E	B v C	D v E	
		C v E		

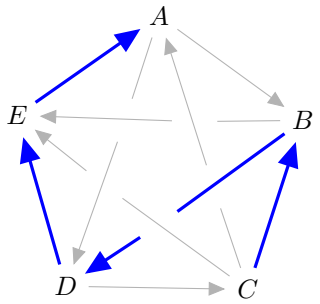
Example 13.5. We say that a tournament T is *consistent* if the players can be listed as $a_1, \ldots, a_n$ in such a way that a_i beats a_j whenever $i < j$. (In particular, a_1 beats everyone else and so is the "number one player" in the usual sense. Similarly, a_2 is the "number two player" and so on.) This is the kind of tournament that we expect when the players have consistently different levels of skill, and the better player always wins, with no effect of randomness or anything else. In a realistic tournament, things will usually be more complicated than this.

Interactive demo

Definition 13.6. A *winning line* for a tournament $T \subseteq A \times A$ is a list $a_1, \ldots, a_n$ containing each player exactly once, such that a_i beats a_{i+1} for $i = 1, \ldots, n-1$.

Remark 13.7. Although a_1 beats a_2 and a_2 beats a_3, we are not assuming here that a_1 beats a_3. Thus, having a winning line is much weaker than having a consistent ranking as in Example 13.5.

Example 13.8. In Example 13.3, the sequence C, B, D, E, A is a winning line. This is most easily seen using the arrow graph:



Note that the effect mentioned in Remark 13.7 appears here: C beats B and B beats D, but C does not beat D.

Proposition 13.9. *Every tournament has a winning line.*

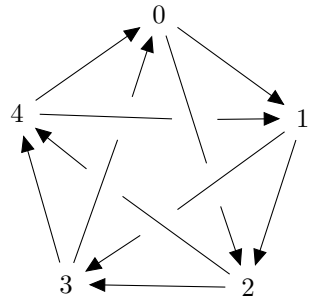
Proof. This is clear if there are 0, 1 or 2 players. Suppose instead that the set of players is P, with $|P| = n > 2$, and argue by induction on n. Choose a player a^*, and put $P' = P \setminus \{a^*\}$. We can apply the induction hypothesis to P', and thus list the elements of P' as $a_1, \dots, a_{n-1}$, in such a way that a_i beats a_{i+1} for $i = 1, \dots, n-2$. If a^* does not beat any of these players, then $a_1, \dots, a_{n-1}, a^*$ is a winning line for T. Suppose instead that a^* does beat some of the players a_i. Let a_k be the first one that a^* beats. If $k = 1$, then $a^*, a_1, \dots, a_{n-1}$ is a winning line for T. Suppose instead that $k > 1$ (so a_{k-1} is meaningful). As a_k is the first player that a^* beats, we see that a_{k-1} must beat a^*. It follows that the sequence $a_1, \dots, a_{k-1}, a^*, a_k, \dots, a_{n-1}$ is a winning line. □

Definition 13.10. The *score* of a player in a tournament is the number of games that they win. (More formally, the score of $a \in A$ with respect to a tournament $T \subseteq A \times A$ is $|\{b \in A \mid (a, b) \in T\}|$.) The *score sequence* of T is the list of scores of all players, written in decreasing order. We write scores(T) for the score sequence of T.

Example 13.11. In Example 13.11, the scores of A, B, C, D and E are $2, 2, 3, 2$ and 1. The score sequence is therefore $3, 2, 2, 2, 1$. Consider instead a consistent tournament (as in Example 13.5) with n players. Then player 1 beats players $2, \dots, n$ and so scores $n-1$, and player 2 beats players $3, \dots, n$ and so scores $n-2$, and so on. The last player (number n) beats no one and so scores zero. The score sequence is therefore $n-1, n-2, \dots, 1, 0$.

Example 13.12. Interactive demo

Given $m > 1$, we can construct a tournament with $2m+1$ players as follows. The set of players is $\mathbb{Z}/(2m+1)$, so all expressions with player numbers must be interpreted modulo $2m+1$. For each $i \in \mathbb{Z}/(2m+1)$, player i beats players $i+1, \dots, i+m$, and is beaten by players $i-1, \dots, i-m$. We call this the *odd modular tournament* of size $2m+1$. For example, when $m = 2$ (so $2m+1 = 5$) we have the following pattern:



Note that in this kind of tournament, each player has a score of m. This is in some sense opposite to the case of a consistent tournament: the scores give no reason to think that any player is better than any other player.

Interactive demo

We next want to investigate some properties of score sequences.

Definition 13.13. Consider a sequence $s = (s_1, \dots, s_n)$ of nonnegative integers with $s_1 \geq \cdots \geq s_n$. We define

$$\begin{aligned}
\text{length}(s) &= n \\
\text{first}_k(s) &= s_1 + s_2 + \cdots + s_k = \sum_{1 \leq i \leq k} s_i \\
\text{last}_k(s) &= s_{n-k+1} + s_{n-k+2} + \cdots + s_n = \sum_{n-k < i \leq n} s_i \\
\text{total}(s) &= \text{first}_n(s) = \text{last}_n(s) = \sum_{1 \leq i \leq n} s_i.
\end{aligned}$$

More generally, given any subset $U \subseteq \{1, \dots, n\}$, we put $s_U = \sum_{i \in U} s_i$.

Definition 13.14. We say that a sequence s is *realisable* if there is a tournament T with $\text{scores}(T) = s$. Any such tournament is a *realisation* of s.

Our main aim in this section is to prove a theorem of Landau, which will tell us exactly which sequences are realisable. The following simple example illustrates the key ingredients:

Interactive demo

The most basic property is as follows:

Lemma 13.15. *If s is a realisable sequence of length n, then* $\text{total}(s) = \binom{n}{2}$.

Proof. As s is realisable, we can find a tournament with score sequence s. Put $m = \sum_i s_i$. This is the sum of the scores of all players. Each game contributes a score of one to one or the other player, so the sum of all scores is just equal to the number of games. We have one game between each pair of players, so the number of games is the number of pairs, which is $\binom{n}{2}$. □

Another basic observation is as follows:

Lemma 13.16. *If a sequence is realisable, then it cannot contain two zeros.*

Proof. Suppose we have a tournament T, and two distinct players a and b. Then a and b play each other, and one of them must win, thereby gaining a score of one; so they cannot both have a score of zero. Thus, the score sequence $\text{scores}(T)$ contains at most one zero. Thus, if we have a sequence with two or more zeros, it cannot be realisable. □

Video (Lemma 13.17 to Lemma 13.20)

To formulate Landau's Theorem, it will be convenient to have the following result.

Lemma 13.17. *For $0 \leq k \leq n$ we have* $\binom{n}{2} = \binom{k}{2} + k(n-k) + \binom{n-k}{2}$.

Algebraic proof.

$$\begin{aligned}
\binom{k}{2} + k(n-k) + \binom{n-k}{2} &= \tfrac{1}{2}k(k-1) + k(n-k) + \tfrac{1}{2}(n-k)(n-k-1) \\
&= \tfrac{1}{2}k^2 - \tfrac{1}{2}k + nk - k^2 + \tfrac{1}{2}n^2 - nk + \tfrac{1}{2}k^2 - \tfrac{1}{2}n + \tfrac{1}{2}k \\
&= \tfrac{1}{2}n^2 - \tfrac{1}{2}n = \tfrac{1}{2}n(n-1) = \binom{n}{2}.
\end{aligned}$$

□

Bijective proof. Put $N = \{1, \dots, n\}$, and let $K \subseteq N$ be any subset of size k, so $N \setminus K$ has size $n - k$. Then

- $\binom{n}{2}$ is the number of subsets $T \subseteq N$ with $|T| = 2$.
- $\binom{k}{2}$ is the number of such subsets T with both elements of T in K.
- $k(n-k)$ is the number of such subsets T with one element in K and the other in $N \setminus K$.
- $\binom{n-k}{2}$ is the number of such subsets T with both elements of T in $N \setminus K$.

From this it is clear that $\binom{n}{2} = \binom{k}{2} + k(n-k) + \binom{n-k}{2}$. □

Proposition 13.18. *Let $s = (s_1, \dots, s_n)$ be a sequence of nonnegative integers in descending order, and suppose that* $\text{total}(s) = \binom{n}{2}$. *Put $N = \{1, \dots, n\}$. Then the following conditions are equivalent:*

(a) *For all $U \subseteq N$ with $|U| = k$ we have $s_U \geq \binom{k}{2}$.*
(b) *For all $U \subseteq N$ with $|U| = k$ we have $s_U \leq k(n-k) + \binom{k}{2}$.*
(c) *For all k we have* $\text{last}_k(s) \geq \binom{k}{2}$.
(d) *For all k we have* $\text{first}_k(s) \leq k(n-k) + \binom{k}{2}$.

Proof. First suppose that (a) holds; we will prove that (b) also holds. Indeed, if $U \subseteq N$ with $|U| = k$, then $|U^c| = n - k$, so (a) tells us that $s_{U^c} \geq \binom{n-k}{2}$, so

$$\binom{n}{2} - s_{U^c} \leq \binom{n}{2} - \binom{n-k}{2}.$$

On the other hand, we have $s_U + s_{U^c} = \text{total}(s) = \binom{n}{2}$, so $\binom{n}{2} - s_{U^c} = s_U$. Moreover, Lemma 13.17 shows that $\binom{n}{2} - \binom{n-k}{2} = k(n-k) + \binom{k}{2}$. Thus, the above inequality can be rewritten as $s_U \leq k(n-k) + \binom{k}{2}$, as required. This completes the proof that (a) implies (b), and the whole argument can be reversed in a straightforward way to prove that (b) implies (a), so (a) and (b) are equivalent.

Now consider all the numbers s_U for subsets U with $|U| = k$. Note that $\text{first}_k(s)$ is one of these numbers (for $U = \{1, \dots, k\}$) and $\text{last}_k(s)$ is also one of these numbers (for $U = \{n-k+1, \dots, n\}$). In fact, since the numbers s_i are in decreasing order, it is clear that $\text{first}_k(s)$ is the largest of these numbers s_U, and $\text{last}_k(s)$ is the smallest. Thus, if $\text{last}_k(s) \leq \binom{k}{2}$, then $s_U \leq \binom{k}{2}$ for all U with $|U| = k$. Similarly, if $\text{first}_k(s) \leq k(n-k) + \binom{k}{2}$, then $s_U \leq k(n-k) + \binom{k}{2}$ for all U with $|U| = k$. This shows that (a) is equivalent to (c) and (b) is equivalent to (d). □

Definition 13.19. We say that a sequence s is *plausible* if it has the equivalent conditions described in Proposition 13.18.

The easy half of Landau's result is as follows.

Lemma 13.20. *Every realisable sequence is plausible.*

Proof. Let s be a realisable sequence of length n. As s is realisable, we can find an n-player tournament T such that s_i is the score of player i. Let U be a set of k players, so that s_U is the total of their scores. We then have $s_U = p + q$, where p is the total of scores that people in U earn by playing each other, and q is the total of scores that they earn by playing other people. The members of U play $\binom{k}{2}$ games against each other, and each of these earns a score of 1 for one or the other of the players, so we get $p = \binom{k}{2}$. On the other hand, the members of U play $k(n-k)$ games against people who are not in U. They might lose all of these games, or they might win all of them, or something in between; so $0 \leq q \leq k(n-k)$. From this it follows that

$$\binom{k}{2} \leq s_U \leq k(n-k) + \binom{k}{2},$$

so conditions (a) and (b) in Proposition 13.18 are satisfied. We have shown that conditions (a) to (d) are all equivalent, so in fact they are all satisfied. □

Example 13.21. For a consistent tournament of size n we have $s_i = n - i$ for all i, and it is not hard to deduce that $\text{last}_k(s) = \binom{k}{2}$ and $\text{first}_k(s) = k(n-k) + \binom{k}{2}$, so the score sequence is plausible, as predicted by Lemma 13.20. For an odd modular tournament of size $2m+1$, we have $s_i = m$ for all i. Thus, for any set U with $|U| = k$, we have $s_U = mk$. On the other hand, we have $2m + 1 = n \geq k$ so $m \geq (k-1)/2$ so

$mk \geq k(k-1)/2 = \binom{k}{2}$. This means that $s_U \geq \binom{k}{2}$, so again the score sequence is plausible, as predicted by Lemma 13.20. Finally, the score sequence for Example 13.3 was $s = (3, 2, 2, 2, 1)$. This has

$$\begin{aligned}
\text{last}_1(s) &= 1 = 1 &&\geq 0 = \tbinom{1}{2} \\
\text{last}_2(s) &= 2 + 1 = 3 &&\geq 1 = \tbinom{2}{2} \\
\text{last}_3(s) &= 2 + 2 + 1 = 5 &&\geq 3 = \tbinom{3}{2} \\
\text{last}_4(s) &= 2 + 2 + 2 + 1 = 7 &&\geq 6 = \tbinom{4}{2} \\
\text{last}_5(s) &= 3 + 2 + 2 + 2 + 1 = 10 &&= 10 = \tbinom{5}{2}.
\end{aligned}$$

Once again, we see that the score sequence is plausible.

The hard part is the converse:

Theorem 13.22 (Landau's Tournament Theorem)**.** *Every plausible sequence is realisable.*

Proof. Interactive demo

Video

Let $s = (s_1, \dots, s_n)$ be a plausible sequence, so $s_1 \geq \cdots \geq s_n$ and $s_1 + \cdots + s_k \geq \binom{k}{2}$ for all k, and $s_1 + \cdots + s_n = \binom{n}{2}$. We will set up a kind of team allocation problem, to which we can apply Proposition 12.11. Consider a pair of players, say $\{\text{Alice}, \text{Bob}\}$. As in Remark 13.4, we can imagine making a medal which says "Alice vs Bob" (but does not say who won). Alice or Bob could receive this medal, but no one else could. We can repeat this for every pair of players, giving a set of medals $x_1, \dots, x_N$, where $N = \binom{n}{2}$. For the medal ceremony, we recruit children $c_1, \dots, c_N$, and give medal x_j to child c_j. To specify the tournament, we just need to specify which children present medals to each player (because this determines who gets each medal, and thus who wins each game). Player i must win s_i games, so they must receive s_i medals, so they need a team of s_i children to present medals to them. These children must be qualified, in the sense that they must carry a medal with player i's name on it. If we can show that this team allocation problem is solvable, this will give a tournament with the required scores. For this, we use the numerical condition in Proposition 12.11. Consider a set U of players, and put $k = |U|$. Let C_U be the set of candidates for the corresponding jobs. Equivalently, C_U is the set of children who carry a medal that could be presented to one of the players in U. There are $\binom{k}{2}$ medals where both names are in U, and $k(n-k)$ medals where one name is in U and the other is not; thus, we have $|C_U| = \binom{k}{2} + k(n-k)$. The total number of children needed to present medals to players in U is $\sum_{i \in U} s_i = s_U$. Thus, the numerical condition in Proposition 12.11 is that $\binom{n}{k} + k(n-k) \geq s_U$ for all U. This is is precisely the same as the plausibility condition from Proposition 13.18, and we are assuming that that condition is satisfied. Thus, Proposition 12.11 tells us that the team allocation problem is solvable, as required. □

Example 13.23. Consider the sequence $s = (5, 3, 2, 2, 2, 1)$, of length 6. We have

$$\begin{aligned}
1 &= 1 &&\geq 0 = \tbinom{1}{2} \\
2 + 1 &= 3 &&\geq 1 = \tbinom{2}{2} \\
2 + 2 + 1 &= 5 &&\geq 3 = \tbinom{3}{2} \\
2 + 2 + 2 + 1 &= 7 &&\geq 6 = \tbinom{4}{2} \\
3 + 2 + 2 + 2 + 1 &= 10 &&\geq 10 = \tbinom{5}{2} \\
5 + 3 + 2 + 2 + 2 + 1 &= 15 &&= 15 = \tbinom{6}{2}
\end{aligned}$$

This shows that the sequence is plausible, so Landau's theorem tells us that it is possible to find a corresponding tournament. In fact, the following tournament works:

	1	2	3	4	5	6
1		W	W	W	W	W
2	L		L	W	W	W
3	L	W		L	L	W
4	L	L	W		L	W
5	L	L	W	W		L
6	L	L	L	L	W	

Example 13.24. For the sequence $s = (4, 4, 4, 2, 1, 1, 1)$ we have $\text{last}_4(s) = 5$, which is strictly less than $\binom{4}{2} = 6$, so the sequence is not plausible, so it cannot be the score sequence for a tournament.

We next discuss two ways of combining tournaments.

Video (Definition 13.25 to Example 13.31)

Definition 13.25. Let T be a tournament with n players and scores $t_1 \geq \cdots \geq t_n$. Let U be another tournament, with m players and scores $u_1 \geq \cdots \geq u_m$. We then let $T : U$ denote the combined tournament with $n + m$ players given by the following rules:

(a) For a game between two players in T, the result is the same as it was in T.
(b) For a game between two players in U, the result is the same as it was in U.
(c) For a game with one player from T and one player from U, the player from T always wins.

Note that the players from U have the same score in the combined tournament as they did in U. However, the players from T have their original score plus an extra m points for beating all the players from U. In particular, the players from T all have at least m points, but the players from U all have less than m points. Thus, the score sequence for the combined tournament is

$$t_1 + m \geq t_2 + m \geq \cdots \geq t_n + m \geq u_1 \geq \cdots \geq u_m.$$

Remark 13.26. You can think of T as the Sheffield primary school football league, and U as the premier league. One year they decide to have a joint tournament for charity, and of course the professionals always let the children win; the resulting tournament is $T : U$.

Example 13.27. Suppose we want to construct a tournament with score sequence $(7, 7, 7, 7, 7, 2, 2, 2, 2, 2)$. We start with an odd modular tournament T with players $(1, 2, 3, 4, 5)$ and scores $(2, 2, 2, 2, 2)$. We then let U be another copy of T, with players $(6, 7, 8, 9, 10)$ and scores $(2, 2, 2, 2, 2)$ again. The combined tournament $T : U$ then has scores $(7, 7, 7, 7, 7, 2, 2, 2, 2, 2)$ as required.

Definition 13.28. Again let T and U be two tournaments, with n and m players respectively. We can combine them in a different way to produce a tournament $T * U$ as follows. The players are pairs (i, j), where i is a player from T and j is a player from U, so there are nm players altogether. Consider two players (i, j) and (i', j') with $(i, j) \neq (i', j')$, so either $i \neq i'$ or $(i = i'$ and $j \neq j')$.

(a) If $i \neq i'$ then the result of (i, j) playing (i', j') in $T * U$ is the same as the result of i playing i' in T.
(b) If $i = i'$ and $j \neq j'$ then the result of (i, j) playing (i', j') in $T * U$ is the same as the result of j playing j' in U.

Remark 13.29. You can think of T as a popularity contest between various potential birthday presents, and U as a popularity contest between various kinds of fancy bags. Then $T * U$ is the corresponding popularity contest between presents-in-bags, where we assume that the recipient mostly cares about the present, and only compares the bags if the presents are the same.

Lemma 13.30. *Suppose that the scores in T are $t_1 \geq \cdots \geq t_n$ and the scores in U are $u_1 \geq \cdots \geq u_m$. Then the score for player (i, j) in $T * U$ is $m\,t_i + u_j$.*

Proof. Let P_i be the set of players beaten by i in T, and let Q_j be the set of players beaten by j in U. We then have $|P_i| = t_i$ and $|Q_j| = u_j$. The players beaten by (i, j) are then the players (i', j') with either $i' \in P_i$ and j' arbitrary, or $i' = i$ and $j' \in Q_j$. There are $m\,t_i$ possibilities of the first type, and u_j possibilities of the second type. This shows that the score for (i, j) is $m\,t_i + u_j$ as claimed. □

Example 13.31. Let T be an odd modular tournament of size $n = 3$, so the scores are $(1, 1, 1)$, or in other words $t_i = 1$. Let U be a consistent tournament of size $m = 5$, so the scores are $(4, 3, 2, 1, 0)$, or in other words $u_j = 5 - j$. In $T * U$, the score for (i, j) is $m\,t_i + u_j = 5 + 5 - j = 10 - j$. Each number appears three times, because there are three possible choices for i, so the full score sequence is

$$(9, 9, 9, 8, 8, 8, 7, 7, 7, 6, 6, 6, 5, 5, 5).$$

We could instead consider the tournament $U * T$. Here the score for (j, i) is $n\,u_j + t_i = 3(5 - j) + 1 = 16 - 3j$. The full score sequence is

$$(13, 13, 13, 10, 10, 10, 7, 7, 7, 4, 4, 4, 1, 1, 1).$$

14. Latin squares

Definition 14.1. Given nonempty finite sets P, Q and N, a *Latin rectangle* L is a system of elements $L_{ij} \in N$ for $i \in P$ and $j \in Q$ such that

(a) For each $i \in P$, all the elements in the row L_{i*} are distinct. In more detail, if $i \in P$ and $j, j' \in Q$ with $j \neq j'$, then we must have $L_{ij} \neq L_{ij'}$.
(b) For each $j \in Q$, all the elements in the column L_{*j} are distinct. In more detail, if $j \in Q$ and $i, i' \in P$ with $i \neq i'$, then we must have $L_{ij} \neq L_{i'j}$.

We will usually write $p = |P|$ and $q = |Q|$ and $n = |N|$. Often (but not always) we will have $P = \{1, \dots, p\}$ or $P = \{0, \dots, p - 1\}$ and similarly for Q and N.

Remark 14.2. In each column we have p entries from N which must all be different, and in each column we have q entries from N which must all be different. This can only work if $p, q \leq n$. Thus, if we fix N with $|N| = n$, then the maximum possible size of a Latin rectangle is $n \times n$.

Definition 14.3. Interactive demo

A *Latin square* of size n is a Latin rectangle with $|P| = |Q| = |N| = n$.

Note that in a Latin square each row contains n different entries taken from N, but $|N| = n$, so each row must contain each element of N precisely once. Similarly, each column must contain each element of N precisely once.

Example 14.4. The matrix

$$\begin{bmatrix} 1 & 4 & 3 \\ 5 & 2 & 1 \end{bmatrix}$$

gives a Latin rectangle with $P = \{1, 2\}$ and $Q = \{1, 2, 3\}$ and $N = \{1, 2, 3, 4, 5\}$ so $p = 2$ and $q = 3$ and $n = 5$.

Example 14.5. The matrix

$$\begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 3 & 4 & 1 \\ 4 & 1 & 2 & 3 \\ 3 & 2 & 1 & 4 \end{bmatrix}$$

gives a Latin square with $P = Q = N = \{1, 2, 3, 4\}$ and $p = q = n = 4$.

Example 14.6. Let G be any finite group, with $|G| = n$. Take $P = Q = N = G$ and $L_{g,h} = g * h$. I claim that this is a Latin square. Indeed, if $L_{g,h} = L_{g,h'}$ then $g * h = g * h'$ and we can multiply on the left by g^{-1} to see that $h = h'$. By the contrapositive, if $h \neq h'$ then $L_{g,h} \neq L_{g,h'}$. By a similar argument, if $g \neq g'$ then $L_{g,h} \neq L_{g',h}$, as required.

Interactive demo

Example 14.7. As a special case of the above, we can consider the group $\mathbb{Z}/n = \{0, \dots, n-1\}$, with addition mod n as the group operation. This gives a Latin square with $P = Q = N = \{0, \dots, n-1\}$ and $L_{ij} = i + j \pmod{n}$. For example, when $n = 5$ we get

$$L = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 0 \\ 2 & 3 & 4 & 0 & 1 \\ 3 & 4 & 0 & 1 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix}.$$

This example shows that for any n, there is at least one $n \times n$ Latin square.

Interactive demo

Theorem 14.8. *Video*

Let L be a Latin rectangle with $p < q = n$ (so L has the maximum possible width, but not the maximum possible height). Then L can be extended by adding extra rows to make an $n \times n$ Latin square.

Remark 14.9. To prove this theorem and the next theorem, we will apply Hall's Theorem and related results to a certain matching problem. This can be interpreted as a job allocation problem, in the following way. Suppose that N is a set of n students on a work experience scheme, and that Q is a set of n jobs, and that $P = \{1, \dots, n\}$. Each student is supposed to do each of the jobs for one day. On day i, job j is done by student L_{ij}. No student can do more than one job at the same time, so for fixed i, the entries L_{ij} must all be different. No student does the same job more than once, so for fixed j, the entries L_{ij} must all be different. Thus, the matrix L must be a Latin square.

In the context of Theorem 14.8, we have constructed the rota for days 1 to p, and our task is to complete the rota for days $p+1$ to n.

The proof of Theorems 14.8 and 14.22 will also depend on some extra definitions which we now explain.

Definition 14.10. Let L be a Latin rectangle with parameters p, q, n. For $k \in N$ we let $m_L(k)$ denote the number of occurrences of k in L, and we call this the *multiplicity* of k. We also put $e_L(k) = m_L(k) + n - p - q$ and call this the *excess* of k.

Example 14.11. For $L = \begin{bmatrix} 1 & 4 & 3 & 5 \\ 5 & 3 & 1 & 4 \end{bmatrix}$ we have $(p, q, n) = (2, 4, 5)$ so $n - p - q = -1$ and

$$\begin{array}{lllll} m_L(1) = 2 & m_L(2) = 0 & m_L(3) = 2 & m_L(4) = 2 & m_L(5) = 2 \\ e_L(1) = 1 & e_L(2) = -1 & e_L(3) = 1 & e_L(4) = 1 & e_L(5) = 1. \end{array}$$

Remark 14.12. The occurrences of k in L must appear in different rows, so $m_L(k)$ can also be described as the number of rows that contain k. Similarly, the occurrences of k in L must appear in different columns, so $m_L(k)$ can also be described as the number of columns that contain k.

Lemma 14.13. *Suppose that $q = n$, so that L has the maximum possible width. Then we have $m_L(k) = p$ and $e_L(k) = 0$ for all k. Similarly, if $p = n$ (so that L has the maximum possible height) then $m_L(k) = q$ and $e_L(k) = 0$ for all k.*

Proof. Suppose that $q = n$. Then each row has n different elements but $|N| = n$ so each element $k \in N$ must occur precisely once in each of the p rows. From this we see that $m_L(k) = p$ and so $e_L(k) = p + n - p - q$. As $q = n$ this simplifies to $e_L(k) = 0$. The case where $p = n$ is essentially the same. □

Proof of Theorem 14.8. Let L be a $p \times n$ Latin rectangle. It will be enough to show that we can add one more row to get a $(p+1) \times n$ Latin rectangle, because we can then repeat the process if necessary. We will interpret the problem as in Remark 14.9.

For $j \in Q$, let $C_j \subseteq N$ be the set of students who are allowed to do job j on day $p+1$. These are just the students who have not already done job j on any of days $1, \dots, p$, or in other words

$$C_j = N \setminus \{L_{1j}, \dots, L_{pj}\}.$$

We are assuming that L is a Latin rectangle, so we have followed the rules on days 1 to p, so students $L_{1j}, \dots, L_{pj}$ must all be different, so $|C_j| = n - p$. To make the new row, we just need to solve the usual job allocation problem with candidate sets C_j. For each $k \in N$ put

$$R_k = \{j \mid k \in C_j\},$$

which is the set of jobs for which student k is qualified (in the sense that they have not already done that job). We will use Proposition 12.2: if there is a constant d such that $|C_j| = d$ for all j and $|R_k| \leq d$ for all k, then the allocation problem is solvable. We will take $d = n - p$; we have already seen that $|C_j| = d$ for all j. On the other hand, R_k is just the set of columns where we are allowed to put k in the new row, or in other words, the set of columns that do not already contain k. The number of columns that contain k is $m_L(k)$, which is p as we explained in Lemma 14.13. Thus, the number of columns that do not contain k is $|R_k| = n - p = d$ as required. Thus, Proposition 12.2 is applicable, so we can solve the allocation problem, and the solution gives us an extra row. □

Example 14.14. Consider the following Latin rectangle, with $p = 2$ and $q = n = 5$:

$$L = \begin{bmatrix} 1 & 2 & 4 & 5 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{bmatrix}.$$

Recall that C_j is the set of possibilities for position j in the next row. For example, in column 2 we already have a 2 and a 1, so these are not allowed, so $C_2 = N \setminus \{2, 1\} = \{3, 4, 5\}$. We can display all the sets C_j as follows:

$$\begin{bmatrix} 1 & 2 & 4 & 5 & 3 \\ 5 & 1 & 2 & 3 & 4 \\ 234 & 345 & 135 & 124 & 125 \end{bmatrix}$$

(We have used abbreviated notation, e.g. 234 for $\{2, 3, 4\}$.) To make the new row, we must choose one element from the possibilities in each column, making sure that we never choose the same element twice. Corollary 55 tells us that this is possible, but does not tell us exactly how to do it. However, in this case it is not difficult: in each column we can just take the first choice that has not already been used. This gives $2, 3, 1, 4, 5$ as the new row. We can write in this new row and display the possibilities for row 4 as follows:

$$\begin{bmatrix} 1 & 2 & 4 & 5 & 3 \\ 5 & 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 & 5 \\ 34 & 45 & 35 & 12 & 12 \end{bmatrix}$$

Again, in each column we can take the first choice that has not already been used. This gives row 4 and leaves only one possibility for row 5. We end up with the following Latin square:

$$\begin{bmatrix} 1 & 2 & 4 & 5 & 3 \\ 5 & 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 3 & 2 & 1 \end{bmatrix}$$

Corollary 14.15. *Let L be a Latin rectangle with $q < p = n$ (so L has the maximum possible height, but not the maximum possible width). Then L can be extended by adding extra columns to make an $n \times n$ Latin square.*

Proof. Note that the transpose L^T is a Latin square of maximum possible width, so we can use Theorem 14.8 to extend it to a Latin square, then take the transpose again at the end. This just amounts to doing the same steps as before, but with the roles of rows and columns exchanged. □

Now consider a Latin rectangle where both p and q are strictly less than n, so neither Theorem 14.8 nor Corollary 14.15 is applicable. Can we still extend it to give an $n \times n$ Latin square? It is not hard to find examples where this is not possible.

Example 14.16. Take $P = Q = \{1, 2\}$ and $N = \{1, 2, 3\}$ and $L = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$. We could try to extend this to a 3×3 Latin square as follows:

$$\left[\begin{array}{cc|c} 2 & 3 & a \\ 3 & 2 & b \\ \hline c & d & e \end{array}\right].$$

To avoid a clash in row 1, we must take $a = 1$. To avoid a clash in row 2, we must also take $b = 1$. However, this creates an unavoidable clash in column 3. Thus, it is impossible to extend L.

Example 14.17. Take $P = Q = \{1, 2, 3, 4\}$ and $N = \{1, \dots, 6\}$ and

$$L = \begin{bmatrix} 6 & 1 & 2 & 3 \\ 5 & 6 & 3 & 1 \\ 1 & 3 & 6 & 2 \\ 3 & 2 & 1 & 4 \end{bmatrix}.$$

It turns out that it is not possible to extend this to a 6×6 Latin square. It is a good exercise to prove this directly. However, we will deduce it from a general theorem instead. We can list the multiplicity and excess of the elements of N as follows:

k	1	2	3	4	5	6
$m_L(k)$	4	3	4	1	1	3
$e_L(k)$	2	1	2	-1	-1	1

It turns out that the key point is that some excesses are negative.

Video (Definition 14.18 to Theorem 14.22)

Definition 14.18. Let L be a Latin rectangle. We say that an element $k \in N$ is *plausible* if $e_L(k) \geq 0$. More precisely, we say that k is *barely plausible* if $e_L(k) = 0$, and *very plausible* if $e_L(k) > 0$.

Proposition 14.19. *If L can be extended to an $n \times n$ Latin square, then every element $k \in N$ is plausible for L.*

Proof. Choose a Latin square extending L. This will have the form

$$\left[\begin{array}{c|c} L & L' \\ \hline L'' & L''' \end{array}\right],$$

where L', L'' and L''' have shape $p \times (n-q)$, $(n-p) \times q$ and $(n-p) \times (n-q)$. Let L^* be the top part, consisting of L and L'. This is a $p \times n$ Latin rectangle, so Lemma 14.13 tells us that

$$m_L(k) + m_{L'}(k) = m_{L^*}(k) = p,$$

so $m_L(k) = p - m_{L'}(k)$. On the other hand, L' has $n - q$ columns, and there is at most one occurrence of k per column, so $m_{L'}(k) \leq n - q$, so $p - m_{L'}(k) \geq p + q - n$. Putting this together, we get $m_L(k) \geq p + q - n$ and so $e_L(k) = m_L(k) + n - p - q \geq 0$. □

In Example 14.17, we see that 4 and 5 have negative excess so they are not plausible, so there cannot be any extension to a 6×6 Latin square. We now discuss another example.

Example 14.20. Consider the following Latin rectangle with $p = 4$ and $q = 5$ and $n = 7$:

$$L = \begin{bmatrix} 5 & 6 & 1 & 3 & 2 \\ 6 & 5 & 2 & 4 & 7 \\ 1 & 4 & 3 & 5 & 6 \\ 4 & 7 & 5 & 6 & 1 \end{bmatrix}.$$

The multiplicities and excesses are as follows:

k	1	2	3	4	5	6	7
$m_L(k)$	3	2	2	3	4	4	2
$e_L(k)$	1	0	0	1	2	2	0

We have $e_L(k) \geq 0$ so all elements are plausible, so we might guess that L can be extended to a 7×7 Latin square. However, Proposition 14.19 does not give us any guarantees about this. If we had found that $e_L(k) < 0$ for some k, then Proposition 14.19 would tell us that is definitely no extension. However, when $e_L(k) \geq 0$ for all k we can only say (for the moment) that the question remains open. To go beyond this we need another lemma and another theorem.

Lemma 14.21. *Let L be a $p \times q$ Latin rectangle, where $0 < p < n$ and $0 < q \leq n$, and suppose that every $k \in N$ is plausible for L. Then we can add an extra row to obtain a $(p+1) \times q$ Latin rectangle L' such that every $k \in N$ is still plausible for L'.*

Proof. We will again interpret this in terms of Remark 14.9. Here we have chosen students to do jobs 1 to q on days 1 to p, but we have not yet decided anything about jobs $q+1$ to n (perhaps the relevant managers are still on holiday). Our immediate task is to allocate students to jobs 1 to q on day $p+1$.

We again put $d = n-p > 0$. We again have a job for each $j \in Q$, with candidates $C_j = N \setminus \{L_{1j}, \dots, L_{pj}\}$, so $|C_j| = n-p = d$. We again put $R_k = \{j \mid k \in C_j\}$, which corresponds to the set of columns not containing k, or the set of jobs that student k can still do. Remark 14.12 tells us that the number of columns that do contain k is $m_L(k)$, so the number of columns that do not contain k is $|R_k| = q - m_L(k)$. The plausibility condition says that $m_L(k) + n - p - q \geq 0$, which translates to $q - m_L(k) \leq n - p = d$, so we see that $|R_k| \leq d$. In fact, we have $|R_k| = d$ iff $e_L(k) = 0$ iff k is barely plausible. The students with $|R_k| = d$ will be called "talented" (although in this model, the fact that they can still do many jobs is not really related to talent). By Corollary 12.3, we can solve the job allocation problem in such a way that every talented student gets assigned a job for day $p+1$. By adding this as a new row, we get a new Latin rectangle L' of size $(p+1) \times q$. Now note that

$$m_{L'}(k) = \begin{cases} m_L(k) + 1 & \text{if } k \text{ is in the new row} \\ m_L(k) & \text{otherwise} \end{cases}$$

so

$$e_{L'}(k) = m_{L'}(k) + n - p - q - 1 = \begin{cases} e_L(k) & \text{if } k \text{ is in the new row} \\ e_L(k) - 1 & \text{otherwise} , \end{cases}$$

so in particular $e_{L'}(k) \geq e_L(k) - 1$ in all cases. Thus, if $e_L(k) > 0$ then $e_{L'}(k) \geq 0$. On the other hand, if $e_L(k) = 0$ then k is barely plausible for L, so student k is "talented", so k appears in the new row by construction, so $e_{L'}(k) = e_L(k) = 0$. Thus, in all cases we have $e_{L'}(k) \geq 0$. □

Theorem 14.22. *Let L be a $p \times q$ Latin rectangle, and suppose that every $k \in N$ is plausible for L. Then L can be extended to an $n \times n$ Latin square.*

Proof. We can apply the lemma repeatedly until we get a $n \times q$ Latin rectangle, then we can apply Corollary 14.15 to get an $n \times n$ Latin rectangle. □

Example 14.23. We now show how to extend the the rectangle from Example 14.20. The process is controlled by the following two tables.

5	6	1	3	2	47^1	**4**
6	5	2	4	7	13^4	**3**
1	4	3	5	6	27^2	**7**
4	7	5	6	1	23^3	**2**
237	12**3**	46**7**	**1**27	3**45**	56^6	**6**
37	**1**2	**4**6	2**7**	3**5**	16^5	**1**
7	**1**	**6**	**2**	**3**	45^7	**5**

k	1	2	3	4	5	6	7
$e_L(k)$	1	0	0	1	2	2	0
$e_{L'}(k)$	1	0	0	1	1	1	0
$e_{L''}(k)$	0	0	0	1	1	0	0

In the left hand table, the top left block is the original matrix L. In the right hand table, the second row shows the excesses of $1, \dots, 7$ in L; in particular, the numbers 2, 3 and 7 have $e_L(k) = 0$ so they are barely plausible. We want to add a new row, making sure that we include the barely plausible numbers 2, 3 and 7. The possibilities for columns $1, \dots, 6$ are 237, 123, 467, 127 and 345, as shown in row 5 on the left. From these sets we choose 2, 3, 7, 1 and 4, as indicated by the bold entries in row 5. This gives a 5×5 Latin rectangle which we call L'. For the next step, we need to know the excesses for L', which we denote by

$e_{L'}(k)$. As we saw in the proof of Lemma 14.21, we have $e_{L'}(k) = e_L(k)$ if k appears in the new row, and $e_{L'}(k) = e_L(k) - 1$ if k does not appear in the new row. The resulting values are shown in row 3 of the right hand table. In particular, 2, 3 and 7 are barely plausible for L'. To get the potential entries for row 6, we simply take the sets of potential entries from row 5 and remove the bold ones, leaving 37, 12, 46, 27 and 35. We must choose five distinct numbers, one from each of these sets, in such a way that the barely plausible numbers 2, 3 and 7 are included. We choose $3, 2, 4, 7, 5$, as indicated by the bold entries in row 6. This gives a 6×5 Latin rectangle which we call L''. The excesses for E'' are again shown in the right hand table. However, we do not really need them, because there is now only one possible way to fill in row 7, namely $(7, 1, 6, 2, 3)$. This gives a 7×5 Latin rectangle. As this has the maximum possible height, we are back in the context of Corollary 14.15, and we do not need to keep track of excesses any more. To the right of the vertical bar, we have written the possible entries for column 7. As our first step (indicated by the superscript 1) we decide to try choosing 7 for the entry in row 1. For the second step (indicated by the superscript 2) we consider row 3. The possible choices there are 2 and 7, but we already used 7 for row 1, so we must use 2 for row 3. For the third step, we consider row 4. The possible choices there are 2 and 3, but we already used 2 for row 3, so we must use 3 for row 4. Continuing in the same way, we must use 1 in row 2, then 6 in row 6, then 5 in row 5, then 4 in row 7. This gives $(7, 1, 2, 3, 5, 6, 4)$ as column 6, and leaves $(4, 3, 7, 2, 6, 1, 5)$ as the only possibility for column 7. We end up with the following Latin square:

$$\begin{bmatrix} 5 & 6 & 1 & 3 & 2 & 7 & 4 \\ 6 & 5 & 2 & 4 & 7 & 1 & 3 \\ 1 & 4 & 3 & 5 & 6 & 2 & 7 \\ 4 & 7 & 5 & 6 & 1 & 3 & 2 \\ 2 & 3 & 7 & 1 & 4 & 5 & 6 \\ 3 & 2 & 4 & 7 & 5 & 6 & 1 \\ 7 & 1 & 6 & 2 & 3 & 4 & 5 \end{bmatrix}$$

Interactive demo

We now start to discuss the theory of orthogonal Latin squares. We will give an example before the definition.

Example 14.24. Consider the following matrices:

$$L = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix} \qquad M = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{bmatrix} \qquad L * M = \begin{bmatrix} 00 & 12 & 21 \\ 11 & 20 & 02 \\ 22 & 01 & 10 \end{bmatrix}$$

Both L and M are Latin squares. The matrix $L * M$ is formed by merging L and M in an obvious way: in symbols, the entry $(L*M)_{ij}$ is the ordered pair (L_{ij}, M_{ij}). There are 9 possible pairs uv with $u, v \in \{0, 1, 2\}$, as follows:

$$00,\ 01,\ 02, \quad 10,\ 11,\ 12, \quad 20,\ 21,\ 22.$$

It is not hard to check that each of these pair occurs precisely once in $L * M$.

Definition 14.25. Let L and M be two $n \times n$ Latin squares, with the same sets P, Q and N. Let $L * M$ be the matrix with entries $(L * M)_{ij} = (L_{ij}, M_{ij}) \in N^2$ for each $i \in P$ and $j \in Q$. We say that L and M are *orthogonal* if each of the n^2 elements of N^2 occurs precisely once in $L * M$. Equivalently, L and M are orthogonal if the entries in $L * M$ are all different.

Interactive demo

Example 14.26. Consider the following matrices:

$$L = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \qquad M = \begin{bmatrix} 1 & a & b \\ c & d & e \\ f & g & 2 \end{bmatrix} \qquad L * M = \begin{bmatrix} 11 & 2a & 3b \\ 3c & 1d & 2e \\ 2f & 3g & 12 \end{bmatrix}$$

We will try to find $a, \ldots, g$ such that M is a Latin square and is orthogonal to L.

- $L * M$ must contain 13 somewhere, and this can only happen if $d = 3$ so that 13 appears as the middle entry.
- In M, entry b lies in the same row as 1 and in the same column as 2, so it must be different from 1 and 2, so it must be equal to 3. By the same logic we also have $f = 3$.
- Now M is as shown on the left below. To make this a Latin square, each row must contain $1, 2$ and 3, and each column must contain $1, 2$ and 3. The only way to achieve this is to take $a = c = 2$ and $e = g = 1$, giving the matrix shown on the right below.

$$M = \begin{bmatrix} 1 & a & 3 \\ c & 3 & e \\ 3 & g & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

- We now have

$$L * M = \begin{bmatrix} 11 & 22 & 33 \\ 32 & 13 & 21 \\ 23 & 31 & 12 \end{bmatrix}.$$

  Inspection shows that each of the 9 possible pairs $11,\ 12,\ 13, 21,\ 22,\ 23, 31,\ 32,\ 33$ appears precisely once in $L * M$, so we have succeeded in finding a Latin square that is orthogonal to L.

We now discuss some facts about the number of possible $n \times n$ Latin squares.

Definition 14.27. We let $\mathcal{L}_n$ denote the set of all Latin squares L with $P = Q = N = \{1, \dots, n\}$. We will find $|\mathcal{L}_n|$ for $n \leq 4$. We say that a Latin square $L \in \mathcal{L}_n$ is *reduced* if the first row is $(1, 2, \dots, n)$ and the first column is also $(1, 2, \dots, n)$. We write $\mathcal{R}_n$ for the set of reduced Latin squares.

Example 14.28. For the degenerate case $n = 1$ the only possible Latin square is $L = \begin{bmatrix} 1 \end{bmatrix}$, so $\mathcal{L}_1 = \mathcal{R}_1$ and $|\mathcal{L}_1| = |\mathcal{R}_1| = 1$.

Example 14.29. For $n = 2$ we have $\mathcal{L}_2 = \left\{ \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \right\}$. The first of these lies in $\mathcal{R}_2$ but the second does not. We therefore have $|\mathcal{R}_2| = 1$ and $|\mathcal{L}_2| = 2$.

Example 14.30. For $n = 3$ there are 12 Latin squares, as follows:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 3 \\ 1 & 3 & 2 \\ 3 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 3 \\ 3 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{bmatrix} \quad \begin{bmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \\ 2 & 1 & 3 \end{bmatrix}$$

Of these only the first lies in $\mathcal{R}_3$. Thus, we have $|\mathcal{R}_3| = 1$ and $|\mathcal{L}_3| = 24$.

Proposition 14.31. *For any n we have $|\mathcal{L}_n| = n!(n-1)!|\mathcal{R}_n|$.*

Proof.
(a) We can permute the columns of a Latin square and it will still be a Latin square.
(b) We can also permute the rows of a Latin square and it will still be a Latin square.
(c) There is a unique way to permute the columns so that the first row becomes $(1, \dots, n)$.
(d) After we have done this, the top left entry will be 1, and the first entries in columns $2, \dots, n$ will therefore be $2, \dots, n$ in some order. Thus, there is a unique way to permute rows $2, \dots, n$ so that the first column becomes $1, \dots, n$. We now have a Latin square in $\mathcal{R}_n$.

By thinking about these steps in the reverse order, we obtain the following fact. We can obtain any Latin square in $\mathcal{L}_n$ by starting with a square in $\mathcal{R}_n$, permuting rows $2, \dots, n$ in any of $(n-1)!$ possible ways, then permuting the columns in any of $n!$ possible ways. The claim is clear from this. □

Proposition 14.32. *We have $|\mathcal{R}_4| = 4$ and so $|\mathcal{L}_4| = 4! \times 3! \times 4 = 576$.*

Proof. We claim that $\mathcal{R}_4$ consists of the following 4 squares. The superscripts are just there to help us follow the proof.

$$L^1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3^{0*} & 4^{2} & 1^{1} \\ 3 & 4^{4} & 1^{6} & 2^{5} \\ 4 & 1^{3} & 2^{7} & 3^{8} \end{bmatrix} \quad L^2 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4^{0*} & 1^{1} & 3^{2} \\ 3 & 1^{3} & 4^{6} & 2^{5} \\ 4 & 3^{4} & 2^{7} & 1^{8} \end{bmatrix} \quad L^3 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1^{0*} & 4^{1} & 3^{2} \\ 3 & 4^{3} & 1^{5*} & 2^{6} \\ 4 & 3^{4} & 2^{7} & 1^{8} \end{bmatrix} \quad L^4 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1^{0*} & 4^{1} & 3^{2} \\ 3 & 4^{3} & 2^{5*} & 1^{6} \\ 4 & 3^{4} & 1^{7} & 2^{8} \end{bmatrix}$$

The numbers in the superscripts indicate the order in which we should think about the entries; the stars indicate places where we have a choice about what to do.

The first row and column have to be $(1, 2, 3, 4)$. Thus, the first place where we have any choice is the $(2, 2)$ position, which we have marked with the superscript $0*$. We already have a 2 in the corresponding row (and also in the corresponding column), so we cannot put a 2 in this position; we must have a 3, a 4 or a 1. For square L^1 we choose to put a 3 in the $(2, 2)$ position. It turns out that we then have no more choices. To see this, consider the superscript 1, which appears in position $(2, 4)$ in L^1. There we have already placed a 2 and a 3 in the same row and a 4 in the same column, so we have to put a 1 in that slot. Now consider the superscript 2, which appears in position $(2, 3)$ in L^1. We have already placed 1, 2 and 3 in the same row, so we are forced to put 4 in this slot. Now continue with the positions with superscripts $3, 4, \ldots, 8$; we again see that there is never any choice, and we have to fill in the entries as in L^1. Thus, L^1 is the only possible Latin square in $\mathcal{R}_4$ that has a 3 in position $(2, 2)$. Similarly, L^2 is the only Latin square in $\mathcal{R}_4$ that has a 4 in position $(2, 2)$. The only other possibility is to put a 1 in position $(2, 2)$, as in L^3. Just as in the case of L^1 and L^2, we find that there is no choice about what to put in the positions with superscripts $1, \ldots, 4$. However, when we get to the superscript $5*$ in position $(3, 3)$, we find that we do have a choice: we can either put in a 1 or a 2. If we put in a 1 then we are forced to fill in the remaining 3 slots as in L^3, but if we put in a 2 then we are forced to fill in the remaining 3 slots as in L^4. We thus have $\mathcal{R}_4 = \{L^1, L^2, L^3, L^4\}$ as claimed.

Interactive demo

□

Remark 14.33. The numbers $|\mathcal{R}_n|$ grow very quickly as n increases:

n	$\lvert\mathcal{R}_n\rvert$
1	1
2	1
3	1
4	4
5	56
6	9,408
7	16,942,080
8	535,281,401,856

We will not prove any of this.

We now consider a different problem. Given $n > 0$, can we find a long list of $n \times n$ Latin squares $L^1, \ldots, L^r$ such that L^u and L^v are orthogonal when $u \neq v$? Our first result gives an upper bound on the possible length of such a list.

Theorem 14.34. *Suppose we have a list $L^1, \ldots, L^r$ of mutually orthogonal Latin squares of size n. Then $r \leq n - 1$.*

Proof. Look at the first two rows of L^u:

	1		m_u		n
1			x		
2	x				

In position $(2,1)$ (at the beginning of the second row), we have some number $x \in \{1,\dots,n\}$. Because L^u is a Latin square, every number must appear in every row. In particular, x must also appear in row 1. It cannot appear in position $(1,1)$, because we would then have two x's in the first column. So x must also appear at position $(1,m_u)$ for some $m_u \in \{2,\dots,n\}$. We have now defined numbers $m_1,\dots,m_r$; we claim that they are all different. Indeed, suppose that $v \neq u$, so the first two rows of L^v have the form

	1		m_v		n
1			y		
2	y				

If m_u and m_v were the same, then in $L^u * L^v$ we would have a pattern like this:

	1		$m_u = m_v$		n
1			xy		
2	xy				

so xy would appear twice in $L^u * L^v$. However, L^u and L^v are assumed to be orthogonal, so each pair occurs precisely once in $L^u * L^v$, so xy cannot appear twice, so m_v must be different from m_u.

We now know that the numbers $m_1,\dots,m_r$ are all different and all lie in $\{2,\dots,n\}$. This is clearly only possible if $r \leq n-1$. □

We now see that the maximum possible length of a list of mutually orthogonal $n \times n$ Latin squares is at most $n-1$. Can we achieve this upper bound? A key example is as follows.

Proposition 14.35. *Let p be a prime. For $0 < u < p$ define $L^u_{ij} = i + uj \pmod p$. Then L^u is a Latin square (with $P = Q = N = \mathbb{Z}/p$). Moreover, L^u and L^v are orthogonal if $u \neq v$ (so we have a list of $p-1$ mutually orthogonal Latin squares of size p).*

Proof. First recall that $\mathbb{Z}/p$ is a field. Thus, if $a, b \in \mathbb{Z}/p$ with $b \neq 0$ then a/b makes sense as an element of $\mathbb{Z}/p$, and fractions like this have all the usual properties.

Now fix u with $0 < u < p$, and put $L^u_{ij} = i + uj$. If $L^u_{ij} = L^u_{i'j}$ we have $i + uj = i' + uj$ so $i = i'$. Similarly, if $L^u_{ij} = L^u_{ij'}$ then $i + uj = i + uj'$ in $\mathbb{Z}/p$. We can rearrange to get $u.(j - j') = 0$ but u is invertible in $\mathbb{Z}/p$ so we can multiply by u^{-1} to get $j - j' = 0$ and so $j = j'$ (in $\mathbb{Z}/p$). This shows that L^u is a Latin square.

Now consider $L^u * L^v$, where $0 < u, v < p$ with $u \neq v$, so $(L^u * L^v)_{ij} = (i + uj, i + vj)$. We want to show that every pair $(x, y) \in (\mathbb{Z}/p)^2$ appears precisely once in this table, or in other words that there is a unique pair (i, j) with$(i + uj, i + vj) = (x, y)$, or that the simultaneous equations $i + uj = x$ and $i + vj = y$ have a unique solution. These equations can be solved in the standard way to give $i = (vx - uy)/(v - u)$ and $j = (x - y)/(u - v)$ as required. □

Remark 14.36. Now consider a number n that is a prime power, say $n = p^v$ for some prime number p and some $v > 1$. In this case the ring $\mathbb{Z}/n$ is not a field, but there is a more complicated way to define a field F with $|F| = n$. We will not discuss the construction here, but it can be found in most books on field theory. Now suppose that $u \in F$ with $u \neq 0$ (so there are $n-1$ possible choices for u). We can again define a Latin square L^u with $P = Q = N = F$ by $L^u_{ij} = i + uj$, and we again find that L^u and L^v are orthogonal when $u \neq v$. Thus, we have a list of $n-1$ mutually orthogonal Latin squares of size n.

The first number that is not a prime or prime power is 6. This case is already hard.

Theorem 14.37. *There are not even two mutually orthogonal Latin squares of size 6.*

Proof. This was conjectured by Euler in the 18th century, and proved by Tarry in 1900. A more digestible proof was given by Stinson in 1982. We will not give any details here. □

15. Block Designs

We now consider matching problems again, but from a rather different point of view. Before, we were given a matching problem and we tried to solve it, or count the number of possible solutions. Here instead we will try to find matching problems that have certain special properties, which in particular make them highly symmetrical. Highly symmetrical combinatorial objects are always interesting and often have applications. In particular, the material in this chapter can be used for efficient design of experiments where one wants to test multiple interacting factors without performing more tests than necessary. It can also be used to design computer communication systems that can detect and correct some transmission errors.

Before, we had a set A of people and a set B of jobs, and for each job $b \in B$ we had a subset $C_b \subseteq A$ of people who are qualified to do that job. For each person a we also considered the set R_a of jobs that they are qualified to do. This can be expressed in symbols as $R_a = \{b \in B \mid a \in C_b\}$.

The framework in this chapter will be mathematically equivalent but we will follow tradition in using slightly different terminology. We will have a set B of "blocks" and a set V of "varieties". For each block $j \in B$ we have a corresponding subset $C_j \subseteq V$. For any variety $p \in V$ we again define $R_p = \{j \in B \mid p \in C_j\}$.

Video (Definition 15.1 and Proposition 15.4)

Definition 15.1. Consider numbers $v, b, r, k, \lambda > 0$ with $k < v$ and $r < b$. A *block design* with parameters (v, b, r, k, λ) is a matching problem as above, with the following properties:

(a) $|V| = v$
(b) $|B| = b$
(c) $|R_p| = r$ for all $p \in V$
(d) $|C_j| = k$ for all $j \in B$
(e) $|R_p \cap R_q| = \lambda$ for all $p, q \in V$ with $p \neq q$.

In words: there are v varieties and b blocks, every variety is in precisely r blocks, every block contains precisely k varieties, every pair of distinct varieties is in precisely λ blocks.

Remark 15.2. As $C_j \subseteq V$ and $|C_j| = k$ and $|V| = v$ it is automatic that $k \leq v$. If k were equal to v then that would mean that $C_j = V$ for all j, which is like a job allocation problem in which every person is qualified to do every job. However, we specified as part of the definition that $k < v$, so as to exclude this uninteresting case. The condition $r < b$ also has the same effect.

Example 15.3. Put $B = \{1, \dots, 12\}$ and $V = \{1, \dots, 9\}$ and

$$C_1 = \{1,2,3\} \quad C_2 = \{4,5,6\} \quad C_3 = \{7,8,9\}$$
$$C_4 = \{1,4,7\} \quad C_5 = \{2,5,8\} \quad C_6 = \{3,6,9\}$$
$$C_7 = \{1,5,9\} \quad C_8 = \{2,6,7\} \quad C_9 = \{3,4,8\}$$
$$C_{10} = \{1,6,8\} \quad C_{11} = \{2,4,9\} \quad C_{12} = \{3,5,7\}$$

The corresponding sets R_p are

$$R_1 = \{1,4,7,10\} \quad R_2 = \{1,5,8,11\} \quad R_3 = \{1,6,9,12\}$$
$$R_4 = \{2,4,9,11\} \quad R_5 = \{2,5,7,12\} \quad R_6 = \{2,6,8,10\}$$
$$R_7 = \{3,4,8,12\} \quad R_8 = \{3,5,9,10\} \quad R_9 = \{3,6,7,11\}.$$

It is now visible that $|V| = 9$ and $|B| = 12$ and $|C_j| = 3$ for all j and $|R_p| = 4$ for all p. We also have

$$R_1 \cap R_2 = \{1\} \quad R_3 \cap R_4 = \{9\} \quad R_3 \cap R_6 = \{6\} \quad R_4 \cap R_9 = \{11\}.$$

In fact, we have $|R_p \cap R_q| = 1$ for all $p \neq q$, as we can see by a long but easy check of cases. Thus, the above sets give a $(9, 12, 4, 3, 1)$ block design.

Interactive demo

Proposition 15.4. *If there is a (v, b, r, k, λ)-block design, then $bk = vr$ and $bk(k-1) = \lambda v(v-1)$ and $r(k-1) = \lambda(v-1)$ and $\lambda < r$.*

Proof. Put

$$\begin{aligned} X &= \{(j,p) \in B \times V \mid p \in C_j\} \\ &= \{(j,p) \in B \times V \mid j \in R_p\}. \end{aligned}$$

We can use the first description to find $|X|$: there are b ways to choose $j \in B$, and then $|C_j| = k$ ways to choose $p \in C_j$, so $|X| = bk$. Alternatively, we can use the second description. There are v ways to choose $p \in V$, and then $|R_p| = r$ ways to choose $j \in R_p$, so $|X| = vr$. By comparing these, we see that $bk = vr$. Now put

$$\begin{aligned} Y &= \{(j,p,q) \in B \times V \times V \mid p, q \in C_j,\ q \neq p\} \\ &= \{(j,p,q) \in B \times V \times V \mid q \neq p, j \in R_p \cap R_q\}. \end{aligned}$$

We can again use the first description to find $|Y|$: there are b ways to choose $j \in B$, then k ways to choose $p \in C_j$, then $k-1$ ways to choose a different element $q \in C_j$, giving $|Y| = bk(k-1)$. Alternatively, we can use the second description: there are v ways to choose $p \in V$, then $v-1$ ways to choose a different element $q \in V$, then $|R_p \cap R_q| = \lambda$ ways to choose $j \in R_p \cap R_q$, giving $|Y| = \lambda v(v-1)$. By comparing these, we get $bk(k-1) = \lambda v(v-1)$. We can now substitute our first equation $bk = vr$ into our second equation $bk(k-1) = \lambda v(v-1)$ and then divide by v to get $r(k-1) = \lambda(v-1)$. Rearranging this, we get $\lambda/r = (k-1)/(v-1)$. As one of our axioms we assumed that $k < v$, so $(k-1)/(v-1) < 1$, so $\lambda/r < 1$, so $\lambda < r$. □

It is useful to have a slight variant of the above proposition. This shows that axiom (c) in Definition 15.1 is not really needed, because it follows from the other axioms.

Proposition 15.5. *Suppose that we have a matching problem as before, and numbers $v, b, k, \lambda > 0$ with $k < v$. Suppose that $|V| = v$ and $|B| = b$ and $|C_j| = k$ for all j and $|R_p \cap R_q| = \lambda$ for all $p \neq q$, so axioms (a), (b), (d) and (e) from Definition 15.1 are satisfied. Then the number $r = \lambda(v-1)/(k-1)$ is an integer and is the same as bk/v. Moreover, we have $|R_p| = r$ for all p, so axiom (c) is also satisfied, and we have a block design.*

Proof. We define X and Y as in the proof of Proposition 15.4. Most of that proof still works: we have $|X| = bk$ and $|Y| = bk(k-1) = \lambda v(v-1)$. However, if we use the second description of X we just get $|X| = \sum_p |R_p|$, and we do not yet know that the sets R_p all have the same size. For this, we fix p and define

$$\begin{aligned} Z_p &= \{(j,q) \in B \times V \mid q \neq p \text{ and } j \in R_p \cap R_q\} \\ &= \{(j,q) \in B \times V \mid j \in R_p \text{ and } q \in C_j \setminus \{p\}\} \end{aligned}$$

In the first description, there are $v-1$ ways to choose q and then λ ways to choose j, so $|Z_p| = \lambda(v-1)$. In the second description, there are $|R_p|$ ways to choose j. We then have $|C_j| = k$ and $p \in C_j$ (because $j \in R_p$) so $|C_j \setminus \{p\}| = k-1$, so there are $k-1$ possible choices for q. This gives $|Z_p| = (k-1)|R_p|$, and we can compare our two formulae for $|Z_p|$ to get $|R_p| = \lambda(v-1)/(k-1)$. The right hand side is precisely the number called r in the statement of this proposition, so $|R_p| = r$ for all p. The left hand side here is clearly a nonnegative integer, so r is an integer. The formula $|X| = \sum_p |R_p|$ now becomes $|X| = vr$, so we again have $bk = vr$ and so $r = bk/v$. All claims are now clear. □

We next discuss an interesting construction that uses some number theory to produce a block design.

Video (Definition 15.6 to Lemma 15.12)

Definition 15.6. Let p be a prime number of the form $p = 4n+3$, so

$$\mathbb{Z}/p = \{0, \pm 1, \pm 2, \dots, \pm(2n+1)\}.$$

We put

$$Q = \{i \in \mathbb{Z}/p \mid i = j^2 \text{ for some } j \in \mathbb{Z}/p \text{ with } j \neq 0\},$$

and call this the set of *quadratic residues*. We then have a matching problem with $B = V = \mathbb{Z}/p$ and $C_j = j + Q$.

Remark 15.7. We have $m \in C_j$ iff $m \in j + Q$ iff $m - j \in Q$ iff $j \in m - Q$, so $R_m = m - Q$.

Example 15.8. Take $p = 7$, so $p = 4n + 3$ with $n = 1$ and $\mathbb{Z}/p = \{0, \pm 1, \pm 2, \pm 3\}$. We have $(\pm 1)^2 = 1$ and $(\pm 2)^2 = 4 = -3 \pmod 7$ and $(\pm 3)^2 = 9 = 2 \pmod 7$, so $Q = \{1, 2, -3\}$. This gives

$$\begin{aligned} C_0 &= \{1, 2, -3\} & R_0 &= \{-1, -2, 3\} \\ C_1 &= \{2, 3, -2\} & R_1 &= \{0, -1, -3\} \\ C_2 &= \{3, -3, -1\} & R_2 &= \{1, 0, -2\} \\ C_3 &= \{-3, -2, 0\} & R_3 &= \{2, 1, -1\} \\ C_{-1} &= \{0, 1, 3\} & R_{-1} &= \{-2, -3, 2\} \\ C_{-2} &= \{-1, 0, 2\} & R_{-2} &= \{-3, 3, 1\} \\ C_{-3} &= \{-2, -1, 1\} & R_{-3} &= \{3, 2, 0\}. \end{aligned}$$

One can check that $|R_l \cap R_m| = 1$ whenever $l \neq m$, so this is a $(7, 7, 3, 3, 1)$-block design.

Interactive demo

Example 15.9. Take $p = 11$, so $p = 4n + 3$ with $n = 2$ and $\mathbb{Z}/p = \{0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5\}$. We have $(\pm 1)^2 = 1$ and $(\pm 2)^2 = 4$ and $(\pm 3)^2 = 9 = -2 \pmod{11}$ and $(\pm 4)^2 = 16 = 5 \pmod{11}$ and $(\pm 5)^2 = 25 = 3 \pmod{11}$, so

$$Q = \{1, -2, 3, 4, 5\}.$$

I particular, we have $|Q| = 5$ and so $|C_j| = 5$ for all j and $|R_m| = 5$ for all m. We also have

$$R_0 \cap R_1 = (-Q) \cap (1 - Q) = \{-1, 2, -3, -4, -5\} \cap \{0, 3, -2, -3, -4\} = \{-3, -4\},$$

so $|R_0 \cap R_1| = 2$. In fact we have $|R_l \cap R_m| = 2$ for all $l \neq m$, so we have a $(11, 11, 5, 5, 2)$-block design. This will follow from Theorem 15.14, which we will prove below.

Lemma 15.10. *For each $i \in \{1, \dots, 2n + 1\}$, precisely one of i and $-i$ is in Q. Thus, $|Q| = 2n + 1$.*

This result is clearly visible in the cases $p = 7$ (where $Q = \{1, 2, -3\}$) and $p = 11$ (where $Q = \{1, -2, 3, 4, 5\}$).

Proof. This is a standard piece of number theory. One key ingredient is a theorem saying that $(\mathbb{Z}/p) \setminus \{0\}$ is cyclic of order $4n + 2$ when considered as a group under multiplication. If g is a generator, one can deduce that g^{2n+1} must be equal to -1. We will not give further details here. □

From Lemma 15.10 it is clear that $|C_j| = 2n + 1$ for all j, and that $|R_m| = 2n + 1$ for all m. However, it is not yet clear what we can say about $|R_l \cap R_m|$ when $l \neq m$. For this we need some more definitions.

Definition 15.11. We put $D = \{(u, v) \in Q \times Q \mid u \neq v\}$, so $|D| = |Q|(|Q| - 1)$. As $|Q| = 2n + 1$, this gives $|D| = (4n + 2)n$. Also, for $x \in \mathbb{Z}/p$ with $x \neq 0$ we put $D_x = \{(u, v) \in D \mid u - v = x\}$. We note that D is the disjoint union of the subsets D_x, so $|D| = \sum_x |D_x|$.

Lemma 15.12. *$|D_x| = n$ for all x.*

Proof. Recall from Lemma 15.12 that either x or $-x$ is a square. Suppose for the moment that x is a square. Suppose that $(u, v) \in D_1$, so u and v are squares with $u - v = 1$. It is clear that the product of two squares is a square, so ux and vx are squares with $ux - vx = x$, so $(ux, vx) \in D_x$. Conversely, if $(u', v') \in D_x$ then $(u'/x, v'/x) \in D_1$. From this it is clear that $|D_x| = |D_1|$.

Now suppose instead that $-x$ is a square. If $(u, v) \in D_1$ then $-vx$ and $-ux$ are squares with $(-vx) - (-ux) = (u - v)x = x$, so $(-vx, -ux) \in D_x$. Conversely, if $(u', v') \in D_x$ then $(-v'/x, -u'/x) \in D_1$. From this it is again clear that $|D_x| = |D_1|$.

We now see that $|D_x| = |D_1|$ in all cases, and the number of possibilities for x is $p - 1 = 4n + 2$. The equation $|D| = \sum_x |D_x|$ now becomes $|D| = (4n + 2)|D_1|$. However, we saw previously that $|D| = (4n + 2)n$, so $|D_1| = n$, so $|D_x| = n$ for all x. □

Example 15.13. We will show how the above lemma works out in the case where $p = 11$ and so $n = 2$ and $Q = \{1, -2, 3, 4, 5\}$. The table below shows the differences $u - v$ for $u, v \in Q$ with $u \neq v$.

u \ v	1	−2	3	4	5
1	■	3	−2	−3	−4
−2	−3	■	−5	5	4
3	2	5	■	−1	−2
4	3	−5	1	■	−1
5	4	−4	2	1	■

We can read off the sets D_x from this. For example, to find D_5 we look in the table and see that 5 appears in the position where $u = -2$ and $v = 4$, and also in the position where $u = 3$ and $v = -2$. We therefore have $D_5 = \{(-2, 4), (3, -2)\}$. The complete list of sets D_x is as follows:

$$D_1 = \{(4,3),(5,4)\} \qquad D_{-1} = \{(3,4),(4,5)\}$$
$$D_2 = \{(3,1),(5,3)\} \qquad D_{-2} = \{(1,3),(3,5)\}$$
$$D_3 = \{(1,-2),(4,1)\} \qquad D_{-3} = \{(-2,1),(1,4)\}$$
$$D_4 = \{(-2,5),(5,1)\} \qquad D_{-4} = \{(5,-2),(1,5)\}$$
$$D_5 = \{(-2,4),(3,-2)\} \qquad D_{-5} = \{(4,-2),(-2,3)\}$$

We find that $|D_x| = 2 = n$ in every case, as predicted by the lemma.

Theorem 15.14. *The matching problem in Definition 15.6 is a $(4n+3, 4n+3, 2n+1, 2n+1, n)$-block design.*

Proof. Video

All that is left is to show that $|R_l \cap R_m| = n$ for all $l \neq m$. Recall that $R_l = l - Q$, so $j \in R_l$ iff $l - j \in Q$. Thus, if $j \in R_l \cap R_m$ we see that $l-j, m-j \in Q$ and of course $(l-j)-(m-j) = l-m$ so $(l-j, m-j) \in D_{l-m}$. We can therefore define a map $f\colon R_l \cap R_m \to D_{l-m}$ by $f(j) = (l-j, m-j)$. In the opposite direction, suppose that $(u, v) \in D_{l-m}$, so $u, v \in Q$ with $u - v = l - m$ or equivalently $l - u = m - v$. If we put $j = l - u = m - v$ then we find that $j \in R_l$ (because $R_l = l - Q$ and $j = l - u$) and also $j \in R_m$ (because $R_m = m - Q$ and $j = m - v$), so $j \in R_l \cap R_m$. Using this we see that f is a bijection, so $|R_l \cap R_m| = |D_{l-m}|$. We also know from Lemma 15.12 that $|D_{l-m}| = n$, so $|R_l \cap R_m| = n$ as required. □

We now discuss some relationships between matrix algebra and the theory of block designs.

Definition 15.15. Suppose we have a matching problem as before, with $|B| = b$ and $|V| = v$, but we do not necessarily assume that this is a block design. We define a $b \times v$ matrix M by

$$M_{jp} = \begin{cases} 1 & \text{if } p \in C_j \\ 0 & \text{if } p \notin C_j. \end{cases}$$

Remark 15.16. In earlier chapters, we drew a board with the square in position (j, p) coloured white if $p \in C_j$, and coloured black if $p \notin C_j$. The matrix M is essentially the same thing but with 1's corresponding to white squares and 0's corresponding to black squares.

Example 15.17. For the block design in Example 15.3 we have

$$M = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Note that the transpose M^T is a $v \times b$ matrix, so the product $M^T M$ is defined and is a $v \times v$ matrix. We will need some other $v \times v$ matrices, as follows.

Definition 15.18. We let I denote the $v \times v$ identity matrix, so I_{pq} is 1 if $p = q$ and 0 if $p \neq q$. We also let 1_m denote the column vector of length m with all entries equal to one. We let J denote the $v \times v$ matrix with $J_{pq} = 1$ for all p and q, so all of the columns are 1_v. Note that $(J - I)_{pq}$ is 0 if $p = q$ and 1 if $p \neq q$.

Example 15.19. In the case $v = 4$ we have

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad J = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad J - I = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

In the theorem below, we will consider the matrix $rI + \lambda(J - I)$, which has diagonal entries equal to r and off-diagonal entries equal to λ. In the case $n = 4$ we get

$$rI + \lambda(J - I) = \begin{bmatrix} r & \lambda & \lambda & \lambda \\ \lambda & r & \lambda & \lambda \\ \lambda & \lambda & r & \lambda \\ \lambda & \lambda & \lambda & r \end{bmatrix}.$$

Theorem 15.20. *Our matching problem is a (v, b, r, k, λ)-block design iff $M.1_v = k.1_b$ and $M^T M = rI + \lambda(J - I) = (r - \lambda)I + \lambda J$.*

Proof. We will use the following standard formulae for vector and matrix algebra: the product of a matrix P and a vector u is $(Pu)_i = \sum_j P_{ij} v_j$, and the product of two matrices is $(PQ)_{ik} = \sum_j P_{ij} Q_{jk}$.

From these we get $(M.1_v)_j = \sum_p M_{jp}(1_v)_p$. Here $(1_v)_p = 1$ for all p, so we just get $(M.1_v)_j = \sum_p M_{jp}$. In this sum we get a contribution of 1 for each $p \in C_j$, and 0 for each $p \notin C_j$, so the sum is just equal to $|C_j|$. Thus $M.1_v$ is just the vector $(|C_1|, \dots, |C_b|)$, and this is equal to $k.1_b$ iff $|C_j| = k$ for all j. Thus, the condition $M.1_v = k.1_b$ is equivalent to axiom (d) in Definition 15.1.

We now consider the product $M^T M$. We have $(M^T)_{pj} = M_{jp}$, so

$$(M^T M)_{pq} = \sum_j (M^T)_{pj} M_{jq} = \sum_j M_{jp} M_{jq}.$$

The j'th term here is 1 if both p and q lie in C_j, and zero otherwise. Equivalently, the j'th term is 1 if $j \in R_p \cap R_q$, and zero otherwise. Taking the sum over j, we get $(M^T M)_{pq} = |R_p \cap R_q|$. In the case $p = q$ this just reduces to $(M^T M)_{pp} = |R_p|$.

We want to compare $M^T M$ with the matrix $N = rI + \lambda(J - I)$, which has $N_{pp} = r$ and $N_{pq} = \lambda$ for $p \neq q$. We find that $M^T M = N$ iff $|R_p| = r$ for all p, and $|R_p \cap R_q| = \lambda$ for all $p \neq q$. These are just axioms (c) and (e) in Definition 15.1, so the claim is now clear. □

Lemma 15.21. *The matrix J has eigenvalues 0 (repeated $v - 1$ times) and v (repeated once).*

Proof. Let $e_1, \dots, e_v$ be the standard basis of $\mathbb{R}^v$. Define an alternative basis $a_1, \dots, a_v$ by $a_1 = e_1 - e_2$, $a_2 = e_2 - e_3$ and so on up to $a_{v-1} = e_{v-1} - e_v$, followed by $a_n = e_1 + \cdots + e_v = 1_v$. For example, when $v = 4$ we have

$$a_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \qquad a_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \qquad a_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \qquad a_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

It is then easy to check that $Ja_1 = \cdots = Ja_{v-1} = 0$ but $Ja_v = v\, a_v$. In other words, the vectors a_i are eigenvectors for J with eigenvalues $0, 0, \dots, 0, v$. The claim is clear from this. □

Corollary 15.22. *If our matching problem is a block design, then* $\det(M^T M) > 0$ *and* $M^T M$ *is invertible.*

Proof. By the theorem, $M^T M = (r - \lambda)I + \lambda J$. By Lemma 15.21, the matrix J has eigenvalues 0 and v, so λJ has eigenvalues 0 and λv, so $(r-\lambda)I + \lambda J$ has eigenvalues $r - \lambda$ and $r - \lambda + \lambda v = r + (v-1)\lambda$. Moreover, the first of these is repeated $v - 1$ times, but the second occurs only once. The determinant is the product of the eigenvalues, which is $(r - \lambda)^{v-1}(r + (v-1)\lambda)$. We also know from Proposition 15.4 that $\lambda < r$, so $r - \lambda > 0$, so $\det(M^T M) > 0$. It is a standard fact from linear algebra that a square matrix with nonzero determinant is invertible. □

Corollary 15.23. *If our matching problem is a block design, then* $b \geq v$.

Proof. Suppose instead that $b < v$; we will derive a contradiction. Let the columns of M be $u_1, \dots, u_v$, so $u_p \in \mathbb{R}^b$ for all p. As $b < v$, the length of this list is larger than the dimension of the space $\mathbb{R}^b$, so the vectors u_p must be linearly dependent, by a standard theorem of linear algebra. Thus, there is an equation $c_1 u_1 + \cdots + c_v u_v = 0$ where the coefficients c_i are not all zero. We can form a vector c with entries $(c_1, \dots, c_v)$, so $c \neq 0$. After thinking about how matrix multiplication works, we see that $Mc = 0$ and so $M^T Mc = 0$. As $M^T M$ is invertible, it follows that $c = 0$, which is a contradiction. □

Note that the conclusion $b \geq v$ is a purely combinatorial fact, so it is interesting that we have had to make a detour into linear algebra to prove it.

Definition 15.24. A *symmetric* design is one in which $b = v$. Corollary 15.23 tells us that in some sense these are maximally efficient. Recall from Proposition 15.4 that $bk = rv$. From this we see that a symmetric design also satisfies $k = r$.

Note that the quadratic residue design from Definition 15.6 and Theorem 15.14 is symmetric, but the design in Example 15.3 is not.