Dragan Mašulović

The Discrete Charm of Discrete Mathematics

Linz, January 2006

Preface

The development of Discrete Mathematics has gained momentum in the second half of the 20th century with the introduction of computers. Today, it is one of the most vivid mathematical disciplines, a *must* for every mathematician/computer scientist of the 21st century.

The objective of the course is to provide an overview of the main topics and techniques of Discrete Mathematics. The emphasis will be on the investigation of the most fundamental combinatorial structures. In this course we address some of the most importand topics in Discrete Mathematics:

- Elementary combinatorial configurations (permutations and variations) and basic counting;
- Systems of distinct representatives and latin squares;
- Combinatorial designs and finite geometries;
- Eulerian and Hamiltonian graphs and NP-hard problems;
- Planarity and the Four Colour Problem.

These lecture notes have been compiled during my stay at the Institute of Algebra of the Johannes Kepler University in Linz, Austria, where I gave a course on Discrete Mathematics in the Winter Semester of the academic year 2005/6.

I would like to express my deepest gratitute to Prof. Dr. Günter Pilz, the Head of the Institute of Algebra and Vice-Rector of the Johannes Kepler University in Linz, and Dr. Erhard Aichinger from the Institute of Algebra. None of this would have been possible without their help, support and friendship.

Linz, January 2006 Dragan Mašulović

Contents

Chapter 1

Words and Sets

This chapter confronts us with the most basic abstract structures:

- words (or strings), which represent the simplest *ordered* structures, and
- sets (or collections), which represent the simplest *unordered* structures.

As we shall see, a *permutation* is nothing but a word over an appropriately chosen alphabet, while a *combination* is just a subset of a finite set. It is natural to ask why should one invent so complicated names for such simple objects. The answer is simple. In the dark past of Discrete Mathematics the terminology used to be as obscure as the ages that gave birth to it. Since the introduction of the names such as *permutation* and *combination* mathematics has gone a long way and brought many simplifications, both in terminology and understanding of the phenomena.

Throughout the course we shall use the followin notation

$$\begin{aligned}
&\mathbb{N} = \{1,2,3,\ldots\} \quad \text{for the set of positive integers,}\\
&\mathbb{N}_0 = \{0,1,2,3,\ldots\} \quad \text{for the set of nonnegative integers, and}\\
&\mathbb{N}_0^\infty = \{0,1,2,3,\ldots\} \cup \{\infty\}.
\end{aligned}$$

The set $\mathbb{N}_0^\infty$ is a usual extension of $\mathbb{N}_0$ with the greatest element ∞: $x+\infty = \infty + x = x \cdot \infty = \infty \cdot x = \infty$ for all $x \in \mathbb{N}_0^\infty$, and $x < \infty$ for all $x \in \mathbb{N}_0$. Also, we define the *factorial* of an integer $n \in \mathbb{N}_0$ as usual:

$$\begin{aligned}
&0! = 1\\
&n! = 1 \cdot 2 \cdot \ldots \cdot n, \text{ for } n \geqslant 1.
\end{aligned}$$

1.1 Words

An *alphabet* is any finite nonempty set. Elements of an alphabet A will be referred to as *letters*, and a *word in* A is a string of symbols from A. More precisely, a *word of length* k *over an alphabet* A is any tuple from A^k. We follow a simple convention to omit commas and parentheses when writing words.

Example 1.1 Here are some words over an alphabet $A = \{a, b, n\}$: *banana*, *abba*, *aa*, or simply *n*. The first of the words has six letters, then comes a four-letter word, a two-letter word and finaly a word with only one letter.

We also alow words with no letters. On any alphabet there is precisely one such word called the *empty word* and denoted by ε. It is a word with length 0. It is important to note that words we deal with in this course are *formal words*, that is, strings of symbols to which no meaning is attached. So, from this point of view *nbbaaa* is just as good a word as *banana*. We shall leave the meaning of words to other branches of science and treat words just as plain and simple strings of letters.

Let w be a word over an alphabet A. The length of w will be denoted by $|w|$. For a letter $a \in A$, by $|w|_a$ we denote the number of occurences of a in w.

Example 1.2 Let $A = \{a, b, c, n\}$ and let $w =$ *banana* be a word over A. Then $|w| = 6$, $|w|_a = 3$, $|w|_b = 1$, $|w|_c = 0$ and $|w|_n = 2$.

There is not much structural theory behind so simple objects such as words. The most exciting thing we can do at the moment is to try to count them.

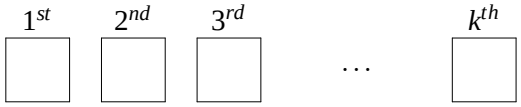
Problem 1.3 *Let $A = \{a_1, a_2, \ldots, a_n\}$ be an alphabet with $n \geqslant 1$ letters and let $k \in \mathbb{N}_0$ be arbitrary.*

(a) How many words with k letters over A are there?

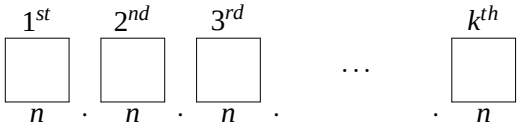
(b) How many words with k letters over A have the property that all the letters in the word are distinct?

(c) How many words over A have the property that every letter from A appears precisely once in the word?

Solution. (a) The set of all words of length k over A is just A^k. Therefore, there are precisely $|A^k| = \underbrace{|A| \cdot \ldots \cdot |A|}_{k} = n^k$ such words. There is a less formal, but more useful way to see this. A word with k letters looks line this:

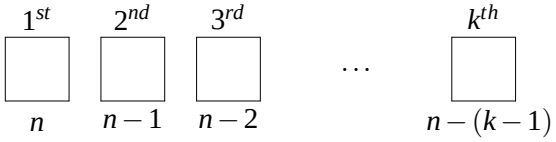


There are n candidates for the first position, n candidates for the second position, $\ldots$, n candidates for the kth position:



Alltogether, there are $\underbrace{n \cdot n \cdot \ldots \cdot n}_{k} = n^k$ possibilites.

(b) Let us again take the informal point of view. Firstly, there are n candidates for the first position, but only $n-1$ candidates for the second position, since the letter used on the first position is not allowed to appear on the second position. Then, there are $n-2$ candidates for the third position since the two letters used on the first two positions are a no-no, and so on. Finally, there will be $n-(k-1)$ candidates for the last position:



and putting it all together we get $n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \dfrac{n!}{(n-k)!}$ possibilites. Of course, this reasoning is valid as long as $k \leqslant n$. If $k > n$ no such word exists.

(c) If every letter from A is required to appear precisely once in the word, then the length of the word is n and all the letters have to be distinct. This is a special case of (b) where $k = n$ and there are $n!$ such words. □

Words where letters are not allowed to repeat are called *permutations of sets*. Words where letters can appear more than once are other kind of permutations — permutations of multisets — and we shall consider them in a separate section.

Definition 1.4 A *permutation of a set A* is a word over A where every letter from the alphabet appears precisely once in the word. A *k-permutation of a set A*, where $k \leqslant |A|$, is a word over A of length k where each letter from the alphabet is allowed to appear at most once (and therefore, all the letters in the word are distinct).

At the end, we apply counting techniques discussed above to determine the number of all the subsets of a finite set. For a set A let $\mathscr{P}(A)$ denote the *power-set of A*, that is, the set of all the subsets of A:

$$\mathscr{P}(A) = \{X : X \subseteq A\}.$$

Let $|A| = n$ and $A = \{a_1, \ldots, a_n\}$. Then every subset B of A can be represented by a string $\chi(B)$ of 0's and 1's as follows:

$$\chi(B) = p_1 \ldots p_n, \quad \text{where} \quad p_i = \begin{cases} 0, & a_i \notin B, \\ 1, & a_i \in B. \end{cases}$$

The word $\chi(B)$ is called the *characteristic vector* of B.

Example 1.5 Let $A = \{a, b, c, d, e, f\}$ and $B = \{b, d, e\}$. Then $\chi(B) = 010110$ since $a \notin B$, $b \in B$, $c \notin B$ etc. Clearly, $\chi(\varnothing) = 000000$ and $\chi(A) = 111111$:

	a	b	c	d	e	f
$\varnothing$	0	0	0	0	0	0
B	0	1	0	1	1	0
A	1	1	1	1	1	1

Theorem 1.6 *Let A be a finite set with n elements. Then $|\mathscr{P}(A)| = 2^n$.*

Proof. The mapping $\chi : \mathscr{P}(A) \to \{0,1\}^n$ that takes a subset of A onto its characteristic vector is a bijection, so $|\mathscr{P}(A)|$ and $|\{0,1\}^n|$ have the same number of elements. Therefore, $|\mathscr{P}(A)|$ equals the number of all words over $\{0,1\}$ whose length is n, so $|\mathscr{P}(A)| = 2^n$. □

The proof of Theorem 1.6 is based on an obvious but important fact we shall use on many occasions in this course:

The Bijection Principle: *Whenever there is a bijection between two sets, they have the same number of elements.*

Words over two-element alphabets will be particularly useful in the sequel. So, we give them a name: a *01-word* is a word over $\{0,1\}$.

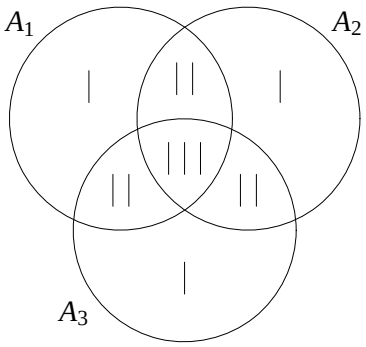
1.2 Sets

One of the most basic things one can do with a set is to count its elements. Clearly,

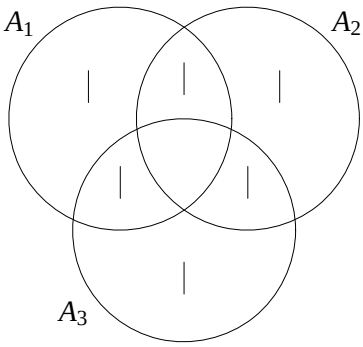
The Product Principle: *If $A_1, \ldots, A_n$ are finite sets, then*
$$|A_1 \times \ldots \times A_n| = |A_1| \cdot \ldots \cdot |A_n|.$$

It is also easy to see that

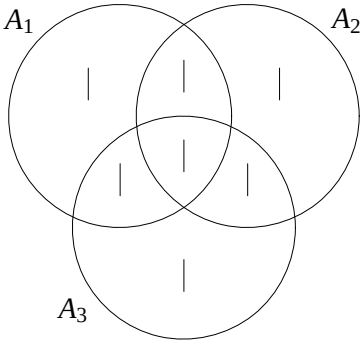
The Sum Principle: *If $A_1, \ldots, A_n$ are mutually disjoint finite sets, then*
$$|A_1 \cup \ldots \cup A_n| = |A_1| + \ldots + |A_n|.$$



$$|A_1| + |A_2| + |A_3|$$



$$|A_1| + |A_2| + |A_3| - \\ -|A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3|$$



$$|A_1| + |A_2| + |A_3| - \\ -|A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + \\ +|A_1 \cap A_2 \cap A_3|$$

Figure 1.1: The cardinality of $A_1 \cup A_2 \cup A_3$

But, what happens if $A_1, \ldots, A_n$ are not mutually disjoint? In case of $n = 2$ we know from the elementary school that

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|,$$

and it is also easy to see that in case $n = 3$ (see Fig. 1.1):

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - \\ -|A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$

Theorem 1.7 (The Principle of Inclusion-Exclusion) *Let $A_1, \ldots, A_n$ be finite sets. Then*

$$\begin{aligned}|A_1 \cup \ldots \cup A_n| = {} & |A_1| + \ldots + |A_n| \\ & - |A_1 \cap A_2| - |A_1 \cap A_3| - \ldots - |A_{n-1} \cap A_n| \\ & + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \ldots + |A_{n-2} \cap A_{n-1} \cap A_n| \\ & - \ldots \\ & + (-1)^{n-1} |A_1 \cap A_2 \cap \ldots \cap A_n|\end{aligned}$$

Proof. The proof is by induction on n. In case $n = 1$ the formula is trivial and we have already seen that the formula is true in case $n = 2$ or $n = 3$. Therefore, assume that the formula is true in case of n finite sets and let us consider the untion of $n+1$ finite sets. Using the formula for the cardinality of the union of two sets:

$$\begin{aligned}|A_1 \cup \ldots \cup A_n \cup A_{n+1}| &= |(A_1 \cup \ldots \cup A_n) \cup A_{n+1}| \\ &= |A_1 \cup \ldots \cup A_n| + |A_{n+1}| - |(A_1 \cup \ldots \cup A_n) \cap A_{n+1}| \\ &= |A_1 \cup \ldots \cup A_n| + |A_{n+1}| - |(A_1 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1})|.\end{aligned}$$

the proof follows straightforwardly by applying the induction hypothesis twice. The calculations are given in Fig. 1.2. □

Corollary 1.8 *Let $A_1, \ldots, A_n$ be finite sets such that $|A_{i_1} \cap \ldots \cap A_{i_k}| = |A_{j_1} \cap \ldots \cap A_{j_k}|$ whenever $i_1, \ldots, i_k$ are k distinct indices and $j_1, \ldots, j_k$ are k distinct indices, $k \in \{1, \ldots, n\}$. Then*

$$\begin{aligned}|A_1 \cup \ldots \cup A_n| = {} & \binom{n}{1} |A_1| - \binom{n}{2} |A_1 \cap A_2| + \binom{n}{3} |A_1 \cap A_2 \cap A_3| - \ldots \\ & + (-1)^{n-1} \binom{n}{n} |A_1 \cap A_2 \cap \ldots \cap A_n|\end{aligned}$$

1.3 Subsets

For historical reasons, a k-element subset of an n-element set is called a *k-combination* of a set. The number of k-combinations of an n-element set is denoted by

$$\binom{n}{k} \qquad \text{[read: “}n\text{ choose }k\text{”]}.$$

The pronounciation comes from the fact that this is the number of ways to choose k objects from a pool of n identical objects.

$$
\begin{aligned}
|A_1 \cup \ldots \cup A_n \cup A_{n+1}| &= |A_1 \cup \ldots \cup A_n| + |A_{n+1}| - |(A_1 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1})| \\
&= \Big(|A_1| + \ldots + |A_n| - |A_1 \cap A_2| - |A_1 \cap A_3| - \ldots - |A_{n-1} \cap A_n| + |A_1 \cap A_2 \cap A_3| + \\
&\quad |A_1 \cap A_2 \cap A_4| + \ldots + |A_{n-2} \cap A_{n-1} \cap A_n| - \ldots + (-1)^{n-1} |A_1 \cap A_2 \cap \ldots \cap A_n| \Big) \\
&\quad + |A_{n+1}| - \Big(|A_1 \cap A_{n+1}| + \ldots + |A_n \cap A_{n+1}| - |A_1 \cap A_2 \cap A_{n+1}| - |A_1 \cap A_3 \cap A_{n+1}| - \ldots \\
&\quad \ldots - |A_{n-1} \cap A_n \cap A_{n+1}| + \ldots + (-1)^{n-1} |A_1 \cap A_2 \cap \ldots \cap A_n \cap A_{n+1}| \Big) \\
&= |A_1| + \ldots + |A_n| + |A_{n+1}| - |A_1 \cap A_2| - |A_1 \cap A_3| - \ldots - |A_n \cap A_{n+1}| + \\
&\quad + |A_1 \cap A_2 \cap A_3| + \ldots + |A_{n-1} \cap A_n \cap A_{n+1}| + \ldots + (-1)^n |A_1 \cap A_2 \cap \ldots \cap A_n \cap A_{n+1}|.
\end{aligned}
$$

Figure 1.2: The calculations from the proof of Theorem 1.7

Theorem 1.9 *Let $n,k \geqslant 0$. If $n \geqslant k$ then $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Otherwise, $\binom{n}{k} = 0$.*

Proof. Let $n \geqslant k$. Although sets seem to be simpler than words due to the lack of structure, ordered structures (words in this case) are always easier to count. A k-element set gives rise to $k!$ different words, e.g.

$$\begin{array}{rcccc} \{a,b,c,d\} \rightarrow & abcd & bacd & cabd & dabc \\ & abdc & badc & cadb & dacb \\ & acbd & bcad & cbad & dbac \\ & acdb & bcda & cbda & dbca \\ & adbc & bdac & cdab & dcab \\ & adcb & bdca & cdba & dcba \end{array}$$

whence immediately follows

$$\boxed{\begin{array}{l}\text{Number of} \\ k\text{-element sets}\end{array}} \quad = \quad \frac{1}{k!} \cdot \boxed{\begin{array}{l}\text{Number of} \\ k\text{-permutations}\end{array}}$$

Since the number of k-permutations of an n-element set is $\frac{n!}{(n-k)!}$, we finaly obtain that

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

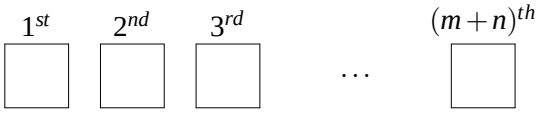
On the other hand, if $k > n$ then trivially $\binom{n}{k} = 0$ since an n-element set cannot have a subset with more than n elements. □

Problem 1.10 *How many 01-words of length $m+n$ are there if they are required to have precisely m zeros and precisely n ones?*

Solution. Consider a set $A = \{a_1, a_2, \ldots, a_{m+n}\}$ with $m+n$ elements. Then each 01-word of length $m+n$ with m zeros and n ones corresponds to an n-element subset of A. Therefore, the number of such 01-words equals the number of n-element subsets of A, which is

$$\binom{m+n}{n}.$$

Here is the other way to see this. Consider a string of $m+n$ empty boxes which are to be filled by m zeros and n ones:



We can choose m boxes in which to write zeros in $\binom{m+n}{m}$ ways. Then the remaining n boxes have to be filled by ones. □

Theorem 1.11 (a) $\binom{n}{k} = \binom{n}{n-k}$ *for all* $n \geqslant k \geqslant 0$*;*

(b) $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ *for all* $n \geqslant k \geqslant 1$ *(Pascal's identity).*

Proof. (a) This follows by an easy calculation:

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

Such proofs are usually called *algebraic proofs*.

Most combinatorial identities can be proven in another way: we find an appropriate collection of objects and then count the elements of the collection in two different ways. The resulting expressions have to be equal because the collection is the same. Such proofs are usually called *combinatorial proofs*.

Let us provide a combinatorial proof of the same identity. Consider 01-words of length n with precisely k zeros. There are $\binom{n}{k}$ ways to choose k places out of n in which to write zeros, so the number of the words under consideration is $\binom{n}{k}$. On the other hand, we can first choose $n-k$ places in which to write ones in $\binom{n}{n-k}$ ways, so the number of the words under consideration is $\binom{n}{n-k}$. Therefore, $\binom{n}{k} = \binom{n}{n-k}$.

(b) The algebraic proof of the Pascal's identity is easy:

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

For the combinatorial proof, let $S = \{1, 2, \ldots, n\}$ be an n-element set and let us count k-element subsets of S. Clearly, the number of k-element subsets is $\binom{n}{k}$.

On the other hand, all k-element subsets of S split into two classes: those that contain 1, and those that do not. The number of k-element subsets of S that contain 1 is $\binom{n-1}{k-1}$ since we have to choose $k-1$ elements from an $(n-1)$-element set $S' = \{2,\ldots,n\}$. The number of k-element subsets of S that do not contain 1 is $\binom{n}{k-1}$ since now we have to choose all k elements from S'. Therefore, $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$. □

Theorem 1.12 (Newton's binomial formula) *For all $n \in \mathbb{N}_0$ we have*

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

Proof. The proof proceeds by induction on n. The first few cases are trivial:

$$(a+b)^0 = 1 = \binom{0}{0}$$
$$(a+b)^1 = a+b = \binom{1}{0}a + \binom{1}{1}b$$
$$(a+b)^2 = a^2+2ab+b^2 = \binom{2}{0}a^2 + \binom{2}{1}ab + \binom{2}{2}b^2$$

Assume that the claim is true for n and let us compute $(a+b)^{n+1}$. By the induction hypothesis:

$$(a+b)^{n+1} = (a+b)\cdot(a+b)^n = (a+b)\cdot\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k.$$

After distributing the sum and multiplying we obtain:

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

Next, we take out the first summand in the first sum and the last summand in the second sum to obtain:

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=0}^{n-1}\binom{n}{k}a^{n-k}b^{k+1} + b^{n+1}$$

and reindex the second sum, which is a standard trick:

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

Putting the two sums together we obtain:

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n} \left(\binom{n}{k} + \binom{n}{k-1} \right) a^{n-k+1} b^k + b^{n+1}.$$

Finally, we apply the Pascal's identity and wrap it up:

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

The combinatorial proof of the Newton's binomial formula is based on a simple observation. Clearly,

$$(a+b)^n = \underbrace{(a+b) \cdot (a+b) \cdot \ldots \cdot (a+b)}_{n \text{ times}}$$

so if one multiplies out and writes down the summands as words of length n (that is, without the usuall abbreviations such as $a \cdot a \cdot a = a^3$), one obtains all possible words od length n in letters a and b. For example,

$$\begin{aligned}(a+b)^4 = {} & aaaa + aaab + aaba + aabb + abaa + abab + abba + abbb \\ & + baaa + baab + baba + babb + bbaa + bbab + bbba + bbbb.\end{aligned}$$

There are $\binom{n}{k}$ words that abbreviate to $a^{n-k}b^k$ since this is the number of ways we can choose k places for b (Problem 1.10). Therefore, $a^{n-k}b^k$ appears $\binom{n}{k}$ times in the sum, whence $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$. □

Combinatorial proofs in Theorem 1.11 are just two instances of another simple but nevertheless very useful fact:

Double Counting: *If the same set is counted in two different ways, the answers are the same.*

We use it again in the proof of the following theorem:

Theorem 1.13 $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n$.

Proof. For the algebraic proof, just note that

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

by the Newton's binomial formula. The combinatorial proof is also not very complicated. Let A be an arbitraty n-element set and let us count the number of subsets of A. According to Theorem 1.6 this number is 2^n. On the other hand, let us split $\mathscr{P}(A)$ into disjoint collections $\mathscr{S}_0$, $\mathscr{S}_1$, ..., $\mathscr{S}_n$ so that $\mathscr{S}_k$ contains all k-element subsets of A. Clearly

$$|\mathscr{P}(A)| = |\mathscr{S}_0| + |\mathscr{S}_1| + \ldots + |\mathscr{S}_n|.$$

But, $|\mathscr{S}_k| = \binom{n}{k}$ according to Theorem 1.9. This concludes the proof. □

1.4 Multisets

Two sets are equal if their elements are the same, or more precisely:

$$A = B \quad \text{if and only if} \quad \forall x (x \in A \Leftrightarrow x \in B).$$

As a consequence, $\{b,a,n,a,n,a\} = \{a,b,n\}$. We usually say that "in a set one can omit repeating elements". But what if we *wish* to put several copies of an object into a set? Well, we have to invent a new type of mathematical object.

Definition 1.14 Let $A = \{a_1, a_2, \ldots, a_n\}$ be a finite set. A *multiset over A* is any mapping $\alpha : A \to \mathbb{N}_0^\infty$.

The idea behind this definition is simple: $\alpha(a_k)$ tells us how many copies of a_k we have in the multiset α. This is why α is sometimes called the *multiplicity function*, and $\alpha(a_k)$ is the *multiplicity* of a_k. In particular, $\alpha(a_k) = 0$ means that a_k does not belong to the multiset, while $\alpha(a_k) = \infty$ means that we have an unlimited supply of copies of a_k.

A multiset $\alpha : A \to \mathbb{N}_0^\infty$ can be compactly represented as

$$\alpha = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ m_1 & m_2 & \ldots & m_n \end{pmatrix}$$

or, even more conveniently, as

$$\{m_1 \cdot a_1, m_2 \cdot a_2, \ldots, m_n \cdot a_n\},$$

where $m_j = \alpha(a_j)$, $j \in \{1, 2, \ldots, n\}$.

Definition 1.15 A multiset $\alpha = \{m_1 \cdot a_1, m_2 \cdot a_2, \ldots, m_n \cdot a_n\}$ is *empty* if $m_1 = \ldots = m_n = 0$. The multiset α is *finite* if $m_1, \ldots, m_n < \infty$. The number of elements of α is denoted by $|\alpha|$ and we define it by $|\alpha| = \sum_{a \in A} \alpha(a)$.

A multiset $\alpha = \{m_1 \cdot a_1, m_2 \cdot a_2, \ldots, m_n \cdot a_n\}$ is a *submultiset* of a multiset $\beta = \{k_1 \cdot a_1, k_2 \cdot a_2, \ldots, k_n \cdot a_n\}$ if $m_j \leqslant k_j$ for all j.

Example 1.16 Let $A = \{a, b, c\}$. Then $\alpha = \{3 \cdot a, 2 \cdot b, 1 \cdot c\}$ and $\beta = \{0 \cdot a, 5 \cdot b, \infty \cdot c\}$ are two multisets over A. Clearly α is a finite multiset with 6 elements, while β is infinite and $|\beta| = \infty$. Both α and β are submultisets of $\gamma = \{\infty \cdot a, 5 \cdot b, \infty \cdot c\}$. Also, β is a submultiset of $\delta = \{1 \cdot a, \infty \cdot b, \infty \cdot c\}$, while α is not.

A *word over a multiset* $\alpha = \{m_1 \cdot a_1, m_2 \cdot a_2, \ldots, m_n \cdot a_n\}$ is any word w over $A = \{a_1, \ldots, a_n\}$ such that $|w|_{a_j} \leqslant m_j$ for all j.

Example 1.17 Let $\alpha = \{3 \cdot a, 2 \cdot b, 2 \cdot n\}$. The following are some words over α: *banana*, *abba*, *aa*, but *abbba* is not. As another example, take $\beta = \{1 \cdot a, \infty \cdot b\}$. Then all these are words over β: *a*, *ab*, *abb*, *abbb*, and so on.

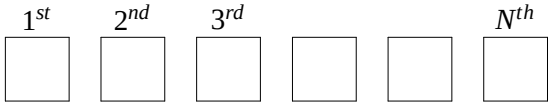
Problem 1.18 *Let $\alpha = \{m_1 \cdot a_1, m_2 \cdot a_2, \ldots, m_n \cdot a_n\}$ be a multiset and let $k \in \mathbb{N}_0$ be arbitrary.*

(a) Suppose $m_1 = m_2 = \ldots = m_n = \infty$. How many words with k letters over α are there?

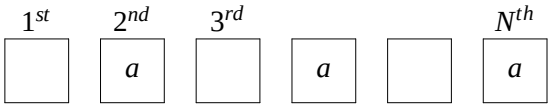
(b) Suppose α is finite. How many words w over α have the property that $|w|_{a_j} = m_j$ for all j?

Solution. (a) Since each letter comes in more than sufficiently many copies, it turns out that the number of such words is n^k. Compare with Problem 1.3 (a).

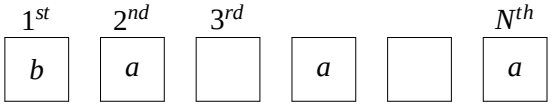
(b) Let $N = |\alpha| = m_1 + \ldots + m_n$. Then the words we are interested are of length N:



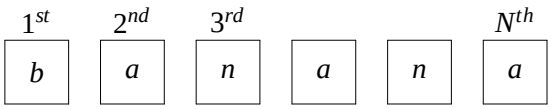
and each letter a_j occurs precisely m_j times. Let us now distribute the letters from α. Out of N free places we can choose m_1 places to put the copies of a_1 in $\binom{N}{m_1}$ ways:



Out of $N-m_1$ remaining free places we can choose m_2 places to put the copies of a_2 in $\binom{N-m_1}{m_2}$ ways:



Out of $N-m_1-m_2$ remaining free places we can choose m_3 places to put the copies of a_3 in $\binom{N-m_1-m_2}{m_3}$ ways, and so on. At the end, out of $N-m_1-m_2-\ldots-m_{n-1}$ remaining free places we can choose m_n places to put the copies of a_n in $\binom{N-m_1-m_2-\ldots-m_{n-1}}{m_n}$ ways:



Therefore, the number fo words we are interested in is given by

$$\binom{N}{m_1}\cdot\binom{N-m_1}{m_2}\cdot\binom{N-m_1-m_2}{m_3}\cdot\ldots\cdot\binom{N-m_1-m_2-\ldots-m_{n-1}}{m_n}=$$

$$=\frac{N!}{m_1!(N-m_1)!}\cdot\frac{(N-m_1)!}{m_2!(N-m_1-m_2)!}\cdot\ldots\cdot\frac{(N-m_1-m_2-\ldots-m_{n-1})!}{m_n!(N-m_1-m_2-\ldots-m_n)!}=$$

$$=\frac{N!}{m_1!\cdot m_2!\cdot\ldots\cdot m_n!},$$

where at the end we use the fact that $N=m_1+m_2+\ldots+m_n$. □

A *permutation of a finite multiset* $\alpha=\{m_1\cdot a_1,m_2\cdot a_2,\ldots,m_n\cdot a_n\}$ is any word w over α such that $|w|_{a_j}=m_j$ for all j. As we have just seen, the number of permutations over a finite multiset α is

$$\binom{N}{m_1,m_2,\ldots,m_n}=\frac{N!}{m_1!\cdot m_2!\cdot\ldots\cdot m_n!}$$

where $N=m_1+m_2+\ldots+m_n$. Finding the number of k-letter words for arbitrary k and over an arbitrary multiset is a *terribly* complicated problem and shall not be discussed here.

We shall now prove an analogon on the Newton's binomial formula in case a sum of more than two expressions is raised to a certain power.

Theorem 1.19 (Multinomial formula) *For all $n \geqslant 0$ we have*

$$(a_1 + a_2 + \ldots + a_k)^n = \sum_{\substack{l_1,l_2,\ldots,l_k \in \mathbb{N}_0 \\ l_1+l_2+\ldots+l_k=n}} \binom{n}{l_1, l_2, \ldots, l_k} a_1^{l_1} a_2^{l_2} \ldots a_k^{l_k}$$

Proof. The proof proceeds by induction on k. In case $k = 2$ this is just the Newton's binomial formula given in Theorem 1.12, see Homework 1.9. Suppose the theorem holds whenever there are less than k summands whose sum we wish to raise to the n-th power and consider the case with k summands. Then by the Newton's binomial formula

$$(a_1 + a_2 + \ldots + a_k)^n = (a_1 + (a_2 + \ldots + a_k))^n = \sum_{l_1=0}^{n} \binom{n}{l_1} a_1^{l_1} (a_2 + \ldots + a_k)^{n-l_1}.$$

The induction hypothesis now yields

$$\begin{aligned}
(a_1 + a_2 + \ldots + a_k)^n &= \sum_{l_1=0}^{n} \binom{n}{l_1} a_1^{l_1} \sum_{\substack{l_2,\ldots,l_k \in \mathbb{N}_0 \\ l_2+\ldots+l_k=n-l_1}} \binom{n-l_1}{l_2, \ldots, l_k} a_2^{l_2} \ldots a_k^{l_k} \\
&= \sum_{l_1=0}^{n} \sum_{\substack{l_2,\ldots,l_k \in \mathbb{N}_0 \\ l_2+\ldots+l_k=n-l_1}} \binom{n}{l_1} \binom{n-l_1}{l_2, \ldots, l_k} a_1^{l_1} a_2^{l_2} \ldots a_k^{l_k} \\
&= \sum_{l_1=0}^{n} \sum_{\substack{l_2,\ldots,l_k \in \mathbb{N}_0 \\ l_2+\ldots+l_k=n-l_1}} \binom{n}{l_1, l_2, \ldots, l_k} a_1^{l_1} a_2^{l_2} \ldots a_k^{l_k} \\
&= \sum_{\substack{l_1,l_2,\ldots,l_k \in \mathbb{N}_0 \\ l_1+l_2+\ldots+l_k=n}} \binom{n}{l_1, l_2, \ldots, l_k} a_1^{l_1} a_2^{l_2} \ldots a_k^{l_k}.
\end{aligned}$$

The combinatorial proof is analogous to the combinatorial proof of Theorem 1.12. □

Problem 1.20 *Let $\alpha = \{\infty \cdot a_1, \infty \cdot a_2, \ldots, \infty \cdot a_n\}$ be a multiset and let $k \in \mathbb{N}_0$ be arbitrary. How many k-element submultisets does α have?*

Solution. If $\beta = \{x_1 \cdot a_1, x_2 \cdot a_2, \ldots, x_n \cdot a_n\}$ is a k-element submultiset of α, then $x_1 + x_2 + \ldots + x_n = k$. Because α has an infinite supply of each of its letters, one easily comes to the following conclusion:

Number of k-element subsets of α	=	Number of solutions of $x_1 + x_2 + \ldots + x_n = k$ in $\mathbb{N}_0$

So, we have reduced the problem to counting nonnegative integer solutions of an equation in n unknowns. Although not at all straightforward, this problem is rather easy to solve. Let

$$\mathscr{S} = \{(x_1, x_2, \ldots, x_n) \in (\mathbb{N}_0)^n : x_1 + x_2 + \ldots + x_n = k\}$$

be the set of all the solutions of the above equation in n unknowns and let

$$\mathscr{W} = \{w \in \{0,1\}^{k+n-1} : |w|_0 = k \text{ and } |w|_1 = n-1\}$$

be the set of all 01-words of length $k+n-1$ with precisely k zeros and $n-1$ ones. Now define $\varphi : \mathscr{S} \to \mathscr{W}$ as follows:

$$\varphi(x_1, x_2, \ldots, x_n) = \underbrace{00\ldots0}_{x_1}\ 1\ \underbrace{00\ldots0}_{x_2}\ 1\ \ldots\ 1\ \underbrace{00\ldots0}_{x_n}.$$

It is easy to see that φ is well defined and bijective. Therefore, $|\mathscr{S}| = |\mathscr{W}|$, and we know from Problem 1.10 that $|\mathscr{W}| = \binom{k+n-1}{k}$. This is at the same time the number of k-element subsets of α. □

A *k-combination of a finite multiset* α is any k-element subset of α. It is again *terribly* complicated to find a number of k-combinations of an arbitrary multiset, but as we have just seen, if $\alpha = \{\infty \cdot a_1, \infty \cdot a_2, \ldots, \infty \cdot a_n\}$, the number of k-combinations is given by $\binom{k+n-1}{k}$.

Homework

1.1. For a real number x, by $\lfloor x \rfloor$ we denote the greatest integer $\leqslant x$. E.g, $\lfloor 1.99 \rfloor = 1$, $\lfloor 4 \rfloor = 4$, $\lfloor 0.65 \rfloor = 0$, while $\lfloor -1.02 \rfloor = -2$.

Let n be an integer and p a prime. Show that the greatest k such that $p^k \mid n!$ is given by

$$k = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \ldots$$

The number 1000! ends with a lot of zeros. How many?

1.2. Show that χ in proof of Theorem 1.6 is a bijection.

1.3. Let A be a set of all 01-words w of length 2005 with the property that $|w|_0 = |w|_1 + 1$, and let B be a set of all 01-words w of length 2005 with the property that $|w|_1 = |w|_0 + 1$. Show that $|A| = |B|$. (Hint: use the Bijection Principle.)

1.4. For $n \in \mathbb{N}$, let $\tau(n)$ denote the number of positive divisors of n. E.g, $\tau(12) = 6$ since 1, 2, 3, 4, 6 and 12 are all positive divisors of 12. Let $n = p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_s^{k_s}$ be the factorisation of n, where $1 < p_1 < p_2 < \ldots < p_s$ are primes. Prove that

$$\tau(n) = (1+k_1)(1+k_2)\ldots(1+k_s).$$

(Hint: note that if $m \mid n$ then $m = p_1^{l_1} \cdot p_2^{l_2} \cdot \ldots \cdot p_s^{l_s}$ where $0 \leqslant l_i \leqslant k_i$ for all i.)

1.5. Prove Corollary 1.8.

1.6. Show that φ defined in the solution to Problem 1.20 is a bijection.

†**1.7.** What do you think, how do "usual" sets fit into the theory of multisets?

†**1.8.** Define the notion of union and intersection for multisets. (Note that there are several possibilities; choose any one you like). Pick a few of your favourite set-theory identities such as

$$\begin{array}{ll} \alpha \cap \alpha = \alpha & \alpha \cup \alpha = \alpha \\ \alpha \cap \varnothing = \varnothing & \alpha \cup \varnothing = \alpha \\ \alpha \cap \beta = \beta \cap \alpha & \alpha \cup \beta = \beta \cup \alpha \\ (\alpha \cap \beta) \cap \gamma = \alpha \cap (\beta \cap \gamma) & (\alpha \cup \beta) \cup \gamma = \alpha \cup (\beta \cup \gamma) \\ (\alpha \cap \beta) \cup \gamma = (\alpha \cup \gamma) \cap (\beta \cup \gamma) & (\alpha \cup \beta) \cap \gamma = (\alpha \cap \gamma) \cup (\beta \cap \gamma) \end{array}$$

and show that they hold for operations you have defined.

1.9. (a) Explain the relationship between $\binom{n}{k}$ and $\binom{n}{k, n-k}$.

(b) Show that $\binom{n}{k, n-k} = \binom{n-1}{k-1, n-k} + \binom{n-1}{k, n-k-1}$ (Hint: this is the Pascal's identity in disguise.)

1.10. Let $m_1, \ldots, m_n \in \mathbb{N}$ be positive integers and let $N = m_1 + \ldots + m_n$. Show that

$$\binom{N}{m_1, m_2, \ldots, m_n} = \binom{N-1}{m_1 - 1, m_2, \ldots, m_n} + \binom{N-1}{m_1, m_2 - 1, \ldots, m_n} + \ldots$$
$$\ldots + \binom{N-1}{m_1, m_2, \ldots, m_n - 1}.$$

1.11. Provide a combinatorial proof of Theorem 1.19.

Exercises

1.12. How much memory can address a processor whose address bus is 32 bits wide?

1.13. FORTRAN IV, being one of the oldest programming languages, had many limitations. One of them concerned identifiers (words used to name variables and procedures). An identifier in FORTRAN IV consists of at most 6 symbols, where each symbol is a figure (0, 1, ..., 9) or an uppercase letter of the English alphabet (A, B, ..., Z), with the exception that the first symbol is obliged to be a letter. How many different identifiers can one declare in FORTRAN IV?

1.14. Two rooks on a chess board are said to be independent if they do not attack each other. In how many different ways can one arrange $n \geqslant 1$ independent identical rooks onto an $n \times n$ chess board?

1.15. In how many different ways can one arrange $k \geqslant 1$ independent identical rooks onto an $n \times m$ chess board, where $n, m \geqslant k$?

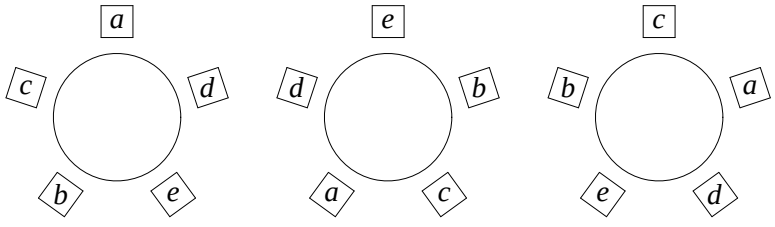
1.16. In how many ways can n students form a queue in front of a mensa so that students A and B

(a) are next to each other in the queue?

(b) are *not* next to each other in the queue?

†**1.17.** In how many ways can n boys $B_1, \ldots, B_n$ and n girls $G_1, \ldots, G_n$ form a queue in front of a mensa so that B_1 is next to G_1 in the queue, B_2 is next to G_2 in the queue, ..., B_n is next to G_n in the queue?

1.18. The round table has entered combinatorial practice at the time of King Arthur and his Knights of the Round Table and has remained an important combinatorial object ever since. Since there is no throne, the trick with the round table is that two arrangements are indistinguishable if it is possible get one of them by rotating the other. For example, the following three arrangements are indistinguishable:



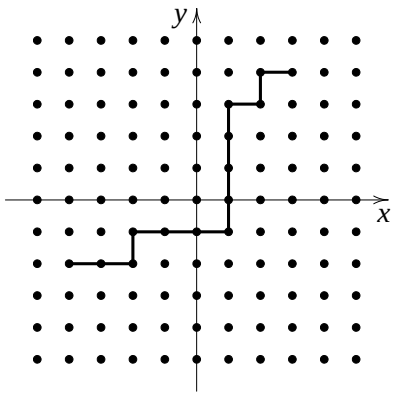
In how many ways can n people be seated around a round table with n

seats?

1.19. The *integer grid* consists of all points in the plane with integer coordinates, which we refer to as *integer points*.

An *increasing path* in the integer grid is a sequence of integer points (x_1,y_1), (x_2,y_2), …, (x_k,y_k) such that for each $i \in \{1,\ldots,k-1\}$ we have:

- either $x_{i+1} = x_i + 1$ and $y_{i+1} = y_i$,
- or $x_{i+1} = x_i$ and $y_{i+1} = y_i + 1$.



Find the number of increasing paths in the integer grid that start at $(0,0)$ and end at (p,q), where $p,q \in \mathbb{N}$.

1.20. Show that

(a) $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$ for all $n \geqslant k \geqslant 1$;

(b) $\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k}$ for all $n \geqslant m \geqslant k \geqslant 0$;

(c) $\binom{n}{0}\binom{m}{k} + \binom{n}{1}\binom{m}{k-1} + \binom{n}{2}\binom{m}{k-2} + \ldots + \binom{n}{k}\binom{m}{0} = \binom{n+m}{k}$
for all $n,m \geqslant k \geqslant 0$.

(d) $\binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \ldots + \binom{n}{n}^2 = \binom{2n}{n}$ for all $n \geqslant 0$.

(e) $\binom{k}{0} + \binom{k+1}{1} + \binom{k+2}{2} + \ldots + \binom{k+j}{j} = \binom{k+j+1}{j}$
for all $k, j \geqslant 0$. (Hint: use mathematical induction on j and (b).)

1.21. Find the number of 01-words of length $2n$ which have the following property: the number of zeroes on the first n places equals the number of zeros on the last n places.

1.22. (a) Using the fact that two points determine precisely one straingt line, find the greatest number of straignt lines that can be drawn through n points in a plane.

(b) Find the greatest number of diagonals a convex polygon with n vertices can have.

(c) Let $A_1, \ldots, A_n$ be n points on a circle, $n \geqslant 4$, and draw all the line segments A_iA_j, $i \neq j$. Find the greatest possible number of intersection points of these line segments.

1.23. Find the number of integer solutions of the equation $x_1 + x_2 + \ldots + x_n = k$ in n unknowns $x_1, x_2, \ldots, x_n$ where $x_i \geqslant 1$ for all i.

1.24. Find the number of integer solutions of the inequality $x_1 + x_2 + \ldots + x_n \leqslant k$ in n unknowns $x_1, x_2, \ldots, x_n$ where $x_i \geqslant 0$ for all i. (Hint: Since $k \in \mathbb{N}_0$, this inequality is equivalent to

$$x_1 + x_2 + \ldots + x_n = 0 \quad \text{or} \quad x_1 + x_2 + \ldots + x_n = 1 \quad \text{or} \ \ldots$$
$$\ldots \ \text{or} \quad x_1 + x_2 + \ldots + x_n = k.$$

Find the number of solutions of each of these $k+1$ equations and then sum up using 1.20 (f).)

1.25. A sequence of numbers $x_1, x_2, \ldots, x_n$ is nondecreasing if $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$. Find the number of nondecreasing sequences $x_1, x_2, \ldots, x_n$ where $x_i \in \{1, \ldots, k\}$ for all i.

1.26. Show that $\displaystyle\sum_{\substack{l_1, l_2, \ldots, l_k \in \mathbb{N}_0 \\ l_1 + l_2 + \ldots + l_k = n}} \binom{n}{l_1, l_2, \ldots, l_k} = k^n$.

Chapter 2

Blocks and Cycles

The simplest way to introduce a structure onto a set is to split it into blocks. In this chapter we consider two such possibilities:

- partitions, where a set is divided into disjoint subsets, and
- permutations (again), which partition a set into cycles.

Counting partitions and permutations leads to Stirling numbers of the second and the first kind, respectively. Stirling numbers of the second kind show up more often than those of the first kind, so we shall consider last things first.

Stirling numbers of the second kind are usually denoted by S_n^k or $\left\{ {n \atop k} \right\}$, while Stirling numbers of the first kind are usually denoted by s_n^k or $\left[{n \atop k} \right]$. The notation with braces and brackets, in analogy to the binomial coefficients, was introduced in 1935 by Jovan Karamata, a famous Serbian mathematician, and promoted later by Donald Knuth. It is referred to as *Karamata notation*. Following Knuth, we verbalise $\left\{ {n \atop k} \right\}$ as "n block k" and $\left[{n \atop k} \right]$ as "n cycle k".

2.1 Partitions

A *partition* of a finite set A is every finite set $\{B_1, \ldots, B_k\}$ of subsets of A which fulfills the following:

- $B_i \neq \varnothing$ for all i,
- $B_i \cap B_j = \varnothing$ whenever $i \neq j$, and
- $B_1 \cup \ldots \cup B_k = A$.

Sets B_i are referred to as the *blocks* of the partition.

Example 2.1 Let $A = \{1,2,3,4,5,6,7\}$. Then $\big\{\{1,3,5\},\{2,6\},\{4\}\big\}$ is a partition of A into three blocks. Instead of $\big\{\{1,3,5\},\{2,6\},\{4\}\big\}$ we can also write 135|26|4, although this notation is not always convenient.

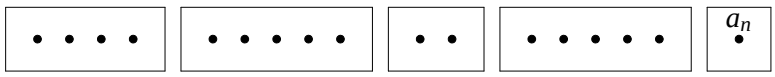
Let $\left\{ n \atop k \right\}$ denote the number of ways to partition an n-element set into k blocks. The number $\left\{ n \atop k \right\}$ is called the *Stirling number of the second kind*.

Example 2.2 For example, $\left\{ 5 \atop 3 \right\} = 25$:

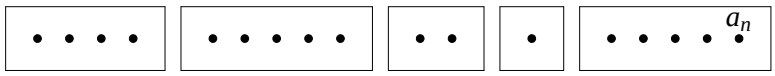
1|2|345	2|3|145	3|4|125	4|5|123	1|23|45	2|13|45	3|12|45	4|12|35
1|3|245	2|4|135	3|5|124		1|24|35	2|14|35	3|14|25	4|13|25
1|4|235	2|5|124			1|25|34	2|15|34	3|15|24	4|15|23
1|5|234							

Theorem 2.3 *Stirling numbers of the second kind fulfill the following:* $\left\{ n \atop 1 \right\} = \left\{ n \atop n \right\} = 1$ *and* $\left\{ n \atop k \right\} = \left\{ n-1 \atop k-1 \right\} + k \left\{ n-1 \atop k \right\}$.

Proof. The only way to partition a set so that in the end we get only one block is to put everything in that block, therefore, $\left\{ n \atop 1 \right\} = 1$. Similarly, the only way to partition an n-element set into n blocks is to put every element in a separate block, so $\left\{ n \atop n \right\} = 1$.



A partition where a_n is alone in its block



A partition where a_n has some company

Figure 2.1: Proof of Theorem 2.3 – two types of partitions

Now, let $\mathscr{P}_n^k$ denote the set of all partitions of a fixed n-element set, say $\{a_1,a_2,a_3,\ldots,a_n\}$, into k blocks, so that $\left\{ n \atop k \right\} = |\mathscr{P}_n^k|$. Consider a_n and note

that all partitions from $\mathscr{P}_n^k$ fall into two disjoint categories: those where a_n is the only element of its block, and those where a_n is not the only element of its block, Fig. 2.1. Let $\mathscr{S}_1 \subseteq \mathscr{P}_n^k$ be the set of all partitions where a_n is the only element of its block, and let $\mathscr{S}_2 \subseteq \mathscr{P}_n^k$ be the set of all partitions where a_n is not the only element of its block. Clearly $\mathscr{S}_1 \cap \mathscr{S}_2 = \varnothing$ and $\mathscr{S}_1 \cup \mathscr{S}_2 = \mathscr{P}_n^k$, i.e. $\{\mathscr{S}_1, \mathscr{S}_2\}$ is a partition of $\mathscr{P}_n^k$.

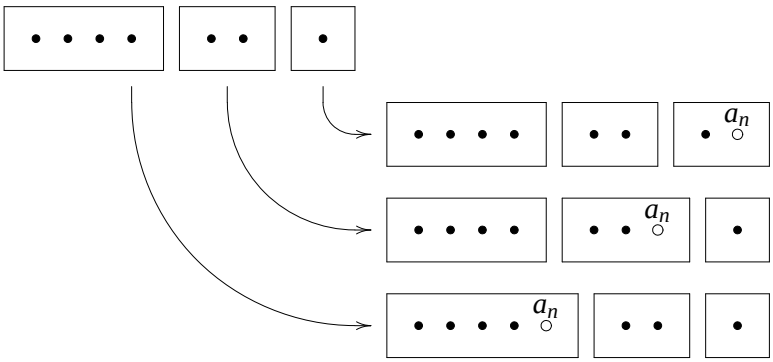


Figure 2.2: Proof of Theorem 2.3 – the second case

For each partition from $\mathscr{S}_1$ the blocks that do not contain a_n form a partition of $\{a_1, \ldots, a_{n-1}\}$ into $k-1$ blocks, so $|\mathscr{S}_1| = \left\{ {n-1 \atop k-1} \right\}$. Now, take any partition from $\mathscr{S}_2$ and remove a_n. What remains is a partition of $\{a_1, \ldots, a_{n-1}\}$ into k blocks. On the other hand, each partition of $\{a_1, \ldots, a_{n-1}\}$ into k blocks determines k different partitions of $\{a_1, \ldots, a_{n-1}, a_n\}$ into k blocks since we can put the missing a_n into each of the k blocks, Fig. 2.2. Therefore, $|\mathscr{S}_2| = k \left\{ {n-1 \atop k} \right\}$, and finally $\left\{ {n \atop k} \right\} = |\mathscr{P}_n^k| = |\mathscr{S}_1| + |\mathscr{S}_2| = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}$. □

Theorem 2.4 *For $n \geqslant k \geqslant 1$,* $\left\{ {n \atop k} \right\} = \dfrac{1}{k!} \displaystyle\sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n$.

Proof. We shall count surjective mappings from $\{1, \ldots, n\}$ onto $\{1, \ldots, k\}$ in two different ways. First, note that every mapping $f : A \to B$ determines an equivalence relation $\sim_f$ on A as follows:

$$a \sim_f b \quad \text{if and only if} \quad f(a) = f(b),$$

which is usually referred to as the *kernel of f*. This equivalence relation then determines a partition of A in the usual way: blocks in the partition are equivalence

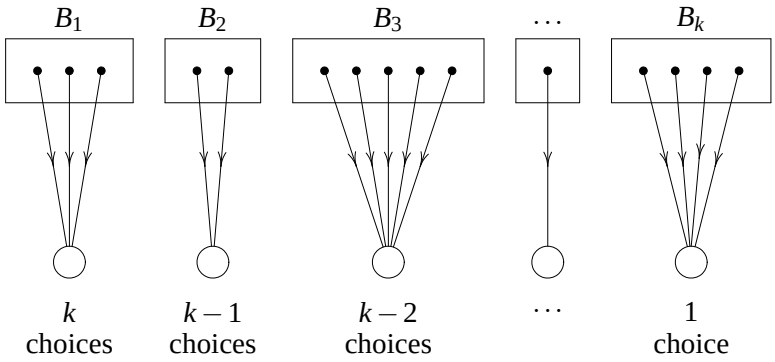


Figure 2.3: Counting surjective mappings

classes of $\sim_f$. So, for every surjective $f : \{1,\dots,n\} \to \{1,\dots,k\}$ the partition of $\{1,\dots,n\}$ that corresponds to $\sim_f$ has precisely k blocks.

Let us now take a look at a somewhat different problem: given a partition $\{B_1,\dots,B_k\}$ of $\{1,\dots,n\}$, how many surjective mappings $f : \{1,\dots,n\} \to \{1,\dots,k\}$ have the property that $\sim_f$ partitions $\{1,\dots,n\}$ into $\{B_1,\dots,B_k\}$? We can choose $f(B_1)$ in k different ways, $f(B_2)$ in $k-1$ different ways, since $f(B_i)$ and $f(B_j)$ have to be distinct whenever $i \neq j$, $f(B_3)$ in $k-2$ different ways, and so on upto $f(B_k)$ for which only one choice remains, Fig. 2.3. Therefore, given a partition $\{B_1,\dots,B_k\}$, there are $k!$ surjective mappings $f : \{1,\dots,n\} \to \{1,\dots,k\}$ such that the equivalence classes of $\sim_f$ are $\{B_1,\dots,B_k\}$. Since every surjective mapping is uniquely determined by its kernel and its values on the equivalence classes of the kernel, it follows that there are $k! \left\{ {n \atop k} \right\}$ surjective mappings from an n-element set onto a k-element set.

On the other hand, let us count surjective mappings using another approach. Clearly,

Number of surjective mappings	=	Number of all mappings	−	Number of nonsurjective mappings

The number of all mappings $\{1,\dots,n\} \to \{1,\dots,k\}$ is k^n (Homework 2.1). As for the nonsurjective mappings, let A_j denote the set of all mappings $f : \{1,\dots,n\} \to \{1,\dots,k\}$ such that $f(x) = j$ for no $x \in \{1,\dots,n\}$:

$$A_j = \{f \in \{1,\dots,n\}^{\{1,\dots,k\}} : f(x) \neq j \text{ for all } x\},$$

$j \in \{1,\ldots,k\}$. Then $A_1 \cup \ldots \cup A_k$ is the set of *all* nonsurjective mappings $\{1,\ldots,n\} \to \{1,\ldots,k\}$, so

$$\boxed{\text{Number of nonsurjective mappings}} \quad = \quad |A_1 \cup \ldots \cup A_k|.$$

According to the special case of Principle of Inclusion-Exclusion, Corollary 1.8,

$$|A_1 \cup \ldots \cup A_k| = \sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} |A_1 \cap \ldots \cap A_j|.$$

But $A_1 \cap \ldots \cap A_j$ is the set of all mappings $\{1,\ldots,n\} \to \{j+1,\ldots,k\}$, so Homework 2.1 yields $|A_1 \cap \ldots \cap A_j| = (k-j)^n$. Putting it all together, we get

$$\boxed{\text{Number of surjective mappings}} \quad = \quad k^n - \sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} (k-j)^n = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n.$$

Therefore, $k! \left\{ {n \atop k} \right\} = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n$ which concludes the proof. □

2.2 Permutations

Recall that a permutation of a finite A set is a word $a_1 a_2 \ldots a_n$ over A where every letter from A appears precisely once (and hence $n = |A|$). However, note that a word $a_1 a_2 \ldots a_n$ over A is just a mapping $f : \{1,\ldots,n\} \to A$ given by

$$f = \begin{pmatrix} 1 & 2 & \ldots & n \\ a_1 & a_2 & \ldots & a_n \end{pmatrix}$$

Therefore, permutations of A correspond to bijective mappings $\{1,\ldots,n\} \to A$, which in case of $A = \{1,\ldots,n\}$ leads to the following important observation:

Observation. *A permutation of* $\{1,\ldots,n\}$ *is any bijective mapping of* $\{1,\ldots,n\}$ *onto itself.*

This simple insight allows us to draw permutations: take any permutation $f : \{1,\ldots,n\} \to \{1,\ldots,n\}$, take n points in the plane and draw an arrow from i to j if $f(i) = j$. For example, Fig. 2.4 depicts the permutation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 8 & 9 & 12 & 7 & 5 & 11 & 10 & 2 & 1 & 3 & 6 & 4 \end{pmatrix}.$$

We see that the permutation splits into *cycles*. The following theorem claims that this is a general phenomenon.

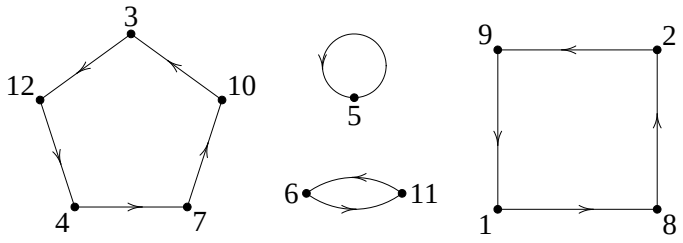


Figure 2.4: A permutation of $\{1,2,\ldots,12\}$

Theorem 2.5 *Every permutation of a finite set splits into cycles.*

Representing a permutation via its cycles is very popular and extremely useful. It is called the *cycle representation* of the permutation. To write a cycle representation of a permutation is easy – just list the elements of each cycle. For example

$$(7 \quad 10 \quad 3 \quad 12 \quad 4) \quad (11 \quad 6) \quad (8 \quad 2 \quad 9 \quad 1) \quad (5)$$

is a cycle representation of the permutation in Fig. 2.4. Since the order of cycles in the cycle representation is not significant and since $(8 \quad 2 \quad 9 \quad 1)$, $(2 \quad 9 \quad 1 \quad 8)$, $(9 \quad 1 \quad 8 \quad 2)$ and $(1 \quad 8 \quad 2 \quad 9)$ are equivalent representations of one and the same cycle, the cycle representation of a permutation is not unique. So, all these are valid cycle representations of the permutation in Fig. 2.4:

$$\begin{array}{l}(7 \quad 10 \quad 3 \quad 12 \quad 4) \quad (11 \quad 6) \quad (8 \quad 2 \quad 9 \quad 1) \quad (5),\\ (2 \quad 9 \quad 1 \quad 8) \quad (5) \quad (6 \quad 11) \quad (4 \quad 7 \quad 10 \quad 3 \quad 12),\\ (5) \quad (9 \quad 1 \quad 8 \quad 2) \quad (11 \quad 6) \quad (7 \quad 10 \quad 3 \quad 12 \quad 4).\end{array}$$

In order to make our lives easier, we shall introduce the *canonical cycle representation of a permutation* as follows:

- each cycle starts with the smallest element in the cycle (call it the *leading element of the cycle*);
- the cycles are arranged according to the increasing leading elements.

So, the canonical cycle representation of the permutation in Fig. 2.4 is:

$$(1 \quad 8 \quad 2 \quad 9) \quad (3 \quad 12 \quad 4 \quad 7 \quad 10) \quad (5) \quad (6 \quad 11).$$

The *Stirling number of the first kind*, $\begin{bmatrix} n \\ k \end{bmatrix}$, is the number of permutations of $\{1,\ldots,n\}$ with precisely k cycles.

Example 2.6 $\begin{bmatrix} 5 \\ 3 \end{bmatrix} = 35$ since

$$
\begin{array}{llll}
(1)(2)(345) & (2)(3)(145) & (3)(4)(125) & (4)(5)(123) \\
(1)(3)(245) & (2)(4)(135) & (3)(5)(124) & \\
(1)(4)(235) & (2)(5)(124) & & \\
(1)(5)(234) & & & \\
(1)(2)(354) & (2)(3)(154) & (3)(4)(152) & (4)(5)(132) \\
(1)(3)(254) & (2)(4)(153) & (3)(5)(142) & \\
(1)(4)(253) & (2)(5)(142) & & \\
(1)(5)(243) & & & \\
(1)(23)(45) & (2)(13)(45) & (3)(12)(45) & (4)(12)(35) \\
(1)(24)(35) & (2)(14)(35) & (3)(14)(25) & (4)(13)(25) \\
(1)(25)(34) & (2)(15)(34) & (3)(15)(24) & (4)(15)(23)
\end{array}
$$

Theorem 2.7 *Stirling numbers of the first kind fulfill the following:* $\begin{bmatrix} n \\ 1 \end{bmatrix} = (n-1)!$, $\begin{bmatrix} n \\ n \end{bmatrix} = 1$, *and* $\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix}$.

Proof. In order to show that $\begin{bmatrix} n \\ 1 \end{bmatrix} = (n-1)!$ it suffices to note that $\begin{bmatrix} n \\ 1 \end{bmatrix}$ is the number of permutations of $\{1,\ldots,n\}$ with presicely one cycle. In other words, $\begin{bmatrix} n \\ 1 \end{bmatrix}$ counts the number of ways to arrange n people around a round table with n seats, which is $(n-1)!$. To show that $\begin{bmatrix} n \\ n \end{bmatrix} = 1$ is even more easy: there is exactly one permutation of $\{1,\ldots,n\}$ that maps each $x \in \{1,\ldots,n\}$ onto itself, namely, the identity.

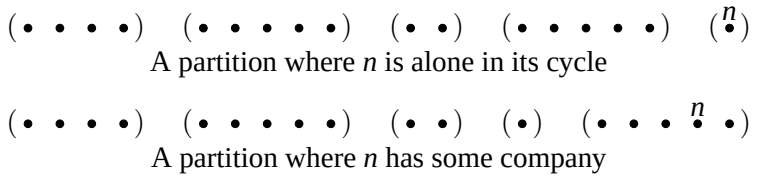


Figure 2.5: Proof of Theorem 2.7 – two types of permutations

Now, let $\mathscr{P}_n^k$ denote the set of all permutations of $\{1,2,\ldots,n\}$, into k cycles, so that $\begin{bmatrix} n \\ k \end{bmatrix} = |\mathscr{P}_n^k|$. Consider n and note that all permutations from $\mathscr{P}_n^k$ fall into two disjoint categories: those where n is the only element in its cycle, and those where n is not the only element in the cycle, Fig. 2.5. Let $\mathscr{S}_1 \subseteq \mathscr{P}_n^k$ be the set

of all permutations where n is the only element in its cycle, and let $\mathscr{S}_2 \subseteq \mathscr{P}_n^k$ be the set of all permutations where n is not the only element in the cycle. Clearly $\mathscr{S}_1 \cap \mathscr{S}_2 = \varnothing$ and $\mathscr{S}_1 \cup \mathscr{S}_2 = \mathscr{P}_n^k$, i.e. $\{\mathscr{S}_1, \mathscr{S}_2\}$ is a partition of $\mathscr{P}_n^k$.

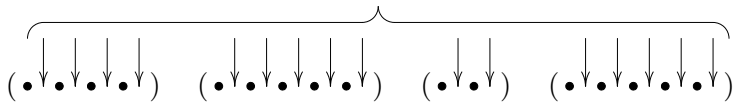


Figure 2.6: Proof of Theorem 2.7 – the second case

For each permutation from $\mathscr{S}_1$ the cycles that do not contain n form a permutation of $\{1,\ldots,n-1\}$ with $k-1$ cycles, so $|\mathscr{S}_1| = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}$. Now, take any permutation from $\mathscr{S}_2$ and remove n. What remains is a permutation of $\{1,\ldots,n-1\}$ with k cycles. On the other hand, each permutation of $\{1,\ldots,n-1\}$ with k cycles determines $n-1$ different permutations of $\{1,\ldots,n-1,n\}$ with k cycles since we can put the missing n after every element of every cycle to produce a new permutation, Fig. 2.6. Therefore, $|\mathscr{S}_2| = (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix}$, and finally $\begin{bmatrix} n \\ k \end{bmatrix} = |\mathscr{P}_n^k| = |\mathscr{S}_1| + |\mathscr{S}_2| = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix}$. □

We have now seen two important representations of permutations of $\{1,\ldots,n\}$: representation by words where each letter in the alphabet appear exactly once, and representation by bijective functions $\{1,\ldots,n\} \to \{1,\ldots,n\}$, and from this point on we shall not distinguish between the two. We shall treat permutations as words or as bijective mappings, whichever is more convenient at that point, since one can easily switch from one representation to another, e.g.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 7 & 1 & 2 & 8 & 6 & 4 & 5 \end{pmatrix} \longleftrightarrow 37128645.$$

Definition 2.8 Let $f = a_1 a_2 \ldots a_n$ be a permutation of $\{1,2,\ldots,n\}$. An *inversion of the permutation* f is a pair (a_i, a_j) such that $i < j$ and $a_i > a_j$. The number of inversions of f is denoted by $\mathrm{inv}(f)$. A permutation is called *even* or *odd* according as $\mathrm{inv}(f)$ is an even or an odd integer.

Example 2.9 Let $f = 37128645$. Then the inversions of f are $(3,1)$, $(3,2)$, $(7,1)$, $(7,2)$, $(7,6)$, $(7,4)$, $(7,5)$, $(8,6)$, $(8,4)$, $(8,5)$, $(6,4)$ and $(6,5)$, Fig. 2.7. Therefore, $\mathrm{inv}(f) = 12$ and this permutation is even.

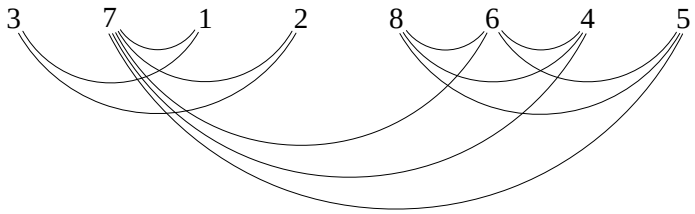


Figure 2.7: The inversions of the permutation 37128645

Theorem 2.10 *Let $f = a_1 \dots a_n$ be a permutation of $\{1,\dots,n\}$.*

(a) Let $g = a_1 \dots a_{i-1}a_{i+1}a_i a_{i+2}\dots a_n$ be a permutation obtained from f by exchanging two adjacent letters. Then $\mathrm{inv}(g) = \mathrm{inv}(f) \pm 1$.

(b) Let $g = a_1 \dots a_{i-1}a_j a_{i+1}\dots a_{j-1}a_i a_{j+1}\dots a_n$ be a permutation obtained from f by exchanging two not necessarily adjacent letters. Then $\mathrm{inv}(g)$ *and* $\mathrm{inv}(f)$ *are not of the same parity (i.e. one of them is even and the other is odd).*

Proof. (a) Note that exchanging two adjacent letters affects neither inversions of the form (a_j, a_k) where $j,k \notin \{i, i+1\}$, nor inversions of the form (a_j, a_i), (a_i, a_j), (a_j, a_{i+1}), (a_{i+1}, a_j) where $j \notin \{i, i+1\}$. So we either add a new inversion if $a_i < a_{i+1}$ in which case $\mathrm{inv}(g) = \mathrm{inv}(f)+1$, or take away an inversion if $a_i > a_{i+1}$ in which case $\mathrm{inv}(g) = \mathrm{inv}(f)-1$.

(b) Exchanging letters a_i and a_j where $i < j$ can be reduced to $2(j-i-1)+1$ operations of exchanging adjacent letters (swaps) as follows:

$$
\left.\begin{array}{l}
\text{swap letters at } j \text{ and } j-1 \\
\text{swap letters at } j-1 \text{ and } j-2 \\
\quad\vdots \\
\text{swap letters at } i+2 \text{ and } i+1
\end{array}\right\} j-i-1 \text{ swaps}
$$

$$
\left.\begin{array}{l}
\text{swap letters at } i+1 \text{ and } i
\end{array}\right\} 1 \text{ swap}
$$

$$
\left.\begin{array}{l}
\text{swap letters at } i+1 \text{ and } i+2 \\
\quad\vdots \\
\text{swap letters at } j-2 \text{ and } j-1 \\
\text{swap letters at } j-1 \text{ and } j
\end{array}\right\} j-i-1 \text{ swaps}
$$

Therefore $\mathrm{inv}(g) = \mathrm{inv}(f) \pm 1 \pm 1 \dots \pm 1$ (where the number of ± 1's is $2(j-i-1)+1$), so $\mathrm{inv}(f)$ and $\mathrm{inv}(g)$ are not of the same parity. □

Theorem 2.11 *If $n \geqslant 2$, the number of even permutations of $\{1,\dots,n\}$ is equal to the number of odd permutations of the same set.*

Proof. Let E_n be the set of even permutations and O_n the set of odd permutations of $\{1,\ldots,n\}$. Define $\varphi : E_n \to O_n$ by

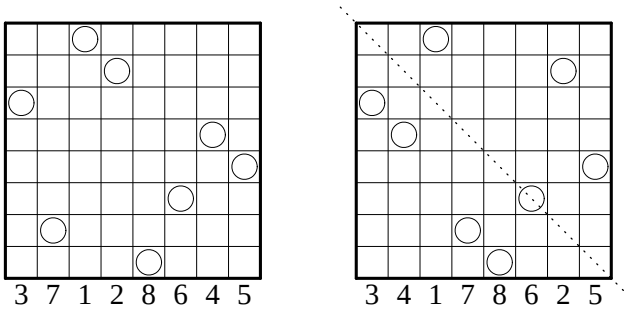
$$\varphi(a_1a_2a_3\ldots a_n) = a_2a_1a_3\ldots a_n.$$

Then one can easily see that φ is a bijection. □

Each permutation f of $\{1,\ldots,n\}$ being a bijective mapping has the inverse f^{-1} which is also a permutation of $\{1,\ldots,n\}$. We conclude this chapter by showing that a permutation and its inverse have the same number of inversions.

Theorem 2.12 *Let f be a permutation of $\{1,\ldots,n\}$. Then* $\mathrm{inv}(f) = \mathrm{inv}(f^{-1})$.

Proof. Let us first note that every permutation of $\{1,\ldots,n\}$ can be represented by an arrangement of n independent rooks on an $n \times n$ chess board. E.g., the representation of the permutation 37128645 as an arrangement of 8 independent rooks is given in Fig. 2.8 (a). Interestingly enough, it is easy to get a representation of the inverse permutation: just take the mirror image with respect to the main diagonal of the chess board. The inverse of 37128645 is 34178625 and it is given in Fig. 2.8 (b).



(a) (b)

Figure 2.8: A permutation and its inverse represented by an arrangement of independent rooks

This representation is also very suitable for the study of inversions. Note that inversions of a permutation correspond to those fields of the chess board where there is a rook below and a rook to the right. The inversions of the permutation in Fig. 2.8 (a) are marked by a $*$ in Fig. 2.9 (a).

The final step in the proof is to observe that the arrangement of stars on the chess board for the inverse permutation is a mirror image with respect to the main

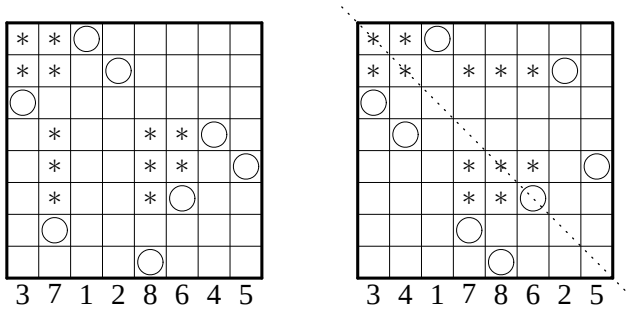


Figure 2.9: Inversions of a permutation and its inverse

diagonal of the chess board of the arrangement of stars for the original permutation, as demonstrated on Fig. 2.9 (b). The reason is simple: as we have seen, an inversion corresponds to an L-shaped structure as the one in Fig. 2.10 (a). Mirror image of such a configuration is again a configuration of the same kind, and hence an inversion. Therefore, f and f^{-1} have the same number of inversions. □

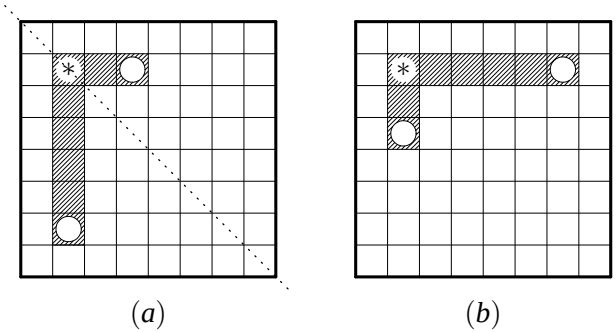


Figure 2.10: An inversion and its mirror image

Homework

2.1. Show that k^n is the number of *all* mappings $\{1,\ldots,n\} \to \{1,\ldots,k\}$.

2.2. Show that $\left\{ {n \atop 2} \right\} = 2^{n-1} - 1$, and $\left\{ {n \atop n-1} \right\} = \binom{n}{2}$.

2.3. Let S_n denote the set of all permutations on $\{1,\dots,n\}$ (understood as bijections $\{1,\dots,n\} \to \{1,\dots,n\}$), and let $\circ$ denote the function composition. Prove that $(S_n, \circ)$ is a group, that is

- if f and g are permutations from S_n, then so is $f \circ g$;
- $f \circ (g \circ h) = (f \circ g) \circ h$, for all $f, g, h \in S_n$;
- there is a permutation $e \in S_n$ such that $f \circ e = e \circ f = f$ for all $f \in S_n$;
- for every $f \in S_n$ there a $g \in S_n$ such that $f \circ g = g \circ f = e$.

This group is commonly referred to as the *symmetric group*.

2.4. Prove Theorem 2.5.

2.5. Show that $\begin{bmatrix} n \\ 2 \end{bmatrix} = (n-1)! \sum_{k=1}^{n-1} \frac{1}{k}$, and $\begin{bmatrix} n \\ n-1 \end{bmatrix} = \binom{n}{2}$.

2.6. Show that the mapping φ defined in the proof of Theorem 2.11 is well defined and that it is bijective.

Exercises

2.7. An ordered k-partition of a finite set A is a k-tuple $(B_1,\dots,B_k)$ such that $\{B_1,\dots,B_k\}$ is a partition of A. Let $n_1, \dots, n_k$ be positive integers such that $n_1 + \dots + n_k = |A|$. Find the number of ordered k-partitions $(B_1,\dots,B_k)$ of A such that $|B_i| = n_i$ for all i.

2.8. Find $\sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix}$.

2.9. Show that $\begin{Bmatrix} n \\ k \end{Bmatrix} \leqslant \begin{bmatrix} n \\ k \end{bmatrix}$ for all $n \geqslant k \geqslant 1$.

†**2.10.** Show that $\begin{Bmatrix} n+1 \\ m+1 \end{Bmatrix} = \sum_{k=m}^{n} \binom{n}{k} \begin{Bmatrix} k \\ m \end{Bmatrix}$.

†**2.11.** A *Bell number* $B(n)$ is the number of equivalence relations on an n-element set. Show that

(*a*) $B(n) = \sum_{k=1}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix}$,

(*b*) $B(n+1) = \sum_{k=0}^{n} \binom{n}{k} B(k)$ (note that $B(0) = 1$).

2.12. A *transposition* is a permutation of the form

$$1 \dots i-1 \quad j \quad i+1 \dots j-1 \quad i \quad j+1 \dots n.$$

Show that every permutation can be represented as a composition of transpositions.

2.13. Using 2.12 show that $(E_n, \circ)$ is a subgroup of $(S_n, \circ)$ (for the definition of E_n see the proof of Theorem 2.11). What is the index of this subgroup? Is it a normal subgroup of S_n?

2.14. Let $a_1 a_2 \dots a_{2005}$ be a permutation of $\{1, 2, \dots, 2005\}$. Show that

$$(a_1 + 1)(a_2 + 2) \dots (a_{2005} + 2005)$$

is an even number.

2.15. Find the number of permutations $a_1 a_2 \dots a_n$ of $\{1, 2, \dots, n\}$, $n \geqslant 3$, having the property that $|a_1 - a_2| > 1$.

2.16. Find the number of permutations $a_1 a_2 \dots a_n$ of $\{1, 2, \dots, n\}$, $n \geqslant 3$, having the property that $a_i < a_{i+2}$ for all $i \in \{1, \dots, n-2\}$.

†**2.17.** Find the number of permutations f of $\{1, 2, \dots, n\}$ such that $\mathrm{inv}(f) = 2$.

Chapter 3

SDRs and Latin Squares

In this chapter we first introduce systems of distinct representatives (SDRs for short) and show the celebrated Hall's Marriage Theorem. We apply the theorem to show that every Latin rectangle can be extended to a Latin square and estimate the number of Latin squares of order n. We then orthogonal Latin squares and there exists a complete system of orthogolan Latin squares of order n whenever n is a power of a prime. Finally, we show the famous result due to Euler that for every $n \geqslant 3$ such that $n \not\equiv 2 \pmod 4$ there exists a pair of orthogonal Latin squares of order n and conclude the chapter by showing that each system of $n-2$ mutually orthogonal Latin squares of order n can be extended to a complete system of orthogonal Latin squares.

3.1 Systems of distinct representatives

Let $\mathscr{A} = (A_1, \ldots, A_n)$ be a sequence of finite sets. A *system of distinct representatives (or SDR for short)* for $\mathscr{A}$ is a sequence $(e_1, \ldots, e_n)$ such that

- $e_i \neq e_j$ whenever $i \neq j$ (i.e., e_i's are distinct) and
- $e_i \in A_i$ for all i (i.e., e_i is a representative of A_i).

The problem we address in this section is: given a finite sequence of finite sets, is there a system of distinct representatives for this sequence?

Example 3.1 (a) Five good friends, Anne, Betty, Cecilia, Dorothy and Emanuela would like to get married to one of the six local boys Fred, George, Horatio, Ian, John and Kevin, and each girl has a list of candidates:

Anne's wish list:	Fred, George, Kevin
Betty's wish list:	Fred, Horatio
Cecilia's wish list:	Fred, Ian, John
Dorothy's wish list:	George, Ian, Kevin
Emanuela's wish list:	Horatio, John

Is it possible to arrange the marriages so that each girl gets married to a boy from her list? Yes it is. There are many possibilities and one of the solutions to the problem is: Anne-Fred, Betty-Horatio, Cecilia-Ian, Dorothy-George and Emanuela-John.

(b) Assume now that the wish lists of the five girls are

Anne's wish list:	Fred, George, Kevin
Betty's wish list:	Fred, Horatio
Cecilia's wish list:	George, Horatio
Dorothy's wish list:	George, Kevin
Emanuela's wish list:	George, Kevin, Horatio

Weeell, the situation is a bit tight this time. Consider Cecilia, Dorothy and Emanuela. Their wish lists all together contain three boys, George, Horatio and Kevin, so if there is a feasible arrangement of marriages, these three girls will have to marry these three boys (say, Cecilia-George, Dorothy-Kevin and Emanuela-Horatio). But now take a look at Anne and Betty. George, Horatio and Kevin are already married to Cecilia, Emanuela and Dorothy, so there is only one candidate left on two wish lists:

Anne's wish list:	Fred, ~~George~~, ~~Kevin~~
Betty's wish list:	Fred, ~~Horatio~~
Cecilia's wish list:	(George), Horatio
Dorothy's wish list:	George, (Kevin)
Emanuela's wish list:	George, Kevin, (Horatio)

and hence there is no feasible arrangement of marriages. A closer look reveals that the five wish lists contained only four boys altogether, so it was impossible from the beginning to make a feasible arrangement of marriages.

It is easy to see that if there is an SDR $(e_1,\ldots,e_n)$ for $(A_1,\ldots,A_n)$ then the union of every k sets in the sequence has at least k elements, for all $k \in \{1,\ldots,n\}$. The remarkable theorem due to Phillip Hall shows that this necessary condition is also sufficient. For $\mathscr{A} = (A_1,\ldots,A_n)$ and $\varnothing \neq J \subseteq \{1,\ldots,n\}$ let

$$\mathscr{A}(J) = \bigcup_{j\in J} A_j.$$

Theorem 3.2 (Hall's Marriage Theorem) *Let $\mathscr{A} = (A_1, \ldots, A_n)$ be a sequence of finite sets. Then $\mathscr{A}$ has an SDR if and only if*

$$|\mathscr{A}(J)| \geqslant |J| \text{ for all } \varnothing \neq J \subseteq \{1, \ldots, n\}. \qquad (\star)$$

Proof. ($\Rightarrow$) This is easy. If $(e_1, \ldots, e_n)$ is an SDR for $\mathscr{A}$ then $\mathscr{A}(J) = \bigcup_{j \in J} A_j \supseteq \{e_j : j \in J\}$ and since all e_j's are distinct, $|\mathscr{A}(J)| \geqslant |\{e_j : j \in J\}| = |J|$.

($\Leftarrow$) The proof proceeds by induction on n, the length of $\mathscr{A}$. If $n = 1$, the condition $(\star)$ guarantees that $|A_1| \geqslant 1$, so there is a representative of A_1. If $n = 2$ then $|A_1| \geqslant 1$, $|A_2| \geqslant 1$ and $|A_1 \cup A_2| \geqslant 2$ so we can easily find an SDR of (A_1, A_2). Assume now that $(\star)$ implies the existence of an SDR of every sequence with less than n sets and consider a sequence $\mathscr{A} = (A_1, \ldots, A_n)$ with n sets.

If $|\mathscr{A}(J)| > |J|$ for all $\varnothing \neq J \subset \{1, \ldots, n\}$, we have enough elements to play with and an SDR can be constructed easily. Take any $e_n \in A_n$ and let $B_j = A_j \setminus \{e_n\}$, $j \in \{1, \ldots, n-1\}$, and $\mathscr{B} = (B_1, \ldots, B_{n-1})$. Let us show that $\mathscr{B}$ satisfies $(\star)$. Take any $\varnothing \neq J \subseteq \{1, \ldots, n-1\}$ and note that $|\mathscr{B}(J)| = |\mathscr{A}(J)| - 1$ or $|\mathscr{B}(J)| = |\mathscr{A}(J)|$ according as $e_n \in \mathscr{A}(J)$ or not. Therefore, $|\mathscr{B}(J)| \geqslant |\mathscr{A}(J)| - 1 > |J| - 1$ since $|\mathscr{A}(J)| > |J|$ by assumption. Since we are working with integers here, $|\mathscr{B}(J)| > |J| - 1$ means that $|\mathscr{B}(J)| \geqslant |J|$ and thus $\mathscr{B}$ satisfies $(\star)$. By the induction hypothesis $\mathscr{B}$ has an SDR, say, $(e_1, \ldots, e_{n-1})$ and it is easy to see that $(e_1, \ldots, e_{n-1}, e_n)$ is an SDR for $\mathscr{A}$.

Assume now that the situation is tight, that is, $|\mathscr{A}(J)| = |J|$ for some $\varnothing \neq J \subset \{1, \ldots, n\}$. Without los of generality we can take $J = \{1, \ldots, s\}$ for some $1 \leqslant s < n$. Then $(A_1, \ldots, A_s)$ satisfies $(\star)$ and by the induction hypothesis there is an SDR $(e_1, \ldots, e_s)$ for $(A_1, \ldots, A_s)$. Since $|\mathscr{A}(J)| = |J|$ it follows that $\mathscr{A}(J) = \{e_1, \ldots, e_s\}$. Let $B_i = A_i \setminus \mathscr{A}(J)$, $s+1 \leqslant i \leqslant n$, and let us show that $\mathscr{B} = (B_{s+1}, \ldots, B_n)$ satisfies $(\star)$. Take any $\varnothing \neq K \subseteq \{s+1, \ldots, n\}$. Then it follows immediately that $\mathscr{B}(K) = \mathscr{A}(J \cup K) \setminus \mathscr{A}(J)$ so $|\mathscr{B}(K)| = |\mathscr{A}(J \cup K)| - |\mathscr{A}(J)|$. Now, $|\mathscr{A}(J \cup K)| \geqslant |J \cup K|$ since $\mathscr{A}$ satisfies $(\star)$, and $|\mathscr{A}(J)| = |J|$ by the assumption. Therefore, $|\mathscr{B}(K)| \geqslant |J \cup K| - |J| = |K|$ since J and K are disjoint. This shows that $\mathscr{B}$ satisfies $(\star)$ and by the induction hypothesis there is an SDR $(e_{s+1}, \ldots, e_n)$ for $\mathscr{B}$. Finally, $(e_1, \ldots, e_s, e_{s+1}, \ldots, e_n)$ is an SDR for $\mathscr{A}$. □

Next, we estimate the number of different SDRs of a sequence of finite sets that has an SDR. For integers $m_1 \leqslant m_2 \leqslant \ldots \leqslant m_n$ let

$$\begin{aligned} F_n(m_1, \ldots, m_n) &= (m_1)^+ (m_2 - 1)^+ (m_3 - 2)^+ \ldots (m_n - (n-1))^+ \\ &= \prod_{i=1}^{n} (m_i - (i-1))^+ \end{aligned}$$

where $(k)^+ = \max\{1, k\}$. Our goal is to show the following theorem:

Theorem 3.3 *Let $\mathscr{A} = (A_1, \ldots, A_n)$ be a sequence of finite sets that has an SDR and let* $\mathrm{SDR}(\mathscr{A})$ *denote the number of different SDRs for $\mathscr{A}$. Assume also that $|A_1| \leqslant \ldots \leqslant |A_n|$. Then* $\mathrm{SDR}(\mathscr{A}) \geqslant F_n(|A_1|, \ldots, |A_n|)$.

The proof of the theorem requires some preparation. Define $f_n : \mathbb{Z}^n \to \mathbb{N}$ by

$$f_n(a_1, \ldots, a_n) = F_n(m_1, \ldots, m_n)$$

where $m_1, \ldots, m_n$ is a permutation of $a_1, \ldots, a_n$ such that $m_1 \leqslant \ldots \leqslant m_n$. So, we can now write $f_n(a_1, \ldots, a_n)$ whenever we are not certain that $a_1 \leqslant \ldots \leqslant a_n$.

Lemma 3.4 *Let $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n, b, c \in \mathbb{Z}$ and assume that $b \leqslant c$. Then*

$$f_n(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n) \leqslant f_n(a_1, \ldots, a_{i-1}, c, a_{i+1}, \ldots, a_n).$$

Proof. Let $m_1 \leqslant \ldots \leqslant m_{k-1} \leqslant m_k = b \leqslant m_{k+1} \leqslant \ldots \leqslant m_l \leqslant m_{l+1} \leqslant \ldots \leqslant m_n$ and $m_1 \leqslant \ldots \leqslant m_{k-1} \leqslant m_{k+1} \leqslant \ldots \leqslant m_l \leqslant c \leqslant m_{l+1} \leqslant \ldots \leqslant m_n$ be nondecreasing rearrangements of $a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n$ and $a_1, \ldots, a_{i-1}, c, a_{i+1}, \ldots, a_n$. Then

$$\frac{f_n(a_1, \ldots, a_{i-1}, c, a_{i+1}, \ldots, a_n)}{f_n(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)} = \\ = \frac{(m_{k+1} - (k-1))^+ \cdot (m_{k+2} - k)^+ \cdot \ldots \cdot (m_l - (l-2))^+ \cdot (c - (l-1))^+}{(b - (k-1))^+ \cdot (m_{k+1} - k)^+ \cdot \ldots \cdot (m_{l-1} - (l-2))^+ \cdot (m_l - (l-1))^+}$$

which is clearly $\geqslant 1$ since $m_{k+1} \geqslant b$, $m_{k+2} \geqslant m_{k+1}$, …, $m_l \geqslant m_{l-1}$ and $c \geqslant m_l$. □

Proof. (of Theorem 3.3) The proof closely follows the outline of the proof of Theorem 3.2. We proceed by induction on n. If $n = 1$ then $\mathrm{SDR}(\mathscr{A}) = |A_1|$ and we are done. Suppose the claim holds for sequences with less than n sets and let $\mathscr{A} = (A_1, \ldots, A_n)$.

If $|\mathscr{A}(J)| > |J|$ for all $\varnothing \neq J \subset \{1, \ldots, n\}$, then take any $e_1 \in A_1$ and let $A_i(e_1) = A_i \setminus \{e_1\}$. We know from the proof of Theorem 3.2 that $\mathscr{A}(e_1) = (A_2(e_1), \ldots, A_n(e_1))$ has an SDR for every $e_1 \in A_1$, and since e_1 can be combined with every SDR of $\mathscr{A}(e_1)$ to produce and SDR of $\mathscr{A}$, we have that

$$\mathrm{SDR}(\mathscr{A}) = \sum_{e_1 \in A_1} \mathrm{SDR}(\mathscr{A}(e_1)).$$

By the induction hypothesis, $\mathrm{SDR}(\mathscr{A}(e_1)) \geqslant f_{n-1}(|A_2(e_1)|, \ldots, |A_n(e_1)|)$ and since $|A_i(e_1)| \geqslant |A_i| - 1$, Lemma 3.4 yields

$$\begin{aligned} \mathrm{SDR}(\mathscr{A}(e_1) &\geqslant f_{n-1}(|A_2(e_1)|, \ldots, |A_n(e_1)|) \\ &\geqslant f_{n-1}(|A_2| - 1, \ldots, |A_n| - 1) = F_{n-1}(|A_2| - 1, \ldots, |A_n| - 1) \end{aligned}$$

due to the assumption that $|A_2| \leqslant \ldots \leqslant |A_n|$. Putting it all together,

$$\begin{aligned}\mathrm{SDR}(\mathscr{A}) &= \sum_{e_1 \in A_1} \mathrm{SDR}(\mathscr{A}(e_1)) \geqslant \sum_{e_1 \in A_1} F_{n-1}(|A_2|-1,\ldots,|A_n|-1) \\ &= |A_1| \cdot F_{n-1}(|A_2|-1,\ldots,|A_n|-1) = F_n(|A_1|,|A_2|,\ldots,|A_n|).\end{aligned}$$

Assume now that $|\mathscr{A}(J)| = |J|$ for some $\varnothing \neq J \subset \{1,\ldots,n\}$. Let $J = \{j_1,\ldots,j_k\}$ and $\{1,\ldots,n\} \setminus J = \{m_1,\ldots,m_l\}$, $k+l=n$. Without loss of generality we may assume that $|A_{j_1}| \leqslant \ldots \leqslant |A_{j_k}|$ and $|A_{m_1}| \leqslant \ldots \leqslant |A_{m_l}|$. Let $E = A_{j_1} \cup \ldots \cup A_{j_k}$ and $A_{m_i}(E) = A_{m_i} \setminus E$, $i \in \{1,\ldots,l\}$. From the proof of Theorem 3.2 we know that each SDR for $\mathscr{A}$ consists of an SDR for $(A_{j_1},\ldots,A_{j_k})$ and an SDR for $(A_{m_1}(E), \ldots, A_{m_l}(E))$, so, by induction hypothesis

$$\begin{aligned}\mathrm{SDR}(\mathscr{A}) &= \mathrm{SDR}(A_{j_1},\ldots,A_{j_k}) \cdot \mathrm{SDR}(A_{m_1}(E),\ldots,A_{m_l}(E)) \\ &\geqslant F_k(|A_{j_1}|,\ldots,|A_{j_k}|) \cdot f_l(|A_{m_1}(E)|,\ldots,|A_{m_l}(E)|).\end{aligned}$$

Since $|E| = k$ and $|A_{m_i}(E)| \geqslant |A_{m_i}| - |E| = |A_{m_i}| - k$, Lemma 3.4 yields

$$\mathrm{SDR}(\mathscr{A}) \geqslant F_k(|A_{j_1}|,\ldots,|A_{j_k}|) \cdot F_l(|A_{m_1}|-k,\ldots,|A_{m_l}|-k).$$

Applying Lemma 3.4 once again, this time to the first factor on the left-hand side of the inequality, we obtain

$$\mathrm{SDR}(\mathscr{A}) \geqslant F_k(|A_1|,\ldots,|A_k|) \cdot F_l(|A_{m_1}|-k,\ldots,|A_{m_l}|-k)$$

since $|A_{j_i}| \geqslant |A_i|$ for $1 \leqslant i \leqslant k$. Next, let us remark that $|A_{j_k}| \leqslant |A_{j_1} \cup \ldots \cup A_{j_k}| = k$, so for $k \leqslant i \leqslant j_k$ we have that $|A_i| \leqslant |A_{j_k}| \leqslant k$ and thus $|A_i| - (i-1) \leqslant 1$. Therefore, $(|A_i| - (i-1))^+ = 1$ whence

$$F_k(|A_1|,\ldots,|A_k|) = \prod_{i=1}^{j_k} (|A_i| - (i-1))^+.$$

On the other hand,

$$F_l(|A_{m_1}|-k,\ldots,|A_{m_l}|-k) = \prod_{i=1}^{l} (|A_{m_i}| - (k+i-1))^+.$$

Now if $m_i \leqslant j_k$ then $|A_{m_i}| \leqslant |A_{j_k}| \leqslant k$, so $(|A_{m_i}| - (k+i-1))^+ = 1$. Since

$$(\{j_1,\ldots,j_k\},\{m_1,\ldots,m_l\})$$

is a partition of $\{1,\dots,n\}$ it follwos that $\{A_{m_i} : m_i > j_k\} = \{A_{j_k+1},\dots,A_n\}$ and thus

$$\begin{aligned} F_l(|A_{m_1}|-k,\dots,|A_{m_l}|-k) &= \prod_{i=1}^{l}(|A_{m_i}|-(k+i-1))^+ \\ &= \prod_{i=j_k+1}^{n}(|A_i|-(i-1))^+. \end{aligned}$$

The last equality is a bit tricky, so we show an example to demonstrate the main idea. Suppose that $n = 10$, $(j_1, j_2, j_3) = (3,5,6)$ and (m_1, m_2, m_3, m_4, m_5, m_6, m_7) = (1, 2, 4, 7, 8, 9, 10), so that

$$\mathscr{A} = (A_{m_1}, A_{m_2}, A_{j_1}, A_{m_3}, A_{j_2}, A_{j_3}, A_{m_4}, A_{m_5}, A_{m_6}, A_{m_7}).$$

Then $|A_{j_1} \cup A_{j_2} \cup A_{j_3}| = 3$ and

$$\begin{aligned} F_5(|A_{m_1}|-3,|A_{m_2}|-3,|A_{m_3}|-3,|A_{m_4}|-3,|A_{m_5}|-3,|A_{m_6}|-3,|A_{m_7}|-3) = \\ = (|A_{m_1}|-3)^+ \cdot (|A_{m_2}|-4)^+ \cdot (|A_{m_3}|-5)^+ \cdot (|A_{m_4}|-6)^+ \cdot \\ \cdot (|A_{m_5}|-7)^+ \cdot (|A_{m_6}|-8)^+ \cdot (|A_{m_7}|-9)^+. \end{aligned}$$

Since $|A_{m_1}| \leqslant |A_{m_2}| \leqslant |A_{m_3}| \leqslant |A_{j_3}| \leqslant 3$ we have that

$$(|A_{m_1}|-3)^+ = (|A_{m_2}|-4)^+ = (|A_{m_3}|-5)^+ = 1,$$

hence

$$\begin{aligned} F_5(|A_{m_1}|-3,|A_{m_2}|-3,|A_{m_3}|-3,|A_{m_4}|-3,|A_{m_5}|-3,|A_{m_6}|-3,|A_{m_7}|-3) = \\ = (|A_{m_4}|-6)^+ \cdot (|A_{m_5}|-7)^+ \cdot (|A_{m_6}|-8)^+ \cdot (|A_{m_7}|-9)^+ \\ = (|A_7|-6)^+ \cdot (|A_8|-7)^+ \cdot (|A_9|-8)^+ \cdot (|A_{10}|-9)^+. \end{aligned}$$

So much for the example. Finally, putting it all together we get

$$\begin{aligned} \mathrm{SDR}(\mathscr{A}) &\geqslant F_k(|A_1|,\dots,|A_k|) \cdot F_l(|A_{m_1}|-k,\dots,|A_{m_l}|-k) \\ &= \left(\prod_{i=1}^{j_k}(|A_i|-(i-1))^+\right) \cdot \left(\prod_{i=j_k+1}^{n}(|A_i|-(i-1))^+\right) \\ &= F_n(|A_1|,\dots,|A_n|). \end{aligned}$$

This completes the proof. □

Corollary 3.5 *Let $\mathscr{A} = (A_1, \ldots, A_n)$ be a sequence of finite sets that has an SDR and suppose $|A_i| \geqslant r$ for all i. Then*

$$\mathrm{SDR}(\mathscr{A}) \geqslant \begin{cases} r!, & r \leqslant n \\ \dfrac{r!}{(r-n)!}, & r > n. \end{cases}$$

Recall that a determinant of an $n \times n$ real matrix $A = [a_{ij}]$ is a number defined by

$$\det(A) = \sum_{f \in S_n} (-1)^{\mathrm{inv}(f)} \cdot a_{1f(1)} \cdot \ldots \cdot a_{nf(n)}$$

where the summation goes over all permutations f of $\{1, \ldots, n\}$. The signless version of the determinant is called a *permanent* of A. Hence the permanent of A is defined by

$$\mathrm{per}(A) = \sum_{f \in S_n} a_{1f(1)} \cdot \ldots \cdot a_{nf(n)}.$$

Although aparently simpler, permanents are almost impossible to compute effectively, which stands in sharp contrast to determinants that can be computed very easily.

Let $\mathscr{A} = (A_1, \ldots, A_n)$ be a sequence of subsets of $\{1, \ldots, n\}$. The *incidence matrix of* $\mathscr{A}$ is the $n \times n$ matrix $M_{\mathscr{A}} = [m_{ij}]$ where

$$m_{ij} = \begin{cases} 1, & A_i \ni j \\ 0, & \text{otherwise.} \end{cases}$$

An example of the incidence matrix of a sequence of sets is given in Fig. 3.1.

	1	2	3	4	5	6
$\{1,3,4\}$	1	0	1	1	0	0
$\{2,3,4,5\}$	0	1	1	1	1	0
$\{1,5\}$	1	0	0	0	1	0
$\{2,4,5\}$	0	1	0	1	1	0
$\{2,3,6\}$	0	1	1	0	0	1
$\{1,3,5,6\}$	1	0	1	0	1	1

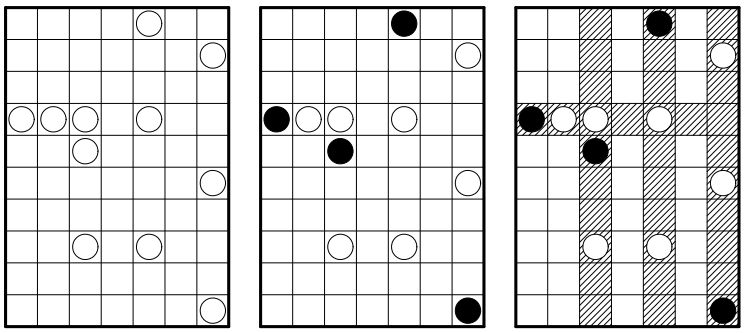
Figure 3.1: The incidence matrix of a sequence of sets

Theorem 3.6 *Let $\mathscr{A} = (A_1, \ldots, A_n)$ be a sequence of subsets of $\{1, \ldots, n\}$. Then* $\mathrm{SDR}(\mathscr{A}) = \mathrm{per}(M_{\mathscr{A}})$.

Proof. Note first that every SDR of $\mathscr{A}$ is a permutation of $\{1, \ldots, n\}$ and that $M_{\mathscr{A}}$ is a 01-matrix. In the sum that defines the permanent, therefore, some summands are 0 and others are 1. A summand that evaluates to 1 corresponds to a permutation f of $\{1, \ldots, n\}$ such that $m_{1f(1)} \cdot \ldots \cdot m_{nf(n)} = 1$. Therefore, $m_{if(i)} = 1$ for all i which is equivalent to $A_i \ni f(i)$ for all i, so $(f(1), \ldots, f(n))$ is an SDR for $\mathscr{A}$. Since every SDR of $\mathscr{A}$ arises from such a permutation, we get the equality. □

We now see that computing the number of SDRs of a sequence of finite sets is as complicated as calculating a permanent of a 01-matrix, which is in general extremely complicated.

The following theorem is a typical example of the *minimax phenomenon* which is one of the most fundamental insights in discrete mathematics. Suppose we are given an arrangement of rooks on a rectangular chess board where some rooks may attack each other. We would like to find the maximal number of independent rooks in this arrangement. Recall that rooks on a chess board are independent if no two of them are on the same line, a *line of a chess board* being a row or a column. In the arrangement of 11 rooks on a 10×7 chessboard in Fig. 3.2 (a) one can find at most four independent rooks, e.g., those four in Fig. 3.2 (b). Interestingly enough, all the rooks can be covered by four lines, Fig. 3.2 (c). The following theorem tells us that this is not a coincidence.

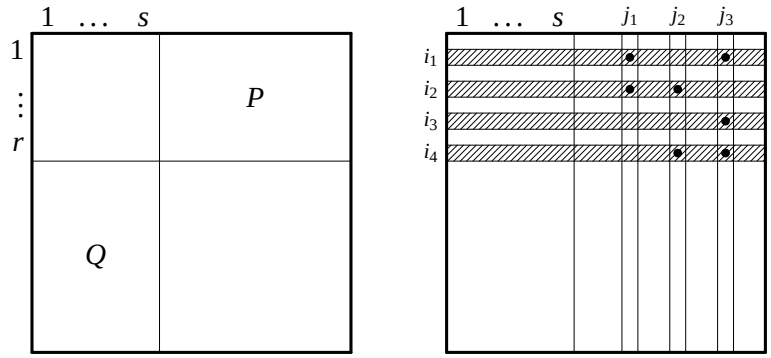


(a) (b) (c)

Figure 3.2: Eleven rooks on a 10×7 chess board

Theorem 3.7 *Suppose we are given an arrangement of rooks on a rectangular chess board. Then the maximum number of independent rooks in the arrangement is equal to the minimum number of lines that cover all the rooks.*

Proof. Let m be the minimum number of lines that cover all the rooks and let M be the maximum number of independent rooks. Since independent rooks are no two on the same line, we have $m \geqslant M$. We have to show that $m = M$.



(a) (b)

Figure 3.3: The proof of Theorem 3.7

Take any minimal collection of lines that cover all the rooks and suppose it consists of r rows and s columns, $r+s=m$. Without loss of generality, let these be the first r rows and the first s columns. Let us show that the region P in Fig. 3.3 (a) contains r independent rooks. For each row i, $1 \leqslant i \leqslant r$, let $A_i = \{j > s :$ there is a rook at the position $(i, j)\}$ and let us show that $(A_1, \ldots, A_r)$ has an SDR. Take any $k \in \{1, \ldots, r\}$ and k indices $i_1, \ldots, i_k \in \{1, \ldots, r\}$. If $|A_{i_1} \cup \ldots \cup A_{i_k}| < k$, the rooks in rows $i_1, \ldots, i_k$ which are not in the first s colums are arranged in less that k columns. Hence, we can replace the rows $i_1, \ldots, i_k$ by some $k-1$ columns and still cover all the rooks, Fig. 3.3 (b). But this is impossible since we have chosen the minimal number of covering lines. This shows that $(A_1, \ldots, A_r)$ has an SDR and consequently there are r independent rooks in region P. By the same argument, there are s independent rooks in region Q, so the number of independent rooks is at least $r+s=m$. Since M is the maximum number of independent rooks, $M \geqslant r+s=m$. □

We conclude the section by an important theorem due to G. Birkhoff from 1946. A *permutation matrix* is a square 01-matrix where each row contains precisely one 1, and each column contains precisely one 1.

Theorem 3.8 *Let $A = [a_{ij}]$ be an $n \times n$ matrix whose entries are nonnegative integers and with the property that the sum of every row and every column is m. Then A is the sum of m permutation matrices.*

Proof. We use induction on m. If $m = 1$ the claim is trivially true. Assume the claim is true for all sums less than m and let A be a matrix where the sum of every row and every column is $m > 1$. Define S_i, $1 \leqslant i \leqslant n$, by $S_i = \{j : a_{ij} > 0\}$. Take any k indices $1 \leqslant i_1 < \ldots < i_k \leqslant n$ and let us show that $|S_{i_1} \cup \ldots \cup S_{i_k}| \geqslant k$. The sum of these k rows is clearly km. Since every column of A has sum m, the nonzero entries in the rows $i_1, \ldots, i_k$ must spread over at least k columns (otherwise, if all the nonzero entries in these k rows are concentrated in $k-1$ columns, the sum of the rows could not exceed $(k-1)m$). Therefore, $|S_{i_1} \cup \ldots \cup S_{i_k}| \geqslant k$ for all k and all choices of $i_1, \ldots, i_k$. This shows that $(S_1, \ldots, S_n)$ has an SDR $(s_1, \ldots, s_n)$ which corresponds to the permutation matrix $P = [p_{ij}]$ where $p_{is_i} = 1$ and all other entries of P are zero. The sum of every row and every column of $A - P$ is $m-1$, so by the induction hypothesis, $A - P$ is the sum of $m-1$ permutation matrices. Therefore, A is a sum of m permutation matrices. □

3.2 Latin squares

Let Q be a finite set with $n \geqslant 2$ elements and let $1 \leqslant r \leqslant n$. A *Latin $r \times n$ rectangle over Q* is an $r \times n$ matrix with entries from Q such that in every row all elements are distinct and in every column all elements are distinct. A *Latin square of order n over Q* is a Latin $n \times n$ rectangle over Q. Fig. 3.4 shows a Latin 3×5 rectangle and a Latin square of order 6. The construction of the Latin square in Fig. 3.4 also shows that for every n there is a Latin square of order n and hence, for every $1 \leqslant r \leqslant n$ there is a Latin $r \times n$ rectangle.

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

1	2	3	4	5	6
2	3	4	5	6	1
3	4	5	6	1	2
4	5	6	1	2	3
5	6	1	2	3	4
6	1	2	3	4	5

Figure 3.4: A Latin rectangle and a Latin square

Let $\varphi : Q \to Q$ be a bijection and let $R = [a_{ij}]_{r \times n}$ be a Latin $r \times n$ rectangle over Q. By $\varphi(R)$ we denote the matrix $\varphi(R) = [\varphi(a_{ij})]_{r \times n}$.

Lemma 3.9 *Let R be a Latin rectangle over Q.*

(a) A matrix obtained from R by permuting rows is again a Latin rectangle.

(b) A matrix obtained from R by permuting columns is again a Latin rectangle.

(c) If $\varphi : Q \to Q$ is a bijection, then $\varphi(R)$ is a Latin rectangle.

Let $Q = \{a_1, \ldots, a_n\}$ be a set of integers and let $a_1 < \ldots < a_n$. A Latin square over Q is said to be *standard* if the elements of the first row of the square are linearly ordered. It is said to be *doubly standard* if the elements of the first row and of the first column of the square are linearly ordered. Fig. 3.5 show a standard and a doubly standard Latin square of order 5.

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

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

Figure 3.5: A standard and a doubly standard Latin square of order 5

Lemma 3.10 (a) *Every Latin square can be turned into a standard or a doubly standard Latin square by permuting rows and columns of the original square.*

(b) *For every Latin square L over Q there is a bijection $\varphi : Q \to Q$ such that $\varphi(L)$ is a standard Latin square.*

The following theorem is yet another important application of the Hall's Marriage Theorem 3.2.

Theorem 3.11 *Let $1 \leqslant r < n$. Then every Latin $r \times n$ rectangle can be extended to a Latin $(r+1) \times n$ rectangle.*

Proof. Consider a Latin $r \times n$ rectangle R over an n element set Q and for each $i \in \{1, \ldots, n\}$ put

$$S_i = \{x \in Q : x \text{ does not appear in the } i\text{-th column of } R\}.$$

Then R can be extended by one row if and only if $(S_1, \ldots, S_n)$ has an SDR, so let us show that $(S_1, \ldots, S_n)$ has an SDR. Take any k indices $j_1, \ldots, j_k \in \{1, \ldots, n\}$, let $S_{j_1} \cup \ldots \cup S_{j_k} = \{a_1, \ldots, a_l\}$ and let us show that $l \geqslant k$. For $a \in Q$ let $N(a)$ denote the number of sets $S_1, \ldots, S_n$ that contain a. Since a appears in every row of R precisely once, $N(a) = n - r$ for all $a \in Q$. It is easy to see that

$$|S_{j_1}| + \ldots + |S_{j_k}| \leqslant N(a_1) + \ldots + N(a_l) = l(n-r).$$

On the other hand every S_i has precisely $n - r$ elements, so

$$|S_{j_1}| + \ldots + |S_{j_k}| = k(n-r).$$

Therefore, $l(n-r) \geqslant k(n-r)$ and thus $l \geqslant k$. This shows that $(S_1, \ldots, S_n)$ satisfies the requirement $(\star)$ in the Hall's Marriage Theorem 3.2, and hence has an SDR. □

Corollary 3.12 *Every Latin rectangle can be extended to a Latin square.*

We conclude the section with an estimate on the number of Latin squares.

Theorem 3.13 *Let λ_n denote the number of distinct Latin squares on an n element set. Then*

$$\prod_{k=1}^{n} k! \leqslant \lambda_n \leqslant \prod_{k=0}^{n-1} (n! - k).$$

Proof. We count Latin squares by adding rows one at a time. Each row of a Latin square is a permutation of Q, so there are $n!$ possibilities for the first row, then there are at most $n! - 1$ possibilities for the second row since the permutation chosen for the first row must not be used again, at most $n! - 2$ possibilities for the third row and so on. Therefore, $\lambda_n \leqslant \prod_{k=0}^{n-1}(n! - k)$.

As for the lower bound, let us first note that there are at least $(n-r)!$ possibilities to extend a Latin $r \times n$ rectangle to a Latin $(r+1) \times n$ rectangle. To see this, form sets S_i as in the proof of Theorem 3.11. Each S_i has $n - r$ elements, so by Corollary 3.5 the sequence $(S_1, \ldots, S_n)$ has at least $(n-r)!$ SDRs. Now, there are $n!$ possibilities to choose the first row of the Latin square, then at least $(n-1)!$ posibilities to find an SDR that consitutes the second row, at least $(n-2)!$ SDRs for the third row, and so on. Therefore, $\lambda_n \geqslant \prod_{k=1}^{n} k!$. □

3.3 Orthogonal Latin squares

Let us start with an old card game which was rather popular in the Middle Ages. From a deck of playing cards take all aces, kings, queens and jacks, and arrange them in a 4×4 array so that each row and column of the array contains an ace (A), a king (K), a queen (Q) and a jack (J), but also a spade (♠), a heart (♡), a club (♣) and a diamond (♢). One possible solution to this ancient problem is given in

A♠	K♡	Q♣	J♢
Q♢	J♣	A♡	K♠
J♡	Q♠	K♢	A♣
K♣	A♢	J♠	Q♡

→

A	K	Q	J
Q	J	A	K
J	Q	K	A
K	A	J	Q

&

♠	♡	♣	♢
♢	♣	♡	♠
♡	♠	♢	♣
♣	♢	♠	♡

Figure 3.6: A medieval problem vs. Latin squares

Fig. 3.6. A careful look reveals that the arrangement of the 16 cards splits into two

Latin squares of order 4: one taking care of ranks and the other one taking care of colours.

Definition 3.14 Latin squares $L_1 = [a_{ij}]_{n\times n}$ over Q_1 and $L_2 = [b_{ij}]_{n\times n}$ over Q_2 are *orthogonal*, in symbols $L_1 \perp L_2$, if $\{(a_{ij}, b_{ij}) : i, j \in \{1,\dots,n\}\} = Q_1 \times Q_2$.

In other words, L_1 and L_2 are orthogonal if the L_2 overlaid with L_1 contains every pair from $Q_1 \times Q_2$. We shall mainly work with Latin squares over the same set Q, although in some examples Q_1 might differ from Q_2.

Lemma 3.15 *Let L_1 and L_2 be orthogonal Latin squares over Q. Then*
(a) $\varphi(L_1) \perp L_2$ for every bijection $\varphi : Q \to Q$;
(b) $\varphi(L_1) \perp \psi(L_2)$ for every pair of bijections $\varphi, \psi : Q \to Q$.

Definition 3.16 Latin squares L_1, L_2, …, L_k are mutually orthogonal if $L_i \perp L_j$ whenever $i \neq j$.

Theorem 3.17 *Let L_1, L_2, …, L_k be mutually orthogonal Latin squares over Q where $|Q| = n$. Then $k \leqslant n-1$.*

Proof. Without loss of generality we can assime that $Q = \{1,\dots,n\}$. According to Lemma 3.10 (b), for every $i \in \{1,\dots,k\}$ there is a bijection $\varphi_i : Q \to Q$ such that $L'_i = \varphi_i(L_i)$ is a standard Latin sqaure and from Lemma 3.15 it follows that L'_1, …, L'_k are mutually orthogonal:

$$\begin{array}{cccc} L_1 & L_2 & \dots & L_k \\ {\scriptstyle\varphi_1}\downarrow & {\scriptstyle\varphi_2}\downarrow & \dots & \downarrow{\scriptstyle\varphi_k} \\ L'_1 & L'_2 & \dots & L'_k \end{array} \qquad\qquad \begin{array}{ccc} L_i & \perp & L_j \\ {\scriptstyle\varphi_i}\downarrow & & \downarrow{\scriptstyle\varphi_j} \\ L'_i & \perp & L'_j \end{array}$$

Let $b_j \in Q$ be the element in L'_j at the position $(2,1)$, $j \in \{1,\dots,k\}$, Fig. 3.7. Then $b_j \neq 1$ for all j since in each L'_j there is an 1 in the first row, just above b_j.

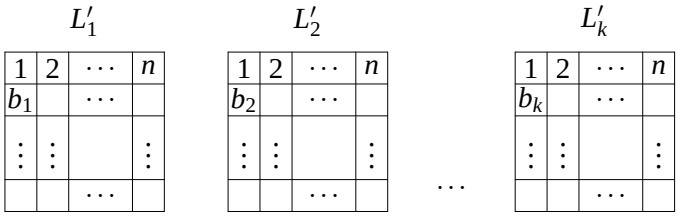


Figure 3.7: The proof of Theorem 3.17

Let us now show that $b_i \neq b_j$ whenever $i \neq j$. Since both L'_i and L'_j are standard,

overlaying the first row of L'_i with the first row of L'_j produces $(1,1)$, $(2,2)$, $\ldots$, (n,n). From $L'_i \perp L'_j$ it follows that every pair of Q^2 appears exactly once when L'_i is overlaid with L'_j so $(b_i, b_j) \notin \{(s,s) : s \in Q\}$ and hence $b_i \neq b_j$. Therefore, $\{b_1, \ldots, b_k\}$ is a k-element subset of $\{2, \ldots, n\}$ whence $k \leqslant n-1$. □

We see from the above theorem that any system of mutually orthogonal Latin squares has at most $n-1$ elements, $n = |Q|$. Systems achieving the upper bound are said to be complete.

Definition 3.18 A *complete system of orthogonal Latin squares over Q* is a system of $n-1$ mutually ortogonal Latin squares, $n = |Q|$.

Theorem 3.19 *Let $|Q| = p^\alpha$ where p is a prime and $\alpha \in \mathbb{N}$. Then there exists a complete system of orthogonal Latin squares over Q.*

Proof. This proof heavily relies on a nontrivial fact that for every prime p and every $\alpha \in \mathbb{N}$ there exists a finite field with p^α elements. Recall that a field is an algebraic structure where we can add, subtract, multiply and divide by any $a \neq 0$ (0 being the neutral element for addition). Therefore, in a field we can perform all the usual arithmetic, regardless of the fact that it need not be one of the number systems we are used to work with, such as $\mathbb{R}$ or $\mathbb{Q}$. In particular, we can solve systems of linear equations using the same strategies we use when solvig systems of linear equations in $\mathbb{Q}$ or $\mathbb{R}$.

So, let Q be a finite field with $n = p^\alpha$ elements and let us denote the operations in Q in a usual way. Let 0 denote the neutral element for $+$. For each $k \in Q \setminus \{0\}$ we define an $n \times n$ matrix $L^k = [a^k_{ij}]$ over Q indexed by $\{0, \ldots, n-1\}$ as follows:

$$a^k_{ij} = i + j \cdot k$$

and let us show that each L^k is a Latin square. If $a^k_{i_1 j} = a^k_{i_2 j}$ then $i_1 + jk = i_2 + jk$ whence $i_1 = i_2$, so in each column of L^k all elements are distinct. If $a^k_{i j_1} = a^k_{i j_2}$ then $i + j_1 k = i + j_2 k$ i.e. $j_1 k = j_2 k$. Since $k \neq 0$ we can divide by k whence $j_1 = j_2$. Therefore, in each row of L^k all elements are distinct.

Finally, let us show that $L^k \perp L^m$ for $k \neq m$. Take any $(\alpha, \beta) \in Q^2$ and let us find indices i and j such that $(a^k_{ij}, a^m_{ij}) = (\alpha, \beta)$. We have to solve the following system of linear equations in unknowns i and j:

$$i + kj = \alpha, \qquad i + mj = \beta.$$

But we are in a field, so this is easy:

$$i = \frac{k\beta - m\alpha}{k-m}, \qquad j = \frac{\alpha - \beta}{k-m}.$$

This shows that each (α,β) appears when L^k is overlaid with L^m so $L^k \perp L^m$. $\square$

For a matrix $L=[b_{ij}]_{n\times n}$ and an a let $a\otimes L=[(a,b_{ij})]_{n\times n}$. For matrices $L_1=[a_{ij}]_{m\times m}$ and $L_2=[b_{ij}]_{n\times n}$ let $L_1\otimes L_2$ be the $mn\times mn$ matrix given in a block representation by

$$L_1\otimes L_2=\begin{bmatrix} a_{11}\otimes L_2 & a_{12}\otimes L_2 & \dots & a_{1m}\otimes L_2 \\ a_{21}\otimes L_2 & a_{22}\otimes L_2 & \dots & a_{2m}\otimes L_2 \\ \dots & \dots & \dots & \dots \\ a_{m1}\otimes L_2 & a_{m2}\otimes L_2 & \dots & a_{mm}\otimes L_2 \end{bmatrix}$$

Clearly, if L_1 is a Latin square over Q_1 and L_2 is a Latin square over Q_2 then $L_1\otimes L_2$ is a Latin square over $Q_1\times Q_2$ (Homework 3.9).

Lemma 3.20 *Let $L_1,\dots,L_k$ be mutually orthogonal Latin squares over Q and $L'_1,\dots,L'_k$ mutually orthogonal Latin squares over Q'. Then $L_1\otimes L'_1,\dots,L_k\otimes L'_k$ are mutually orthogonal Latin squares over $Q\times Q'$.*

Lemma 3.21 *Let $n=p_1^{\alpha_1}p_2^{\alpha_2}\dots p_k^{\alpha_k}$ where p_i's are distinct primes and $\alpha_i>0$ for all i, and let q be the minimum of $p_1^{\alpha_1}$, $p_2^{\alpha_2}$, …, $p_k^{\alpha_k}$. Then there exist $q-1$ mutually orthogonal Latin squares of order n.*

Proof. Let $q_i=p_i^{\alpha_i}$ and $q=\min\{q_1,\dots,q_k\}$. We know that for each i there exists a complete system of orthogonal Latin squares of order q_i. Since any subset of a complete system of orthogonal Latin squares is a set of mutually orthogonal Latin squares and since $q\leqslant q_i$ for all i, it follows that for each i there is a set L_1^i, L_2^i, …, L_{q-1}^i of $q-1$ mutually orthogonal Latin squares of order q_i. Lemma 3.20 now yields that

$$L_1^1\otimes\dots\otimes L_1^k,\quad L_2^1\otimes\dots\otimes L_2^k,\quad \dots,\quad L_{q-1}^1\otimes\dots\otimes L_{q-1}^k$$

is a set of $q-1$ mutually orthogonal Latin squares of order $q_1q_2\dots q_k=n$. $\square$

Theorem 3.22 (Euler) *If $n\not\equiv 2 \pmod 4$, there exists a pair of orthogonal Latin squares of order n.*

Proof. Let $n=p_1^{\alpha_1}p_2^{\alpha_2}\dots p_k^{\alpha_k}$ where p_i's are distinct primes and $\alpha_i>0$ for all i, and let $q=\min\{p_1^{\alpha_1},p_2^{\alpha_2},\dots,p_k^{\alpha_k}\}$. If $q=2$ then $n=2\cdot p_2^{\alpha_2}\dots p_k^{\alpha_k}$ where p_i's are odd primes whence follows that $n\equiv 2 \pmod 4$, which contradicts the assumption. Therefore, $q\geqslant 3$. According to Lemma 3.21 there exists a set of $q-1\geqslant 2$ mutually orthogonal Latin squares of order n, and in particular, there exists a pair of mutually orthogonal Latin squares of order n. $\square$

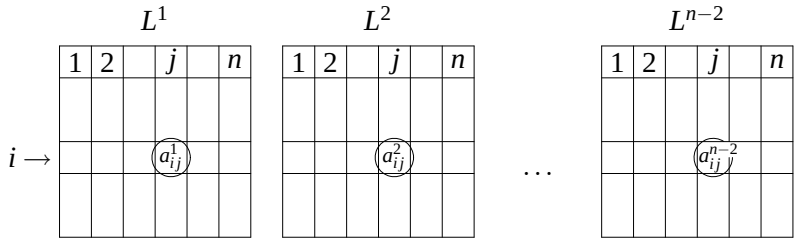


Figure 3.8: The set S_{ij}

Theorem 3.23 *Every set of $n-2$ mutually orthogonal Latin squares of order n can be extended (by adding "a missing square") to a complete system of orthogonal Latin squares.*

Proof. Let $L^1 = [a_{ij}^1], \ldots, L^{n-2} = [a_{ij}^{n-2}]$ be mutually orthogonal Latin squares over $\{1,2,\ldots,n\}$. Without loss of generality we can assume that all L^i's are standard (if this is not the case, we can find permutations $\varphi_1, \ldots, \varphi_{n-2}$ such that each $\varphi_i(L^i)$ is standard, and $\varphi_1(L^1), \ldots, \varphi_{n-2}(L^{n-2})$ are still mutually orthogonal; if a Latin square L^* is orthogonal to each $\varphi_i(L^i)$, it will be orthogonal to each L^i as well).

Let $S_{ij} = \{a_{ij}^1, a_{ij}^2, \ldots, a_{ij}^{n-2}\}$, see Fig. 3.8. Since $L^1, \ldots, L^{n-2}$ are mutually orthogonal standard Latin squares, it is easy to show that $|S_{ij}| = n-2$ and that $j \notin S_{ij}$ for all i and j. Let $a_{1j}^* = j$, $1 \leqslant j \leqslant n$, and for $i \geqslant 2$ let a_{ij}^* be the only element of $\{1,\ldots,n\}$ that does not appear in $S_{ij} \cup \{j\}$. Put $L^* = [a_{ij}^*]$. We are going to show that L^* is a Latin square and that $L^i \perp L^*$ for all i. Note that by construction the first row of L^* is $12\ldots n$.

Let us first show that L^* is a Latin square. Suppose that an element, say 1, is missing in the row $i \geqslant 2$ of L^*. Then $S_{i2}, \ldots, S_{in}$ all contain 1. From each square $L^1, \ldots, L^{n-2}$ take the i-th row without its first cell and arrange these rows in a matrix as in Fig. 3.9. The matrix has $n-1$ columns $S_{i2}, \ldots, S_{in}$ and each column contains a 1. On the other hand, the matrix clearly has $n-2$ rows, so there has to be a row which contains two 1's. But this is impossible because these are rows of Latin sqaures. Therefore, each row of L^* contains each element of $\{1,\ldots,n\}$. The proof that each column of L^* contains each element of $\{1,\ldots,n\}$ is analogous, so L^* is a Latin square.

Now let us show that $L^1 \perp L^*$. Suppose $L^1 \not\perp L^*$. Then there is a pair $(x,y) \in \{1,\ldots,n\}^2$ such that $(x,y) \notin \{(l_{ij}^1, l_{ij}^*) : 1 \leqslant i,j \leqslant n\}$. Since both L^1 and L^* are standard we have $x \neq y$, so without loss of generality we can assume that $(x,y) = (1,2)$. We can also assume that 1's are on the main diagonal of L^1 (if this is not the case, we can simultaneously permute rows of $L^1, \ldots, L^{n-2}, L^*$ to achieve this). Since $(1,2)$ does not appear in $\{(l_{ij}^1, l_{ij}^*) : 1 \leqslant i,j \leqslant n\}$ we see that there are no 2's

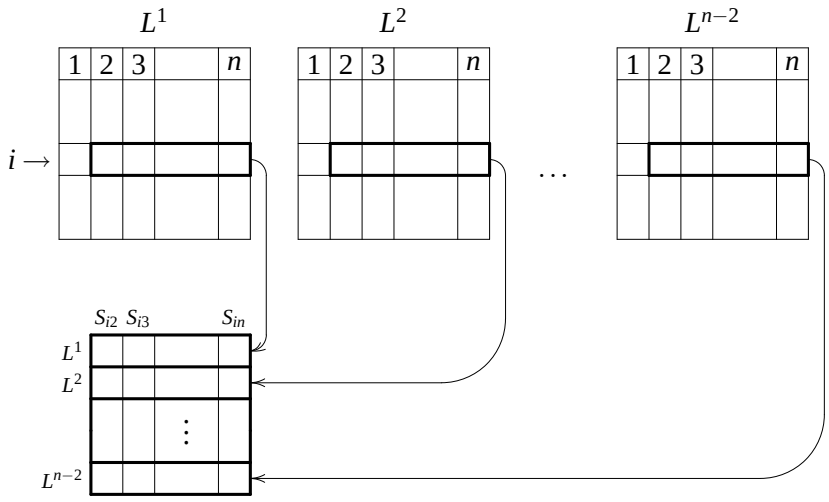


Figure 3.9: The proof that L^* is a Latin square

on the main diagonal of L^*, which, by the construction of L^*, means that $2 \in S_{33}$, $2 \in S_{44}, \ldots, 2 \in S_{nn}$. From each square $L^1, \ldots, L^{n-2}$ take the main diagonal without its first two cells and write these diagonals as rows of a matrix, see Fig. 3.9. The matrix has $n-2$ columns $S_{33}, \ldots, S_{nn}$ and each column contains a 2. On the other hand, the matrix clearly has $n-2$ rows and the first row is $11\ldots1$. So all the 2's appear in the remaining $n-3$ rows and hence there has to be a row, say row s, which contains two 2's. But then $L^1 \not\perp L^s$ since $(1,2)$ appears twice when we overlay L^1 with L^s. The contradiction shows that $L^1 \perp L^*$. The proof that $L^i \perp L^*$ for the remaining i's is analogous, so $L^1, \ldots, L^{n-2}, L^*$ is a complete system of orthogonal Latin squares. □

Homework

3.1. Prove the following generalisation of Hall's Marriage Theorem:
Let $\mathscr{A} = (A_1, \ldots, A_n)$ be a sequence of finite sets, let $r \geqslant 0$ be an integer and assume that

$$|\mathscr{A}(J)| \geqslant |J| - r \text{ for all } \varnothing \neq J \subseteq \{1, \ldots, n\}.$$

Then there are indices $1 \leqslant i_1 < \ldots < i_{n-r} \leqslant n$ such that $(A_{i_1}, \ldots, A_{i_{n-r}})$ has an SDR. (Hint: take r distinct elements $x_1, \ldots, x_r \notin \bigcup_{i=1}^n A_i$ and consider $\mathscr{A}' = (A'_1, \ldots, A'_n)$ where $A'_i = A_i \cup \{x_1, \ldots, x_r\}$.)

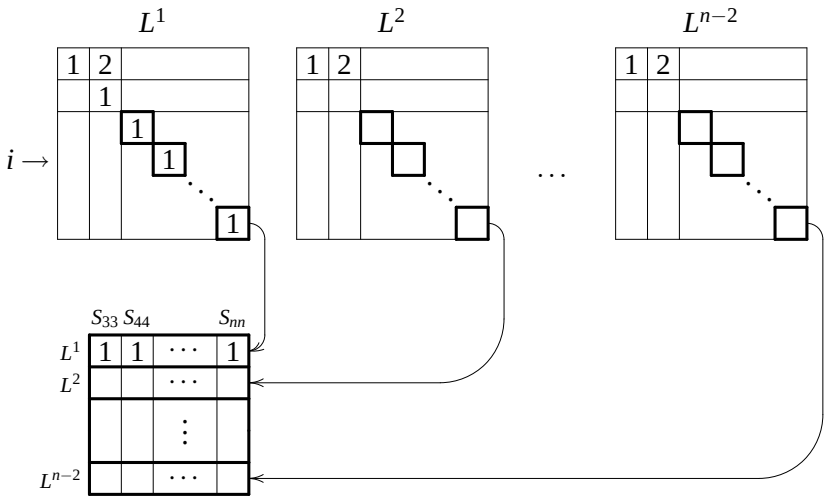


Figure 3.10: The proof that $L^1 \perp L^*$

3.2. Let $1 \leqslant m_1 \leqslant m2 \leqslant \ldots \leqslant m_n$. Show that $m_1 \cdot F_{n-1}(m_2 - 1, \ldots, m_n - 1) = F_n(m_1, m_2, \ldots, m_n)$.

3.3. Prove Corollary 3.5.

3.4. Let $Q = \{1, 2, \ldots, mn\}$ where $m, n \geqslant 2$. Let $\{A_1, \ldots, A_n\}$ be a partition of Q into n blocks of size m, and let $\{B_1, \ldots, B_n\}$ be another partition of Q into n blocks of size m. Show that there is a permutation f of $\{1, \ldots, n\}$ such that $A_i \cap B_{f(i)} \neq \varnothing$ for all i. (Hint: Take the $n \times n$ integer matrix $M = [m_{ij}]$ where $m_{ij} = |A_i \cap B_j|$ and apply Theorem 3.8.)

3.5. Prove Lemmas 3.9 and 3.10.

3.6. Find all doubly standard Latin squares of order 4.

3.7. Let $\lambda_{r \times n}$ denote the number of distinct Latin $r \times n$ rectangles on an n element set. Show that

$$\prod_{k=0}^{r-1}(n-k)! \leqslant \lambda_{r \times n} \leqslant \prod_{k=0}^{r-1}(n! - k).$$

3.8. Prove Lemma 3.15.

3.9. Show that if L_1 is a Latin square over Q_1 and L_2 is a Latin square over Q_2 then $L_1 \otimes L_2$ is a Latin square over $Q_1 \times Q_2$.

3.10. Prove Lemma 3.20.

Exercises

3.11. Find an SDR for $(\{1\},\{1,2,3\},\{3,4\},\{2,4,5\},\{3,6\},\{1,4,7\},\{6\})$. How many SDRs does this sequence of sets have?

3.12. Find subsets A, B, C of $\{1,2,3\}$ such that $\mathrm{SDR}(A,B,C) = 3$.

3.13. For each $n \geqslant 3$ find n subsets $A_1, \ldots, A_n$ of $\{1,2,\ldots,n\}$ such that $|A_1| = \ldots = |A_n|$ and $\mathrm{SDR}(A_1, \ldots, A_n) = 2$.

3.14. Find $\mathrm{SDR}(\mathscr{A})$ where $\mathscr{A} = (\{1,2,3\},\{1,4,5\},\{1,6,7\},\{2,4,6\},\{2,5,7\},\{3,4,7\},\{3,5,6\})$.

†**3.15.** Let $A_i = \{1,\ldots,n\} \setminus \{i\}$, $1 \leqslant i \leqslant n$ and let $f(n) = \mathrm{SDR}(A_1,\ldots,A_n)$.

(a) Find $f(n)$. (Hint: Use the Principle of Inclusion-Exclusion.)

(b) Compute $\lim\limits_{n\to\infty} \dfrac{f(n)}{n!}$.

3.16. Let $\mathscr{A} = (A_1,\ldots,A_n)$ be a sequence of subsets of $\{1,\ldots,n\}$. Show that if $M_{\mathscr{A}}$ is a regular real matrix, then $\mathscr{A}$ has an SDR. (Hint: if $M_{\mathscr{A}}$ is a regular real matrix then $\det(M_{\mathscr{A}}) \neq 0$; conclude that at least one of the summands in the expression for $\mathrm{per}(M_{\mathscr{A}})$ is nonzero using the fact that $M_{\mathscr{A}}$ is a 01-matrix.)

3.17. Turn the following *partial Latin squares* into Latin squares:

(a)

2	1			
			3	4
				3
1				

(b)

1				
	1			
		1		
			1	
				2

†**3.18.** Find the number of Latin $2 \times n$ rectangles, $n \geqslant 2$. (Hint: Use 3.15.)

3.19. Let L be a Latin square of order n where n is an odd integer and suppose L is symmetric with respect to its main diagonal. Show that all the elements on the main diagonal of L are distinct.

3.20. Find all pairs of orthogonal Latin squares of order 3.

3.21. Let $L^+ = [a_{ij}]$ and $L^- = [b_{ij}]$ be Latin squares over $\mathbb{Z}_n = \{0,1,\ldots,n-1\}$ indexed by $\mathbb{Z}_n$, i.e.

$$L^+ = \begin{bmatrix} a_{00} & a_{01} & \ldots & a_{0,n-1} \\ a_{10} & a_{11} & \ldots & a_{1,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n-1,0} & a_{n-1,1} & \ldots & a_{n-1,n-1} \end{bmatrix}$$

and similarly for L^-. Assume that

$$a_{ij} = i +_n j \quad \text{and} \quad b_{ij} = i -_n j$$

where $+_n$ and $-_n$ are addition and subtraction in $\mathbb{Z}_n$. Show that $L^+ \perp L^-$ if $n \geqslant 3$ is odd.

3.22. We say that Latin squares L_1 and L_2 over Q are *isotopic* and write $L_1 \sim L_2$ if there is a permutation $\varphi : Q \to Q$ such that L_2 can be obtained from L_1 by permuting rows and columns of $\varphi(L_1)$.

(a) Show that there are two nonisotopic Latin squares of order 4. (Hint: Use 3.6.)

(b) Show that one of them has an orthogonal "mate" and the other does not.

3.23. Prove that if $L_1 \sim L_2$ and L_1 has an orthogonal "mate", then so does L_2.

3.24. A *cross-section* of a Latin square $L = [l_{ij}]$ of order n is a set $\delta = \{(1,k_1), (2,k_2), \ldots, (n,k_n)\}$ such that $(k_1, \ldots, k_n)$ is a permutation of $\{1, \ldots, n\}$ and $l_{ik_i} \neq l_{jk_j}$ whenever $i \neq j$. In other words, a cross-section is a selection of cells in L such that there is one cell in each row and each column of L, and all the entries in these cells are distinct.

Show that a Latin square L of order n has an orthogonal mate if and only if L has n pairwise disjoint cross-sections. (Hint: if there is an L' such that $L \perp L'$, put $\delta_k = \{(i,j) : l'_{ij} = k\}$ and show that $\delta_1, \ldots, \delta_n$ are pairwise disjoint cross-sections; the other implication uses the same idea to reconstruct an orthogonal mate of L from n pairwise disjoint cross-sections.)

3.25. A Latin square $L = [l_{ij}]$ over Q is said to be *row-complete* if for every $(p,q) \in Q^2$ such that $p \neq q$ there is exactly one pair (i,j) such that $l_{ij} = p$ and $l_{i+1,j} = q$ (that is, every pair of distinct elements of Q occurs exactly once in consecutive positions in the same row). The definition of the *column-complete* Latin square is analogous.

Show that for every even $n \geqslant 4$ there exists a Latin square of order n that is both row-complete and column-complete. (Hint: Let $n = 2k$ and let $(x_1, \ldots, x_n)$ be the following sequence:

$$(0, 1, 2k-1, 2, 2k-2, 3, 2k-3, \ldots, k-1, k+1, k).$$

Show that for every $s \in \{1, \ldots, n-1\}$ there is a unique i such that $s = x_{i+1} - x_i$. Define $L = [l_{ij}]$ over $\{0, 1, \ldots, n-1\}$ by $l_{ij} = x_i +_n x_j$, where $+_n$ denotes addition modulo n, and show that L is both row-complete and column-complete.)

Chapter 4

Finite Geometries and Designs

In this chapter we first present some basic facts about finite geometries. More precisely, we shall consider finite planes only. We show that the existence of a finite projective plane is equivalent to the exisence of a complete system of orthogonal Latin squares. We then move on to designs, one of the most important combinatorial configurations, which are straightforward generalisations of geometries. We characterize projective and affine planes as some special designs, but also show that other structures (such as Hadamard matrices) appear to be designes.

4.1 Projective planes

A *finite projective plane* is a pair $(\pi, \mathscr{L})$ where π is a nonempty set whose elements are called *points*, $\mathscr{L}$ is a set of nonempty subsets of π whose elements are called *lines*, and the following four conditions called the *axioms of projective planimetry* are satisfied:

(P1) For every pair of distinct points $A, B \in \pi$ there exists one and only one line $l \in \mathscr{L}$ such that $A \in l$ and $B \in l$.

(P2) For every pair of distinct lines $l, m \in \mathscr{L}$ there exists one and only one point $A \in \pi$ such that $l \ni A$ and $m \ni A$.

(P3) There exist four distinct points such that no three are on the same line.

(P4) π is finite.

A finite projective plane with 13 points and 13 lines is given in Fig. 4.1. Note that each line consists of exactly four points in this finite projective plane, and the curved lines in the figure are just used as an illustration.

For distinct points A and B, by $A \cdot B$ or just AB we denote the unique line that contains both A and B. Similarly, for distinct lines l and m, by $l \cdot m$ we denote the

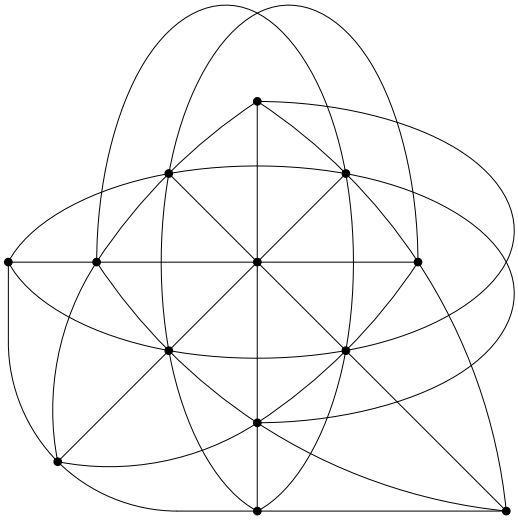
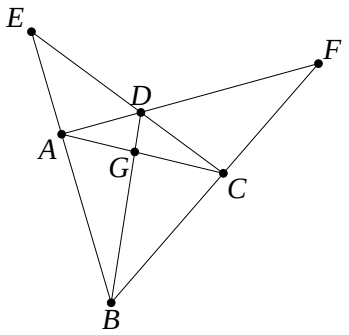


Figure 4.1: A finite projective plane with 13 points and 13 lines

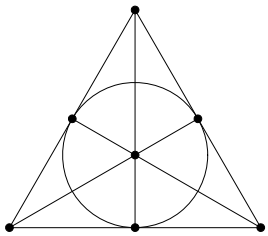
unique point that belongs to both l and m. If three or more points belong to the same line, we say that the points are *collinear*. If three or more lines pass through the same point, we say that the lines are *concurrent*.

Theorem 4.1 *Every finite projective plane has at least 7 points.*

Proof. Let A, B, C, D be the four distinct points in the plane which exist by (P3). Then $E = AB \cdot CD$ is distinct from A, B, C and D (e.g., if $E = A$ then A, D, C have to be collinear, which is impossible). Furthermore, let $F = AD \cdot BC$ and $G = AC \cdot BD$. It is easy to see that $F \notin \{A,B,C,D,E\}$ and $G \notin \{A,B,C,D,E,F\}$. Therefore, A, B, C, D, E, F, G are seven distinct points. □

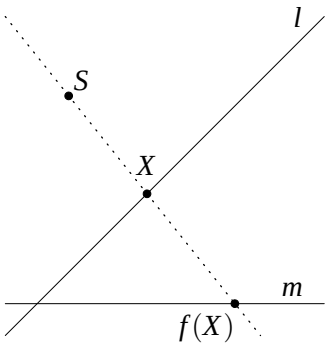


There exists a finite projective plane with exaclty seven points. It is called the *Fano plane* and clearly this is the smallest finite projective plane.



Theorem 4.2 *Every line in a finite projective plane contains at least three points.*

Theorem 4.3 *All lines in a projective plane have the same number of points.*



Proof. Let l and m be two distinct lines in a projective plane. Take a point S such that $S \notin l \cup m$ and define $\varphi_S : l \to m$ by $\varphi_S(X) = SX \cdot m$. Then φ_S is bijective and hence $|l| = |m|$. □

Definition 4.4 The *order* of a projective plane $(\pi, \mathscr{L})$ is a positive integer q such that $|l| = q+1$ for all $l \in \pi$.

So, the order of the Fano plane is 2, while the order of the projective plane in Fig. 4.1 is 3. One can easily show that in a projective plane of order q every point belongs to precisely $q+1$ lines (Exercise 4.13).

Theorem 4.5 *Let $(\pi, \mathscr{L})$ be a finite projective plane of order q. Then $|\pi| = |\mathscr{L}| = q^2+q+1$.*

Proof. Let S be a point in the plane. Then there are $q+1$ lines $l_1, \ldots, l_{q+1}$ that contain S. Now $|l_1| + \ldots + |l_{q+1}| = (q+1)^2$. Note that every point in the plane appears only once in this sum, except for S which was counted $q+1$ times, once for each line. Therefore, the number of points in π is $(q+1)^2 - q = q^2+q+1$.

For $A \in \pi$ let $\mathscr{L}_A$ denote the set of all lines that contain A. Then $\sum_{A \in \pi} |\mathscr{L}_A| = (q^2+q+1)(q+1)$. In this sum each line was counted $q+1$ times, once for each of its points. Therefore, $|\mathscr{L}| = q^2+q+1$. □

The axioms of the finite projective plane (projective space of dimension 2) can easily be extended to alow for higher dimensional projective spaces. Projective space of dimension d and order q is denoted by $\mathrm{PG}(d,q)$. So, projective plane of order q is $\mathrm{PG}(2,q)$ and in particular the Fano plane is just $\mathrm{PG}(2,2)$. Higher dimensional projective spaces have many properties that resemble the projective plane, e.g. an appropriate form of the Duality Principle is always valid (see Exercise 4.15), or $|\mathrm{PG}(d,q)| = \dfrac{q^{d+1}-1}{q-1}$.

Finite projective geometry is a source of very hard problems, e.g., it was shows only recently that a $\mathrm{PG}(2,10)$ does not exist (a long computation by Lam, Swierz, Thiel in 1989). The question for $\mathrm{PG}(2,12)$ is still unresolved. On the other hand, we know that if q is a prime power then there exists a unique projective plane of order q. We show the existence of such planes in the next section.

4.2 Affine planes

A *finite affine plane* is a pair $(\alpha, \mathscr{L})$ where α is a nonempty set whose elements are called *points*, $\mathscr{L}$ is a set of nonempty subsets of α whose elements are called *lines*, and the following four conditions called the *axioms of affine planimetry* are satisfied:

(A1) For every pair of distinct points $A, B \in \alpha$ there exists one and only one line $l \in \mathscr{L}$ such that $A \in l$ and $B \in l$.

(A2) For every line l and every point $A \notin l$ there is a unique line m such that $A \in m$ and $l \cap m = \varnothing$.

(A3) There exist three distinct points not on the same line.

(A4) α is finite.

We say that lines l and m are *parallel* and write $l \parallel m$ if $l = m$ or $l \cap m = \varnothing$. Axiom (A2) is therefore called the Parallel Postulate. It is easy to see that $\parallel$ is an equivalence relation on $\mathscr{L}$ and hence the lines in $\mathscr{L}$ can be divided into equivalence classes of parallel lines, called *parallel pencils*. An affine plane with 9 points is depicted in Fig. 4.2. It has 9 points, 12 lines and 4 parallel pencils one of which is outlined in the figure. Each parallel pencil in this geometry consists of three parallel lines.

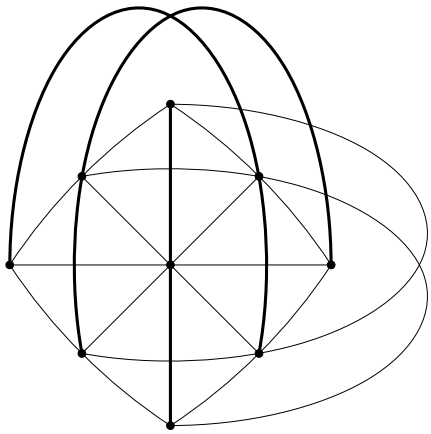


Figure 4.2: A finite affine plane with 9 points and 12 lines

Lemma 4.6 *Let $(\alpha, \mathscr{L})$ be a finite affine plane.*

(a) If $a, b \in \mathscr{L}$ and $a \nparallel b$ then a and b have a unique common point.

(b) α has at least four points such that no three are on the same line.

We are now going to show that affine planes are closely related to projective planes. Let $(\pi,\mathscr{L})$ be a projective plane and let $m \in \mathscr{L}$ be any line in π. Let $\pi - m = \pi \setminus m$ and define $\mathscr{L} - m$ by

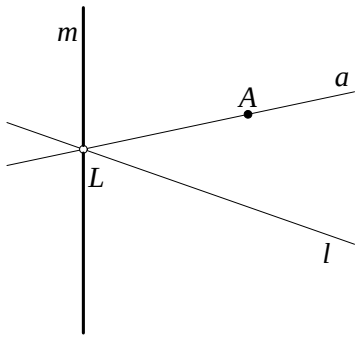
$$\mathscr{L} - m = \{l \setminus m : l \in \mathscr{L} \text{ and } l \neq m\}.$$

So, $(\pi - m, \mathscr{L} - m)$ is a structure obtained from the projective plane $(\pi,\mathscr{L})$ by removing the line m and all its points.

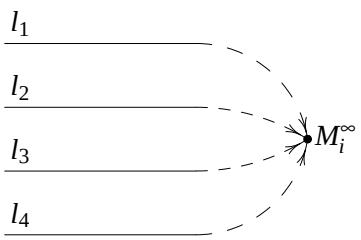
Theorem 4.7 *Let $(\pi,\mathscr{L})$ be a projective plane and let $m \in \mathscr{L}$ be arbitrary. Then $(\pi - m, \mathscr{L} - m)$ is an affine plane.*

Proof. Let us show that $(\pi - m, \mathscr{L} - m)$ satisfies (A1)–(A4). (A4) is obvious, while (A1) is a direct consequence of (P1). Let us show (A3). By (P3) there exist distinct points A, B, C, $D \in \pi$ such that no three are on the same line. If $|m \cap \{A,B,C,D\}| \leqslant 1$ then at least three of these four points lie outside m, that is in $\pi - m$ and we are done. Assume now that $|m \cap \{A,B,C,D\}| = 2$, say, $A, B \in m$. Then $C, D \notin m$. It is easy to see that $E = AC \cdot BD$ does not lie on m and that $\{C,D,E\} \subseteq m$ are three distinct points.

Finally, let us show (A2). Take any $l \in \mathscr{L} - m$ and any $A \in \pi - m$ such that $A \notin l$. By the construction of $\mathscr{L} - m$ there is an $l' \in \mathscr{L}$ such that $l = l' \setminus m$. Since $l' \neq m$ the two lines intersect and let $L = l' \cdot m$. Clearly, $l' = l \cup \{L\}$. Put $a' = AL$ and $a = a' \setminus m \in \mathscr{L} - m$. Then $a \ni A$ and $a \parallel l$. Let $b \in \mathscr{L} - m$ be any other line parallel to l that contains A and take $b' \in \mathscr{L}$ such that $b = b' \setminus m$. Now, b' and l' have an intersection in π while b and l do not. Therefore, the intersection of b' and l' is a point on l' that does not belong to l, and hence $b' \cdot l' = L$. This shows that $b' = AL = a'$ whence $b = a$, and the parallel through A is unique. □



This construction is based on the idea that "parallel lines intersect at infinity", where "infinity" is the line m. Using the same idea we can reverse the construction: starting from an affine plane, for each parallel pencil we add one new "point at infinity" and assume that all these "points at infinity" lie on a new line, the "line at infinity". Each old line goes through all its old points plus the one new point corresponding to its parallel pencil.



More precisely, let $(\alpha, \mathscr{L})$ be a finite affine plane and let $m^\infty = \mathscr{L}/\| = \{M_1^\infty, M_2^\infty, \ldots, M_k^\infty\}$ be the set of parallel pencils (so that each M_i^∞ is a set of lines parallel to each other). Just for the record note that $M_i^\infty \notin \alpha$ for all i since these are sets of lines, and that $m^\infty \notin \mathscr{L}$. Let $\alpha^* = \alpha \cup \{M_1^\infty, \ldots, M_k^\infty\}$. For each line $l \in \mathscr{L}$ let $l^* = l \cup \{M_i^\infty\}$ where $M_i^\infty = l/\|$, and let $\mathscr{L}^* = \{l^* : l \in \mathscr{L}\} \cup \{m^\infty\}$, Fig 4.3.

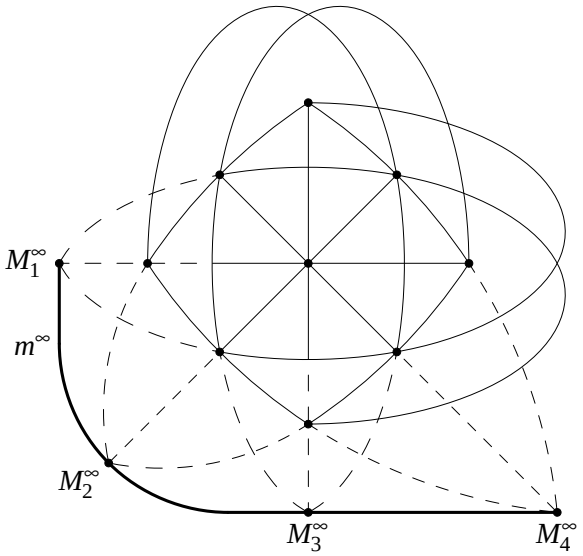


Figure 4.3: Extending an affine plane to a projective plane

Theorem 4.8 *If $(\alpha, \mathscr{L})$ is an affine plane, then $(\alpha^*, \mathscr{L}^*)$ is a projective plane.*

Proof. Let us show that $(\alpha^*, \mathscr{L}^*)$ satisfies (P1)–(P4). Points M_i^∞ will be referred to as points at infinity. Firstly, note that (P4) is obvious and (P3) is a direct consequence of Lemma 4.6 (b).

To show that (P1) holds, take any $A, B \in \alpha^*$, $A \neq B$. If $A, B \in \alpha$ then there is a unique line $l \in \mathscr{L}$ such that $A, B \in l$ so l^* is a unique line in $\mathscr{L}^*$ that contains A and B. If $A = M_i^\infty$ and $B = M_j^\infty$ for some $i \neq j$ then clearly m^∞ is the only line in $\mathscr{L}^*$ that contains A and B since all other lines in $\mathscr{L}^*$ contain precisely one point at infinity. Finally, let $A \in \alpha$ and $B = M_i^\infty$ for some i. Recall that M_i^∞ is a parallel pencil, so take any $l \in M_i^\infty$. If $A \in l$ then l^* is the unique line in $\mathscr{L}^*$ that contains A and B. If $A \notin l$ let a be the unique line in $\mathscr{L}$ parallel to l which passes through A. Then a^* is the unique line in $\mathscr{L}^*$ that contains A and $B = M_i^\infty$.

To show that (P2) holds, take any $a, b \in \mathscr{L}^*$, $a \neq b$. If one of them is m^∞, say $b = m^\infty$, then $a = a_0^*$ for some $a_0 \in \mathscr{L}$ and $a \cdot b = M_i^\infty$ where M_i^∞ is the parallel pencil

of a_0. If $a \neq m^\infty \neq b$ then $a = a_0^*$ for some $a_0 \in \mathscr{L}$ and $b = b_0^*$ for some $b_0 \in \mathscr{L}$. If $a_0 \parallel b_0$ then they have the same point at infinity M_i^∞ so $a \cdot b = M_i^\infty$. Otherwise $a_0 \cap b_0 \neq \varnothing$. Since $a_0 \neq b_0$, it follows by Lemma 4.6 (a) that $a_0 \cap b_0 = \{A\}$ for some $A \in \alpha$, so $a \cdot b = A$. The uniqueness of the point of intersection is immediate. □

Corollary 4.9 *For every finite affine plane $(\alpha, \mathscr{L})$ there is a positive integer q such that $|\alpha| = q^2$, $|\mathscr{L}| = q(q+1)$, each line has precisely q points, each point belongs to precisely $q+1$ lines, each parallel pencil has q lines and there are $q+1$ parallel pencils.*

Proof. Let q be the order of the finite projective plane $(\alpha^*, \mathscr{L}^*)$. Then clearly $|\alpha| = |\alpha^*| - |m^\infty| = (q^2+q+1) - (q+1) = q$, $|\mathscr{L}| = |\mathscr{L}^*| - 1 = q^2 + q$ and for each $l \in \mathscr{L}$ we have $|l| = |l^*| - 1 = q$. Since each point in α^* is incident with $q+1$ lines, and since no $A \in \alpha$ is incident to m^∞ we get that each point from α belongs to precisely $q+1$ lines from $\mathscr{L}$. The number of parallel pencils is $|m^\infty| = q+1$. Since each point in α^* belongs to $q+1$ lines, the same holds for points at infinity. But one of these $q+1$ lines that pass through a point at infinity is $m^\infty \notin \mathscr{L}$, so each parallel pencil of lines in $\mathscr{L}$ consists of q lines. □

Definition 4.10 The integer q from Corollary 4.9 is called the *order* of the finite affine plane $(\alpha, \mathscr{L})$.

We conclude the section on affine planes by showing that the existence of projective and affine planes of a given order is equivalent to the existence of a complete system of orthogonal Latin squares.

Theorem 4.11 *The following statements are equivalent for every integer $q \geqslant 2$:*

(1) There exists a finite projective plane of order q.

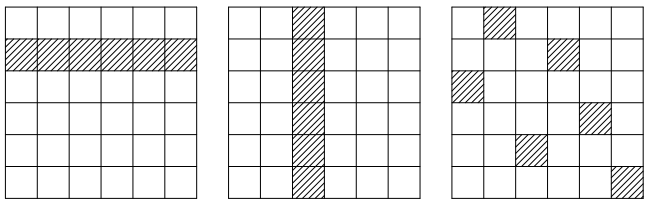
(2) There exists a finite affine plane of order q.

(3) There exists a complete system of orthogonal Latin squares of order q.

Proof. The equivalence of (1) and (2) has been established in Theorems 4.7 and 4.8. Let us show the equivalence of (2) and (3).

(3) $\Rightarrow$ (2): Let $L^1 = [l_{ij}^1], \ldots, L^{q-1} = [l_{ij}^{q-1}]$ be a complete system of orthogonal Latin squares over $Q = \{1, 2, \ldots, q\}$. For the points we take "coordinates", that is, we let $\alpha = Q^2 = \{(i, j) : i, j \in Q\}$ and (A4) is obviously satisfied. For $\mathscr{L}$ we take three types of lines, Fig. 4.4:

- horizontal lines: $h_i = \{(i, j) : j \in Q\}$, for each $i \in Q$,



(a) (b) (c)

Figure 4.4: The three types of lines: (a) horizontal, (b) vertical, (c) skew

- vertical lines: $v_j = \{(i,j) : i \in Q\}$, for each $j \in Q$, and
- skew lines: $s_a^k = \{(i,j) : l_{ij}^k = a\}$, for each $k \in \{1,\ldots,q-1\}$ and each $a \in Q$.

Let us first note that no skew line contains two points from the same row, or two points from the same column (if s_a^k contains (i,j_1) and (i,j_2) then $l_{ij_1}^k = a = l_{ij_2}^k$, which is impossible since L^k is a Latin square).

Let us show that (A1) is valid. Take two distinct points (i_1,j_1) and (i_2,j_2). If $i_1 = i_2 = i$ then h_i contains both points. No other horizontal line contains these two points, and clearly no vertical line and no skew line can contain two points from the same row. So, h_i is the only line that contains the two points. The proof is analogous in case $j_1 = j_2$. Assume now that $i_1 \neq i_2$ and $j_1 \neq j_2$. To show that there is a unique skew line s_a^k that contains both (i_1,j_1) and (i_2,j_2) it suffices to show that there is a unique k and a unique a such that $l_{i_1j_1}^k = l_{i_2j_2}^k = a$. But this is true due to Homework 4.4.

To show (A2) take any line p and a point (i,j) not on the line. If $p = h_{i'}$ for some i' then h_i contains (i,j) and it is parallel to $h_{i'}$. It is easy to see that h_i is the only such line: neither the vertical line v_j nor skew lines that contain (i,j) are parallel to $h_{i'}$. The proof is analogous in case p is a vertical line. Assume now that p is a skew line s_a^k for some k and a. By the assumption, $(i,j) \notin s_a^k$ so $b = l_{ij}^k \neq a$. Now, s_b^k contains (i,j) and it is parallel to s_a^k (if $(u,v) \in s_a^k \cap s_b^k$ then $b = l_{uv}^k = a$, which is not the case). Let us show that no other line through (i,j) is parallel to s_a^k. Clearly, no horizontal and no vertical line is parallel to s_a^k and let us show that no $s_c^m \neq s_b^k$ that contains (i,j) is parallel to s_a^k. Take any s_c^m such that $(i,j) \in s_c^m$ and $s_c^m \neq s_b^k$. Then $m \neq k$ and hence $L^k \perp L^m$. Therefore, there exists an (i_0,j_0) such that $(l_{i_0j_0}^k, l_{i_0j_0}^m) = (a,c)$. So, $(i_0,j_0) \in s_a^k \cap s_c^m$, i.e. $s_a^k \nparallel s_c^m$, and hence s_b^k is the only line parallel to s_a^k that contains (i,j).

To see that (A3) is valid, take $(1,1)$, $(1,2)$ and $(2,1)$. No horizontal or vertical line contains all three points, and no skew line contains $(1,1)$ and $(1,2)$.

$(2) \Rightarrow (3)$: Let $(\alpha, \mathscr{L})$ be a finite affine plane of order q. It has q^2 points and $q+1$ parallel pencils. In order to produce Latin squares out of this configuration,

we shall first introduce "coordinates" as follows. Take two distinct parallel pencils $\mathscr{H} = \{h_1, \ldots, h_q\}$ and $\mathscr{V} = \{v_1, \ldots, v_q\}$. The two pencils will serve as "horizontal" lines and "vertical" lines. Every point $A \in \alpha$ lies on a unique "horizontal" line h_i and a unique "vertical" line v_j, so we say that (i, j) are the coordinates of A. Let $f : \{1, \ldots, q\}^2 \to \alpha$ be the mapping that takes coordinates to their respective points. Clearly, f is a bijection. Each of the remaining $q-1$ parallel pencils $\mathscr{P}_m = \{a_{m1}, \ldots, a_{mq}\}$ determines a matrix $L^m = [l^m_{ij}]_{q\times q}$ over $\{1, \ldots, q\}$ as follows: $l^m_{ij} = k$ if $f(i, j) \in a_{mk}$. Let us show that L^m is a Latin square. Suppose that some k appears twice in a row of L^m, say, $l^m_{ij_1} = k$ and $l^m_{ij_2} = k$. Then $f(i, j_1) \in a_{mk}$ and $f(i, j_2) \in a_{mk}$, and hence h_i and a_{mk} have two distinct points in common: $f(i, j_1)$ and $f(i, j_2)$. Therefore, $h_i = a_{mk}$, but this implies that two distinct parallel pencils $\mathscr{H}$ and $\mathscr{P}_m$ have a line in common, which is not possible. Therefore, for each of the remaining $q-1$ parallel pencils $\mathscr{P}_1, \ldots, \mathscr{P}_{q-1}$ the matrices $L^1, \ldots, L^{q-1}$ are Latin squares. Finally, let us show that $L^m \perp L^k$ for $m \neq k$. Take any i_1, j_1, i_2, j_2 and suppose that $(l^m_{i_1 j_1}, l^k_{i_1 j_1}) = (l^m_{i_2 j_2}, l^k_{i_2 j_2}) = (t, u)$. Then $f(i_1, j_1) \in a_{mt}$, $f(i_2, j_2) \in a_{mt}$, $f(i_1, j_1) \in a_{ku}$ and $f(i_2, j_2) \in a_{ku}$. If $(i_1, j_1) \neq (i_2, j_2)$ then both a_{mt} and a_{ku} contain these two distinct points, so $a_{mt} = a_{ku}$. But if this is true, then the parallel pencils $\mathscr{A}_m$ and $\mathscr{A}_k$ have a line in common, which contradicts the assumption $m \neq k$. Therefore, $(i_1, j_1) = (i_2, j_2)$. This shows that every pair $(t, u) \in \{1, \ldots, q\}^2$ appears at most once when we overlay L^m with L^k, whence follows that $L^m \perp L^k$. □

4.3 Designs

Assume that we wish to compare v varieties of wine. In order to make the testing procedure as fair as possible it is natural to require that each person participating tastes the same number (say k) of varieties being tested so that each person's opinion has the same weight, and each pair of varieties of wine is compared by the same number of persons (say λ) so that each variety of wine gets the same treatment. One possibility would be to let everyone taste all the varieties. But if v is large, this is very impractical. We would like to design a fair experiment so that $k < v$.

Definition 4.12 Let X be a finite set with v elements, and let $2 \leqslant k < v$ and $\lambda > 0$. A pair $(X, \mathscr{B})$ where $\mathscr{B}$ is collection of distinct subsets of X is called a (v, k, λ)-*design* if

- each set in $\mathscr{B}$ contains exactly k elements, and
- each 2-element subset of X is contained in exactly λ sets in $\mathscr{B}$.

The elements of X are called *vertices* and the sets in $\mathscr{B}$ are called the *blocks* of the design. The number of blocks of $(X, \mathscr{B})$ is usually denoted by b.

Example 4.13 (a) Let $X = \{1,\dots,7\}$ and $\mathscr{B} = \{\{1,2,4\}, \{2,3,5\}, \{3,4,6\}, \{4,5,7\}, \{5,6,1\}, \{6,7,2\}, \{7,1,3\}\}$. Then $(X,\mathscr{B})$ is a $(7,3,1)$-design.

(b) Every finite projective plane of order q is a $(q^2+q+1, q+1, 1)$-design.

(c) Every finite affine plane of order q is a $(q^2, q, 1)$-design.

Theorem 4.14 *Let $(X,\mathscr{B})$ be a (v,k,λ)-design. Then each vertex of the design occurs in r blocks, where r satisfies the following two equalities:*

$$r(k-1) = \lambda(v-1) \qquad \textit{and} \qquad bk = vr.$$

Proof. Consider a $a \in X$ and assume that a occurs in r_a blocks. Let

$$\mathscr{H} = \{(x,B) : a \neq x, B \in \mathscr{B}, \{a,x\} \subseteq B\},$$

and let us find $|\mathscr{H}|$. There are $v-1$ possibilities to choose x (since $x \neq a$), and once we have chosen $x \neq a$, there are λ blocks that contain both a and x. Therefore, $|\mathscr{H}| = (v-1)\lambda$. On the other hand, there are r_a blocks that contain a and each block has k elements. Therefore, in each of the r_a blocks that contain a there are $k-1$ possibilities to choose $x \neq a$, so $|\mathscr{H}| = r_a(k-1)$. So, we see that $r_a(k-1) = (v-1)\lambda$, i.e. r_a is uniquely determined by by v, k and λ. This means that $r_{a_1} = r_{a_2}$ for all $a_1, a_2 \in X$ and we have the first equality. For the second equality, let

$$\mathscr{H}' = \{(x,B) : B \in \mathscr{B}, x \in B\}.$$

Since the design has b blocks and each block has k elements, $|\mathscr{H}'| = bk$. On the other hand, there are v ways to choose x, and once we fix x, there are r blocks that contain it. Therefore, $|\mathscr{H}'| = vr$ and finally $bk = vr$. □

This theorem also shows that b and r are uniquely determined by v, k and λ. A main problem in design theory is to determine for which values of v, k and λ there is a (v,k,λ)-design. Certainly, designs do not exist for every choice of v, k and λ.

We have seen in Example 4.13 that every affine plane is a $(n^2, n, 1)$-design. But the converse is also true:

Theorem 4.15 *A (v,k,λ)-design is an affine plane if and only if this is a $(n^2,n,1)$-design for some $n \geqslant 2$.*

Proof. Every affine plane of order q is a $(q^2,q,1)$-design. To show the implication from right to left take any $(n^2,n,1)$-design $(X,\mathscr{B})$, $n \geqslant 2$, and let us show that it satisfies (A1)–(A4). (A1) is satisfied since $(X,\mathscr{B})$ is a design, while (A3) and (A4) are obvious, so let us show (A2). Take any block $B = \{a_1,\dots,a_n\} \in \mathscr{B}$ and any $x \notin B$. For every $i \in \{1,\dots,n\}$ there is a unique block B_i containing x and a_i, and

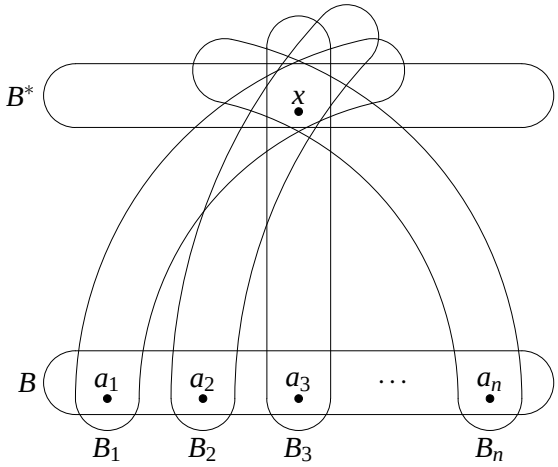


Figure 4.5: The proof of Theorem 4.15

clearly $B_i \neq B_j$ whenever $i \neq j$. According to Theorem 4.14 we have $r = n+1$, i.e. every point lies in $n+1$ blocks. Now, there are n blocks that contain x and intersect B, so the remaining $(n+1)$-th block containing x has to be disjoint from B, Fig. 4.5. Therefore, this is the unique block disjoint from B that contains x. □

The *incidence matrix* of a design $(X, \mathscr{B})$ where $X = \{x_1, \ldots, x_v\}$ and $\mathscr{B} = \{B_1, \ldots, B_b\}$ is the incidence matrix of of the family $\mathscr{B}$, that is, an $b \times v$ matrix $A = [a_{ij}]$ over $\{0,1\}$ such that

$$a_{ij} = \begin{cases} 1, & B_i \ni x_j \\ 0, & \text{otherwise.} \end{cases}$$

Lemma 4.16 *Let A be an incidence matrix of a (v,k,λ)-design, let E be the identity matrix and J a square matrix in which every entry is 1. Then*

$$A^\top A = \begin{bmatrix} r & \lambda & \lambda & \ldots & \lambda \\ \lambda & r & \lambda & \ldots & \lambda \\ \lambda & \lambda & r & \ldots & \lambda \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \lambda & \lambda & \lambda & \ldots & r \end{bmatrix} = (r-\lambda)E + \lambda J,$$

Theorem 4.17 (Fisher's inequality) *If there exists a (v,k,λ)-design then $b \geqslant v$.*

Proof. Let us calculate $\det(A^\top A)$. We first subtract the first row from the others:

$$\det(A^\top A) = \begin{vmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda & r & \lambda & \dots & \lambda \\ \lambda & \lambda & r & \dots & \lambda \\ \dots & \dots & \dots & \dots & \dots \\ \lambda & \lambda & \lambda & \dots & r \end{vmatrix} = \begin{vmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda - r & r - \lambda & 0 & \dots & 0 \\ \lambda - r & 0 & r - \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \lambda - r & 0 & 0 & \dots & r - \lambda \end{vmatrix}$$

and then add all other columns to the first column:

$$\det(A^\top A) = \begin{vmatrix} r + (v-1)\lambda & \lambda & \lambda & \dots & \lambda \\ 0 & r - \lambda & 0 & \dots & 0 \\ 0 & 0 & r - \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & r - \lambda \end{vmatrix}.$$

Therefore, $\det(A^\top A) = (r + (v-1)\lambda)(r-\lambda)^{v-1}$. Now from $r(k-1) = \lambda(v-1)$ we obtain that $\det(A^\top A) = rk(r-\lambda)^{v-1}$. We have assumed that $k < v$, so $r(k-1) = \lambda(v-1)$ implies $r > \lambda$ which, together with $v > 2$ yields $\det(A^\top A) > 0$.

Assume now that $b < v$. Then there are fewer rows than columns in A. Let A_1 be a $v \times v$ matrix obtained by adding $v - b$ rows of zeros to A. It is easy to see that $A_1^\top A_1 = A^\top A$. But since A_1 is a square matrix, the product rule for determinants implies that

$$\det(A^\top A) = \det(A_1^\top A_1) = \det(A_1^\top)\det(A_1) = 0$$

because there is at least one row of zeros in A_1. This contrardicts $\det(A^\top A) > 0$ and hence $b \geqslant v$. □

A (v,k,λ)-design is *symmetric* if $b = v$, i.e., its incidence matrix is a square matrix. Note that the incidence matrix of a symmetric design does not have to be a symmetric matrix!

Lemma 4.18 *In a symmetric (v,k,λ)-design we have $k = r$ and $\lambda < k$.*

In a symmetric design we denote $k - \lambda$ by n. A symmetric design is said to be *trivial* if $n = 1$.

Lemma 4.19 *In a trivial symmetric design every $(v-1)$-element subset of X is a block. For a nontrivial symmetric design we have $k \leqslant v - 2$.*

Theorem 4.20 *Let A be an incidence matrix of a symmetric (v,k,λ)-design. Then $A^\top A = AA^\top$ and the intersection of any two distinct blocks of the design has λ elements.*

Proof. Let J be the $v \times v$ matrix whose all entries are 1. The following is clearly true: $AJ = JA = kJ$, $A^\top J = JA^\top = kJ$ since $k = r$, and $J^2 = vJ$. Using Lemma 4.16, $k = r$ and $A^\top J = JA = kJ$ we see that

$$\left(A^\top - \sqrt{\frac{\lambda}{v}}J\right)\left(A + \sqrt{\frac{\lambda}{v}}J\right) = A^\top A + \sqrt{\frac{\lambda}{v}}(A^\top J - JA) - \frac{\lambda}{v}J^2$$
$$= A^\top A - \lambda J = (k-\lambda)E.$$

Since $k > \lambda$, the above calculation means that $A^\top - \sqrt{\frac{\lambda}{v}}J$ has an inverse and the inverse is $\frac{1}{k-\lambda}\left(A + \sqrt{\frac{\lambda}{v}}J\right)$. Now,

$$\frac{1}{k-\lambda}\left(A + \sqrt{\frac{\lambda}{v}}J\right)\left(A^\top - \sqrt{\frac{\lambda}{v}}J\right) = E$$
$$AA^\top + \sqrt{\frac{\lambda}{v}}(JA^\top - AJ) - \frac{\lambda}{v}J^2 = (k-\lambda)E$$
$$AA^\top - \lambda J = (k-\lambda)E$$
$$AA^\top = (k-\lambda)E + \lambda J,$$

so Lemma 4.16 yields $AA^\top = A^\top A$. The second part of the theorem follows immediately from $AA^\top = (k-\lambda)E + \lambda J$. □

Let $(X, \mathscr{B})$ be a (v,k,λ)-design such that $b - 2r + \lambda > 0$, and let $\overline{\mathscr{B}} = \{X \setminus B : B \in \mathscr{B}\}$. Then $(X, \overline{\mathscr{B}})$ is a $(v, v-k, b-2r+\lambda)$-design called the *complement* of $(X, \mathscr{B})$.

Not every design has a complement simply because it may happen that $b - 2r + \lambda < 0$. However, every nontrivial symmetric (v,k,λ)-design has a complement and its complement is a symmetric $(\overline{v}, \overline{k}, \overline{\lambda})$-design where $\overline{v} = v$, $\overline{k} = v - k$ and $\overline{\lambda} = v - 2k + \lambda$. Moreover, $\overline{n} = n$, where $\overline{n} = \overline{k} - \overline{\lambda}$ (Exercise 4.21).

Theorem 4.21 *If a nontrivial symmetric (v,k,λ)-design exists, then*

$$4n - 1 \leqslant v \leqslant n^2 + n + 1.$$

Proof. Suppose there exists a nontrivial symmetric (v,k,λ)-design. Then it has a complement and it is a $(\overline{v}, \overline{k}, \overline{\lambda})$-design where $\overline{v} = v$, $\overline{k} = v - k$ and $\overline{\lambda} = v - 2k + \lambda$. Let us calculate $\lambda\overline{\lambda}$:

$$\lambda\overline{\lambda} = \lambda(v - 2k + \lambda) = \lambda(v-1) + \lambda - 2k\lambda + \lambda^2$$
$$= k(k-1) + \lambda - 2k\lambda + \lambda^2 \quad [\text{since } \lambda(v-1) = r(k-1) \text{ and } r = k]$$
$$= (k-\lambda)^2 - (k-\lambda) = n^2 - n.$$

From $\lambda \geqslant 1$ and $\overline{\lambda} \geqslant 1$ we get

$$\begin{aligned} 0 \leqslant (\lambda-1)(\overline{\lambda}-1) &= \lambda\overline{\lambda} - (\lambda+\overline{\lambda}) + 1 = (n^2-n) - (v-2k+2\lambda) + 1 \\ &= (n^2-n) - (v-2(k-\lambda)) + 1 = n^2 - n - v + 2n + 1 \end{aligned}$$

whence $v \leqslant n^2+n+1$. For the lower bound let us first note that $(x+y)^2 \geqslant 4xy$ for every pair of reals x and y. Now for $x=\lambda$ and $y=\overline{\lambda}$ we obtain

$$(\lambda+\overline{\lambda})^2 \geqslant 4\lambda\overline{\lambda} = 4n(n-1) > (2n-2)^2$$

since $n \geqslant 2$. Having in mind that $\lambda+\overline{\lambda} \geqslant 1$ and $2n-2 \geqslant 1$ taking the square root of the above inequality yields

$$v - 2n = \lambda + \overline{\lambda} > 2n-2$$

i.e. $v-2n \geqslant 2n-1$, so $v \geqslant 4n-1$. □

We say that a nontrivial symmetric (v,k,λ)-design is *maximal* if $v = n^2+n+1$ and that it is *minimal* if $v = 4n-1$. We conclude this section by showing that maximal symmetric designs correspond to finite geometries. We give the interpretation of minimal symmetric designs in the next section.

Theorem 4.22 *A (v,k,λ)-design is a finite projective plane if and only if it is a maximal nontrivial symmetric $(v,k,1)$-design.*

Proof. ($\Rightarrow$) Let $(X,\mathscr{B})$ be a projective plane of order q. Then it is a $(q^2+q+1, q+1, 1)$-design. Since a projective plane has the same number of points and lines, it is a symmetric design where $n = k-\lambda = (q+1)-1 = q > 1$. Hence, this is a nontrivial symmetric design whose maximality is obvious since $v = n^2+n+1$.

($\Leftarrow$) Let $(X,\mathscr{B})$ be a maximal nontrivial symmetric $(v,k,1)$-design. Then $v = n^2+n+1$ and $k = n+\lambda = n+1$. Hence, this is a $(n^2+n+1, n+1, 1)$-design. Let us show that this is a projective plane. (P4) is trivially satisfied, while (P1) follows from $\lambda = 1$. From Theorem 4.20 we know that in a nontrivial symmetric design the intersection of every two distinct blocks has λ elements, so (P2) is valid since $\lambda = 1$. Finally, to show that (P3) holds, take any two distinct blocks $B, B' \in \mathscr{B}$, $B \neq B'$. Then $|B \cap B'| = \lambda = 1$ and let x be the only element of $B \cap B'$. Since the design is nontrivial, $n \geqslant 2$ so $|B| = |B'| = n+1 \geqslant 3$. Take $y_1, y_2 \in B \setminus \{x\}$ such that $y_1 \neq y_2$ and take $y_3, y_4 \in B' \setminus \{x\}$ such that $y_3 \neq y_4$. Now, y_1, y_2, y_3 and y_4 are four distinct points no three in the same block. □

4.4 Hadamard matrices

One of the important results due to the famous French mathmatician Jacques Salomon Hadamard (1865–1963) is the answer to the following question: Let $A = [a_{ij}]$ be a real square matrix such that $|a_{ij}| \leqslant 1$; how large can $|\det(A)|$ be?

Theorem 4.23 (Hadamard) *Let $A = [a_{ij}]$ be a real $n \times n$ matrix such that $|a_{ij}| \leqslant 1$. Then $|\det(A)| \leqslant n^{n/2}$. The equality holds if and only if $a_{ij} = \pm 1$ for all i and j, and $A^\top A = nE$.*

Proof. (Sketch) It is well known that $|\det(A)|$ is the volume of the parallelepiped in n-dimensional Euclidean space whose sides are verctors that correspond to the columns of A. If $|a_{ij}| \leqslant 1$ for all i and j, then the Euclidean length of such vectors is at most $\sqrt{n}$. The volume of the parallelepiped is at most the product of the lengths of its edges, so $|\det(A)| \leqslant (\sqrt{n})^n = n^{n/2}$. The equality holds if and only if the edges of the parallelepiped are mutually orthogonal and of maximal length $\sqrt{n}$. The edges can achieve the length of $\sqrt{n}$ just in case $a_{ij} = \pm 1$ for all i and j, while the orthogonality requirement means that the scalar product of any two distinct columns in A is zero. Therefore, $A^\top A = nE$. □

Definition 4.24 An *Hadamard matrix of order n* is an $n \times n$ matrix H with entries ± 1 such that $H^\top H = nE$.

Note that $H^\top H = nE$ means that H is invertible and $H^{-1} = \frac{1}{n} H^\top$. Therefore, a matrix H with entries ± 1 is an Hadamard matrix if and only if $HH^\top = nE$. We say that an Hadamard matrix is *normalized* if its first row and its first column are $11\ldots1$.

Lemma 4.25 *If H is an Hadamard matrix and H' is a matrix obtained from H by multiplying a row or a column by -1, then H' is also an Hadamard matrix. Every Hadamard matrix H can be transformed to a normalized Hadamard matrix by multiplying some of its rows and columns by -1.*

Theorem 4.26 *Let H be a normalized Hadamard matrix of order $n > 1$. Then every row other than the first row on the matrix has $n/2$ entries equal to 1 and $n/2$ entries equal to -1. If $n > 2$ then any two rows other than the first row have exactly $n/4$ 1's in common. The analogous statements hold for columns.*

Consequently, if there exists an Hadamard matrix of order n then $n = 1$, $n = 2$ or $n \equiv 0 \pmod 4$.

Proof. The inner product of the first row with any other row is 0. Therefore, in any other row the number of entries equal to 1's and the number of entries equal to -1 are the same. Hence n is even.

Take any two rows $i \neq j$ other than the first row. Let s be the number of 1's these two rows have in common and let t be the number of -1's these two rows have in common:

$$\begin{array}{lcccccccccccccccc}
1^{\text{st}}\text{ row} & 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 \\
i^{\text{th}}\text{ row} & 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 & -1 & -1 & \dots & -1 & -1 & -1 & \dots & -1 \\
j^{\text{th}}\text{ row} & \underbrace{1} & \underbrace{1} & \underbrace{\dots} & \underbrace{1} & \underbrace{-1} & \underbrace{-1} & \underbrace{\dots} & \underbrace{-1} & \underbrace{1} & \underbrace{1} & \underbrace{\dots} & \underbrace{1} & \underbrace{-1} & \underbrace{-1} & \underbrace{\dots} & \underbrace{-1} \\
 & \multicolumn{4}{c}{s} & \multicolumn{4}{c}{n/2-s} & \multicolumn{4}{c}{n/2-t} & \multicolumn{4}{c}{t}
\end{array}$$

The number of 1's in the third row is $n/2$ so $s+(n/2-t)=n/2$, whence follows that $s=t$. The inner product of the second and the third row is 0 so $s+t=(n/2-s)+(n/2-t)$ and since $s=t$ we immediately obtain $s=n/2-s$. Therefore, $s=n/4$ and hence n is divisible by 4. □

The existence of Hadamard matrices is still an open problem. However, it is easy to show that an Hadamard matrix of order n exists whenever n is a power of 2. It is conjectured that an Hadamard matrix of order $n \geqslant 4$ exists if and only if n is divisible by 4.

Example 4.27 Let us show that an Hadamard matrix of order n exists whenever $n=2^k$. Clearly, $H_1=[1]$ and $H_2=\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ are Hadamard matrices of order 1 and 2. Now, if H_n is an Hadamard matrix of order n it is easy to show that $H_{2n}=\begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$ is an Hadamard matrix of order $2n$. Therefore, there exist Hadamard matrices of order n whenever n is a power of 2.

Each normalized Hadamard matrix gives rise to a $(4s-1,2s-1,s-1)$-design as follows. Remove the first row and the first column of the matrix (they carry no information so we don't need them) and in the remaining truncated matrix replace each -1 with 0. The new matrix is an incidence matrix of a $(4s-1,2s-1,s-1)$-design called the *Hadamard design*. The reverse construction shows that every Hadamard design gives rise to an Hadamard matrix.

Theorem 4.28 *A (v,k,λ)-design is an Hadamard design with $v \geqslant 7$ if and only if it is a minimal nontrivial symmetric $(v,k,n-1)$-design (here, n is not the order of the Hadamard matrix but a parameter of the symmetric design).*

Proof. ($\Rightarrow$) If a design is an Hadamard design arising from an Hadamard matrix of order $4s$ then its parameters are $v=4s-1$, $k=2s-1$ and $\lambda=s-1$. According to Theorem 4.14, $r=2s-1$ and $b=4s-1=v$, so the design is symmetric. Then

$n = k - \lambda = s \geqslant 2$ since $v \geqslant 7$. So it is minimal ($v = 4n-1$), nontrivial ($n \geqslant 2$) and $\lambda = s - 1 = n - 1$.

($\Leftarrow$) In a minimal symmetric $(v,k,n-1)$-design we have $v = 4n-1$ and $k = n + \lambda = 2n - 1$. So this is a $(4n-1, 2n-1, n-1)$-design, i.e. an Hadamard design. Since the design is nontrivial, $n \geqslant 2$ and hence $v \geqslant 7$. □

Corollary 4.29 *There exists an Hadamard matrix of order $n \geqslant 8$ if and only if n is divisible by 4 and there exists an Hadamard $(n-1, \frac{1}{2}n-1, \frac{1}{4}n-1)$-design.*

Homework

4.1. (a) Prove Theorem 4.2. (Hint: take any line and consider the four points that exist by (P3).)

(b) Show that every point in a finite projective plane belongs to at least three distinct lines.

4.2. (a) Show that for every pair of lines l, m there is a point S such that $S \notin l \cup m$.

(b) Show that the mapping φ_S in the proof of Theorem 4.3 is bijective.

4.3. Show Lemma 4.6. (Hint: for (a) use (A1); for (b) take three points A, B, C not on the same line using (A3), let l be a line through C parallel to AB and let m be a line through B parallel to AC; show that l and m have a point of intersection D and that A, B, C and D are the points we have been looking for.)

4.4. Let $L^1 = [l^1_{ij}], \ldots, L^{q-1} = [l^{q-1}_{ij}]$ be a complete system of orthogonal Latin squares over $\{1,\ldots,q\}$. Let (i_1, j_1) and (i_2, j_2) be two pairs of indices such that $i_1 \neq i_2$ and $j_1 \neq j_2$. Show that there is a unique k and a unique a such that $l^k_{i_1 j_1} = l^k_{i_2 j_2} = a$. (Hint: Define $\mathscr{A}$ and $\mathscr{B}$ as follows:

$$\mathscr{A} = \big\{(k, a, \{(u_1,v_1),(u_2,v_2)\}) : k \in \{1,\ldots,q-1\},\ a \in Q \text{ and } l^k_{u_1 v_1} = a = l^k_{u_2 v_2}\big\}$$

$$\mathscr{B} = \big\{\{(u_1,v_1),(u_2,v_2)\} : u_1, v_1, u_2, v_2 \in \{1,\ldots,q\} \text{ and } u_1 \neq u_2, v_1 \neq v_2\big\}$$

and let

$$\varphi\big(k, a, \{(u_1,v_1),(u_2,v_2)\}\big) = \{(u_1,v_1),(u_2,v_2)\}.$$

(a) Show that φ is well defined, i.e. $\varphi\big(k, a, \{(u_1,v_1),(u_2,v_2)\}\big)$ belongs to $\mathscr{B}$ for all $\big(k, a, \{(u_1,v_1),(u_2,v_2)\}\big) \in \mathscr{A}$.

(b) Show that φ is injective.

(c) Show that $|\mathscr{A}| = |\mathscr{B}|$ and conclude that φ is bijective.)

4.5. Show that the design in Example 4.13 (a) is the Fano plane.

4.6. Let X be a finite set with v elements and $\mathscr{B}$ is collection of b distinct subsets of X such that

- each set in $\mathscr{B}$ contains at most k elements,
- every element from X belongs to exactly r sets in $\mathscr{B}$, and
- each 2-element subset of X is contained in at least λ sets in $\mathscr{B}$.

Show that $bk \geqslant rv$ and $r(k-1) \geqslant \lambda(v-1)$.

4.7. Prove Lemma 4.16

4.8. Prove Lemma 4.18. (Hint: Use Theorem 4.14 and the assumption $k < v$.)

4.9. Prove Lemma 4.19. (Hint: Show that in a symmetric design we have $k - \lambda = 1$ if and only if $k = v - 1$.)

4.10. Prove Lemma 4.25.

4.11. Show that if H_n is an Hadamard matrix of order n it is easy to show that $H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$ is an Hadamard matrix of order $2n$.

Exercises

=

4.12. Is there a finite projective plane with 9 points? And with 12 points?

4.13. Show that in a projective plane of order q every point belongs to precisely $q+1$ lines.

4.14. How many noncollinear triples of points $\{A,B,C\}$ are there in a finite projective plane of order q?

4.15. (a) Show that in every finite projective plane there exist four distinct lines such that no three of them are concurrent.

(b) The *Duality Principle* for projective planes is a metatheorem of projective geometry. Take any statement about a projective plane and interchange words and phrases as follows: "point" $\leftrightarrow$ "line", "belongs to ($\in$)" $\leftrightarrow$ "contains ($\ni$)", "the point of intersection of the two lines" $\leftrightarrow$ "the line that passes through the two points" etc. The statement you obtain is said

to be *dual* to the original statement. The Duality Principle states that a statement is true in a projective plane if and only if its dual is true.

Show the Duality Principle for projective planes. (Hint: Recall that a proof of a statement S is a sequence of statements $S_1, S_2, \ldots, S_n \equiv S$ such that every S_i is either an axiom or there exist $k, j < i$ such that $S_k \equiv S_j \Rightarrow S_i$. Let S^∂ denote the dual statement of S. Show that the dual od all the axioms are true and show that if $S_1, S_2, \ldots, S_n$ is a proof of S then $S_1^\partial, S_2^\partial, \ldots, S_n^\partial$ is a proof of S^∂.)

4.16. Let V be a 3-dimensional vector space over a finite field $\mathbb{F}_q$ with $q = p^k$ elements. Let π be the set of all 1-dimesional subspaces of V and $\mathscr{L}$ the set of all 2-dimensional subspaces of V.

(a) Show that $(\pi, \mathscr{L})$ is a projective plane.

(b) Find $|\pi|$. What is the order of $(\pi, \mathscr{L})$?

4.17. Is there an $(11, 6, 2)$-design?

4.18. Show that for a (v, k, λ)-design, $b/\lambda = \binom{v}{2} \Big/ \binom{k}{2}$.

4.19. Let $(X, \mathscr{B})$ be a (v, k, λ)-design such that $b - 2r + \lambda > 0$. Show that $(X, \overline{\mathscr{B}})$ is a $(v, v-k, b-2r+\lambda)$-design.

4.20. Let $(X, \mathscr{B})$ be a (v, k, λ)-design and let $\mathscr{B}^* = \{D \subseteq X : |D| = k \text{ and } D \notin \mathscr{B}\}$. Show that $(X, \mathscr{B}^*)$ is a (v, k, λ^*)-design for some λ^*. What is the value of λ^*?

4.21. Show that every nontrivial symmetric (v, k, λ)-design has a complement (i.e. $b - 2r + \lambda > 0$) and its complement is a symmetric $(\overline{v}, \overline{k}, \overline{\lambda})$-design where $\overline{v} = v$, $\overline{k} = v - k$ and $\overline{\lambda} = v - 2k + \lambda$. Moreover, $\overline{n} = n$, where $\overline{n} = \overline{k} - \overline{\lambda}$.

4.22. Show that if $(X, \mathscr{B})$ is a maximal nontrivial symmetric design, then one of the designs $(X, \mathscr{B})$, $(X, \overline{\mathscr{B}})$ is a finite projective plane. (Hint: show that $v = n^2 + n + 1$ implies $\lambda = 1$ or $\overline{\lambda} = 1$; see proof of Theorem 4.21.)

4.23. Show that if $(X, \mathscr{B})$ is a nontrivial symmetric (v, k, λ)-design with $v = 4n - 1$, then one of the designs $(X, \mathscr{B})$, $(X, \overline{\mathscr{B}})$ is a $(4n-1, 2n-1, n-1)$-design. (Hint: Recall that $\lambda\overline{\lambda} = n(n-1)$ and $\lambda + \overline{\lambda} = v - 2n = 2n - 1$. Conclude that $\{\lambda, \overline{\lambda}\} = \{n, n-1\}$ and discuss the two possibilities.)

4.24. Show that for every nontrivial symmetric (v, k, λ)-design there exists a Latin $k \times v$ rectangle whose columns are blocks of the design. (Hint: Use Theorem 3.8.)

4.25. A (v,k,λ)-design $(X,\mathscr{B})$ is *resolvable* if there is a partition $\{\mathscr{B}_1,\ldots,\mathscr{B}_r\}$ of $\mathscr{B}$ in r classes such that each $\mathscr{B}_i$ consists of b/r blocks, and for each $i\in\{1,\ldots,r\}$ and each $x\in X$ there is exactly one $B\in\mathscr{B}_i$ that contains x. The classes $\mathscr{B}_i$ are called the *parallel classes* of the resolvable design.

Show that for every $n>0$ there exists a resolvable $(2n,2,1)$-design. (Hint: Take $X=\{0,1,\ldots,2n-1\}$ and define the parallel classes $\mathscr{B}_1,\ldots,\mathscr{B}_{2n-1}$ as follows: $\{0,i\}\in\mathscr{B}_i$, and if $x,y>0$ and $x+y\equiv 2i \pmod{2n-1}$ then $\{x,y\}\in\mathscr{B}_i$.)

4.26. (a) Show that the definition of an Hadamard design is correct, i.e. that the configuration obtained by the construction is indeed a design.

(b) Show that every Hadamard design indeed gives rise to an Hadamard matrix.

Chapter 5

Graphs and Digraphs

Graphs represent one of the most popular tools for modeling discrete phenomena where the abstraction of the problem involves information about certain objects being connected or not. For example, crossings in a city transportation model are joined by streets, or cities in a country are joined by roads. We will examine two types of such models: graphs which correspond to situations where all the "roads" are bidirectional, and digraphs (*di*rected *graphs*) where one-way "roads" are allowed.

5.1 Graphs

A *graph* is an ordered pair $G = (V, E)$ where V is a nonempty finite set and E is an arbitrary subset of $V^{(2)} = \big\{\{u, v\} \subseteq V : u \neq v\big\}$. Elements of V are called *vertices* of G, while elements of E are called *edges* of G. We shall often write $V(G)$ and $E(G)$ to denote the set of vertices and the set of edges of G, and $n(G)$ and $m(G)$ to denote the number of vertices and the number of edges of G. If $e = \{u, v\}$ is an edge of a graph, we say that u and v are *adjacent*, and that e is *incident* with u and v. We also say that u is a *neighbour* of v. The *neighbour-set of v* is the set $N_G(v) = \{x \in V(G) : x \text{ is a neighbour of } v\}$. The *degree of a vertex v*, denoted by $\delta_G(v)$, is the number of edges incident to v: $\delta_G(v) = |N_G(v)|$. If G is clear from the context, we simply write $N(v)$ and $\delta(v)$. By $\delta(G)$ we denote the least, and by $\Delta(G)$ the greatest degree of a vertex in G. A vertex with degree 0 is said to be an *isolated vertex*. A vertex of degree 1 is called a *leaf of G*. A vertex is said to be *even*, resp. *odd* according as $\delta(v)$ is an even or an odd integer. A graph is *regular* if $\delta(G) = \Delta(G)$. In other words, in a regular graph all vertices have the same degree.

The graphs are called graphs because of a very natural graphical representation they have. Vertices are usually represented as (somewhat larger) points in a plane,

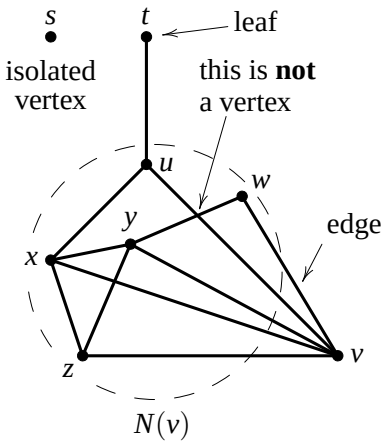


Figure 5.1: An example of a graph

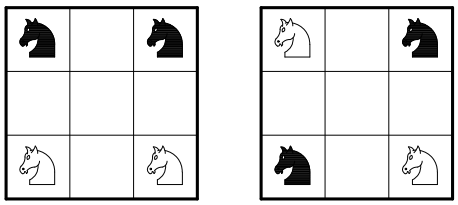
while edges are represented as (smooth non-selfintersecting) curves joining the respective vertices, so that adjacent vertices are joined by a curve.

Example 5.1 Fig. 5.1 depicts a graf G with $V = \{s,t,u,v,w,x,y,z\}$ and $E = \big\{\{t,u\}, \{u,x\}, \{u,v\}, \{w,y\}, \{w,v\}, \{v,x\}, \{v,y\}, \{v,z\}, \{x,y\}, \{x,z\}, \{y,z\}\big\}$. We see that

vertex	s	t	u	v	w	x	y	z
δ	0	1	3	5	2	4	4	3

so $\delta(G) = 0$ and $\Delta(G) = 5$. Also, $N(v) = \{u,w,x,y,z\}$.

Example 5.2 Two black and two white knights are placed on a 3×3 chessboard as in Fig. 5.2 (a). Is it possible to reach the configuration in Fig. 5.2 (b) following the rules of chess?



(a) (b)

Figure 5.2: Example 5.2

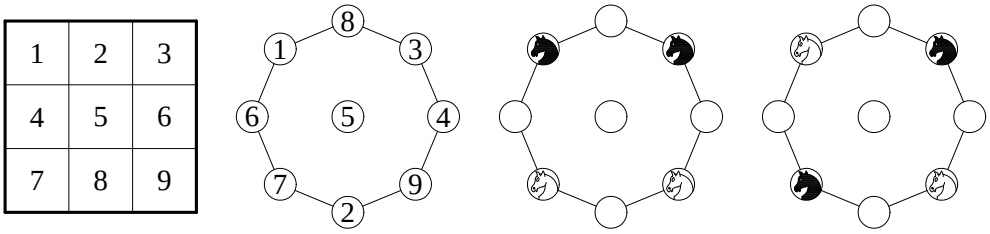


Figure 5.3: Solution to the problem in Example 5.2

Answer: No. Let us enumerate the fields of the chess board by 1, ..., 9 as in Fig. 5.3 (a). To this chess board we can now assign a graph with $\{1,\ldots,9\}$ as the set of vertices by joining i and j if an only if it is possible for a knight to jump from i to j following the general rules of chess. The graph is given in Fig. 5.3 (b). Clearly, regular movements of a knight on the 3×3 chess board correspond to movements of the knight along the edges of the graph in Fig. 5.3 (b). We see now that it is not possible to start from the initial position of the knights given in Fig. 5.3 (c) and reach the final position in Fig. 5.3 (d) by moving one knight at a time along the edges of the graph simply because the white knights separate the black knights in Fig. 5.3 (d), which is not the case in the initial position.

Theorem 5.3 (The First Theorem of Graph Theory) *If $G=(V,E)$ is a graph with m edges, then $\sum_{v\in V}\delta(v)=2m$.*

Proof. Since every edge is incident to two vertices, every edge is counted twice in the sum on the left. □

Corollary 5.4 *In any graph the number of odd vertices is even.*

Theorem 5.5 *If $n(G)\geqslant 2$, there exist vertices $v,w\in V(G)$ such that $v\neq w$ and $\delta(v)=\delta(w)$.*

Proof. Let $V(G)=\{v_1,\ldots,v_n\}$ and suppose that $v_i\neq v_j$ whenever $i\neq j$. Without loss of generality we may assume that $\delta(v_1)<\delta(v_2)<\ldots<\delta(v_n)$. Since there are only n possibilities for the degree of a vertex (0, 1, ..., $n-1$) it follows that $\delta(v_1)=0$, $\delta(v_2)=1$, ..., $\delta(v_n)=n-1$. But then v_n is adjacent to every other vertex of a graph, including the isolated vertex v_1. Contradiction. □

A graph $H=(W,E')$ is a *subgraph* of a graph $G=(V,E)$, in symbols $H\leqslant G$, if $W\subseteq V$ and $E'\subseteq E$. A subgraph H of G is a *spanning subgraph* if $W=V(G)$.

A subgraph H is an *induced subgraph* of G if $E' = E \cap W^{(2)}$. Induced subgraphs are usually denoted by $G[W]$. The edges of an induced subgraph of G are all the edges of G whose both ends are in W. A set of vertices $W \subseteq V(G)$ is *independent* if $E(G[W]) = \varnothing$, i.e. no two vertices in W are adjacent in G. By $\alpha(G)$ we denote the maximum carinality of an independent set of vertices in G. If $A, B \subseteq E(G)$ are disjoint, by $E(A,B)$ we denote the set of all edges in G whose one end is in A and the other in B.

Theorem 5.6 $\alpha(G) \leqslant n(G) - \delta(G)$.

Proof. Let $A \subseteq V(G)$ be an independent set of vertices of G such that $\alpha(G) = |A|$. Take any $v \in A$. Since A is independent all vertices adjacent to v are in $V(G) \setminus A$, so $\delta(v) \leqslant |V(G) \setminus A| = n(G) - |A|$. Now $\delta(G) \leqslant \delta(v) \leqslant n(G) - |A|$ and the statement follows. □

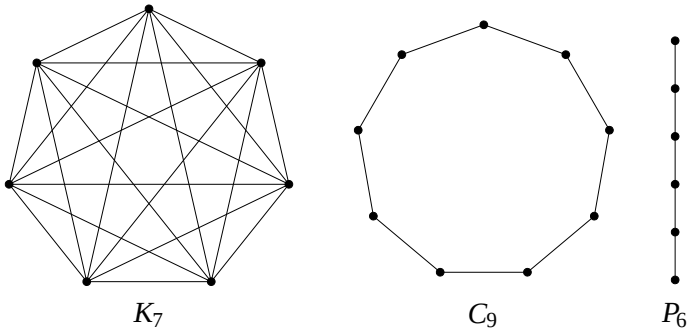


Figure 5.4: K_7, C_9 and P_6

A *complete graph on n vertices* (or an *n-clique*) is a graph with n vertices where each two distinct vertices are adjacent. A complete graph on n vertices is denoted by K_n. A *cycle* of length n, denoted by C_n, is the graph with n vertices where the first vertex is adjacent to the second one, and the second vertex to the third one, and so on, the last vertex is adjacent to the first. A *path* with n vertices is a graph where the first vertex is adjacent to the second one, and the second vertex to the third one, and so on, and the penultimate vertex is adjacent to the last one, but the last vertex is *not* adjacent to the first. We say that the path with n vertices has length $n-1$. Fig. 5.4 depicts K_7, C_9 and P_6.

Theorem 5.7 *If $\delta(G) \geqslant 2$ then G contains a cycle.*

Proof. Let $x_1 \ldots x_{k-1}\, x_k$ be the longest path in G. Since $\delta(x_k) \geqslant \delta(G) \geqslant 2$, x_k has a neighbour v distinct from x_{k-1}. If $v \notin \{x_1, \ldots, x_{k-2}\}$ then $x_1 \ldots x_{k-1}\, x_k\, v$ is a path

with more vertices than the longest path, which is impossible. Therefore, $v = x_j$ for some $j \in \{1,\ldots,k-2\}$ so $x_j \ldots x_k$ are vertices of a cycle in G. □

Theorem 5.8 *If $C_3 \not\leqslant G$ then $n(G) \leqslant \alpha(G)(\alpha(G)+1)$.*

Proof. Let S be an independent set of vertices of G such that $\alpha(G) = |S|$ and let $T = V(G) \setminus S$. Since S is maximal, every vertex from T has a neighbour in S, so $|E(S,T)| \geqslant n - \alpha(G)$. This implies $\sum_{v\in S} \delta(v) \geqslant n - \alpha(G)$. Let $v \in S$ be the vertex of the greatest degree in S. Then $\alpha(G) \cdot \delta(v) \geqslant \sum_{v\in S} \delta(v) \geqslant n - \alpha(G)$ whence $\delta(v) \geqslant \frac{n-\alpha(G)}{\alpha(G)}$. Vertices in T adjacent to v form an independent set since G does not have a C_3 as its subgraph. Now, $\alpha(G)$ is the maximum cardinality of an independent set whence $\delta(v) \leqslant \alpha(G)$. Therefore, $\frac{n-\alpha(G)}{\alpha(G)} \leqslant \delta(v) \leqslant \alpha(G)$ and thus $n \leqslant \alpha(G)(\alpha(G)+1)$. □

Graphs G_1 and G_2 are *isomorhic*, and we write $G_1 \cong G_2$, if there is a bijection $\varphi : V(G_1) \to V(G_2)$ such that $\{x,y\} \in E(G_1) \Leftrightarrow \{\varphi(x),\varphi(y)\} \in E(G_2)$. For example graphs G and G_2 in Fig. 5.5 are isomorphic, while G and G_1 are not.

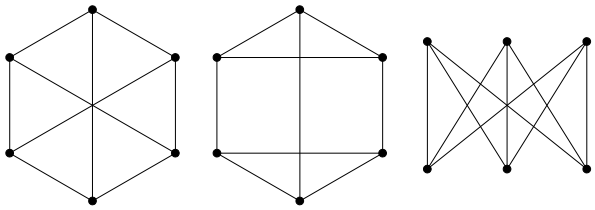


Figure 5.5: $G \cong G_2$, but $G \ncong G_1$

Theorem 5.9 *Let $G_1 \cong G_2$ and let φ be an isomorphism between G_1 and G_2. Then $n(G_1) = n(G_2)$, $m(G_1) = m(G_2)$ and $\delta_{G_1}(x) = \delta_{G_2}(\varphi(x))$ for every $x \in V(G_1)$.*

The *complement* of a graph $G = (V,E)$ is the graph $\overline{G} = (V,\overline{E})$ where $\overline{E} = V^{(2)} \setminus E$. A graph G is *selfcomplementary* if $G \cong \overline{G}$. Clearly, $m(G) + m(\overline{G}) = \binom{n}{2}$.

Lemma 5.10 *Let G and H be graphs.*
(a) $G \cong H$ if and only if $\overline{G} \cong \overline{H}$.
(b) $\delta_{\overline{G}}(x) = (n(G)-1) - \delta_G(x)$ for all $x \in V(G)$.

Theorem 5.11 *If G is a selfcomplementary graph with n vertices then $n \geqslant 4$ and $n \equiv 0,1 \pmod 4$. Conversely, for every integer $n \geqslant 4$ such that $n \equiv 0,1 \pmod 4$ there exists a selfcomplementary graph with n vertices.*

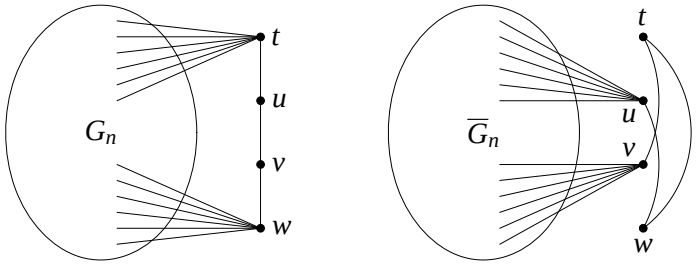


Figure 5.6: The proof of Theorem 5.11

Proof. Let G be a selfcomplementary graph with $n \geqslant 4$ vertices and m edges and let $\overline{m} = m(\overline{G})$. Then $m + \overline{m} = \binom{n}{2}$ and $m = \overline{m}$ since $G \cong \overline{G}$. Therefore $2m = \dfrac{n(n-1)}{2}$ i.e. $m = \dfrac{n(n-1)}{4}$. But m is an integer and n and $n-1$ are not of the same parity, so $4 \mid n$ or $4 \mid n-1$.

For the other part of the statement, for every integer $n \geqslant 4$ such that $n \equiv 0,1 \pmod 4$ we shall construct a selfcomplementary graph $G_n = (V_n, E_n)$ with n vertices. It is obvious that we can take $G_4 = P_4$ and $G_5 = C_5$. Now let G_n be a selfcomplementary graph with n vertices and construct G_{n+4} as follows. Take four new vertices t, u, v, w and put

$$V_{n+4} = V_n \cup \{t,u,v,w\}$$
$$E_{n+4} = E_n \cup \{\{t,u\},\{u,v\},\{v,w\}\} \cup \{\{t,x\} : x \in V_n\} \cup \{\{w,x\} : x \in V_n\},$$

see Fig. 5.6 (a). Then $\overline{G}_{n+4}$ is given in Fig. 5.6 (b) and it is easy to establish that $G_{n+4} \cong \overline{G}_{n+4}$. □

5.2 Connectedness and distance

A *walk* in a graph G is any sequence of vertices and edges $v_0\, e_1\, v_1\, e_2\, v_2 \ldots v_{k-1}\, e_k\, v_k$ such that $e_i = \{v_{i-1}, v_i\}$ for all $i \in \{1,\ldots,k\}$. Note that an edge or a vertex may appear more than once in a walk. We say that k is the *length* of the walk. If $v_0 \neq v_k$ we say that the *walk connects v_0 and v_k*. A *closed walk* is a walk $v_0\, e_1\, v_1\, \ldots\, v_{k-1}\, e_k\, v_k$ where $v_0 = v_k$. Clearly, a path is a walk where neither vertices nor edges are allowed to repeat, and a cycle is a closed walk where neither edges nor vertices are allowed to repeat, except for the first and the last vertex.

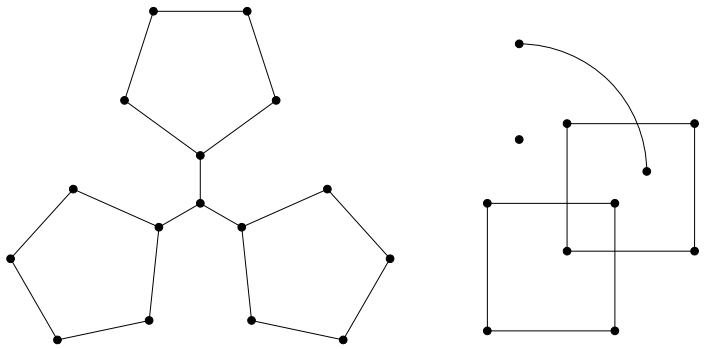


Figure 5.7: (a) A connected graph; (b) A graph with $\omega = 4$

Lemma 5.12 *If there is a walk in G that connects two vertices then there is a path that connects them. Every closed walk of odd length contains an odd cycle.*

We define a binary relation θ on $V(G)$ by $x\theta y$ if $x = y$ or there is a walk that connects x and y. Clearly, θ is an equivalence relation on $V(G)$ and hence partitions $V(K)$ into blocks $S_1, \ldots, S_t$. These blocks or the corresponding induced subgraphs (depending on the context) are called *connected components* of G. The number of connected components of G is denoted by $\omega(G)$. A graph G is *connected* if $\omega(G) = 1$. An example of a connected graph and of a graph with four connected components are given in Fig. 5.7.

Lemma 5.13 *$S \subseteq V(G)$ is a connected component of G if and only if no proper superset $S' \supset S$ induces a connected subgraph of G.*

Theorem 5.14 *A graph G is connected if and only if $E(A,B) \neq \varnothing$ for every partition $\{A,B\}$ of $V(G)$.*

Proof. $(\Rightarrow)$ Let G be a connected graph and $\{A,B\}$ a partition of $V(G)$. Take any $a \in A$ and $b \in B$. Now G is connected, so there is a path $x_1 \ldots x_k$ that connects a and b. Since $x_1 = a$ and $x_k = b$, there is a j such that $x_j \in A$ and $x_{j+1} \in B$ whence $E(A,B) \neq \varnothing$.

$(\Leftarrow)$ Suppose G is not connected and let $S_1, \ldots, S_\omega$ be the connected components. Then Lemma 5.13 yields $E(S_1, \bigcup_{j=2}^{\omega} S_j) = \varnothing$. □

Theorem 5.15 *At least one of the graphs G, $\overline{G}$ is connected.*

Proof. Suppose that G is not connected and let $S_1, \ldots, S_\omega$, $\omega \geqslant 2$, be the connected components of G. Let us show that there is a path that connects any two vertices in G. Take any $x, y \in V(G)$, $x \neq y$. If x and y belong to distinct connected components of G then $\{x,y\} \notin E(G)$ and hence $\{x,y\} \in E(\overline{G})$, so they are connected by an edge. If, however, x and y belong to the same connected component of G, say S_i, take any $j \neq i$ and any $z \in S_j$. Then x and z are connected by an edge in $\overline{G}$ and so are y and z. Therefore, $x\,z\,y$ is a path in $\overline{G}$ that connects x and y. □

We see from the proof of previous theorem that if G is not connected, then $\overline{G}$ is "very connected". We shall now introduce a numerical measure that enables us to express such statements formally.

The *distance* d_G between vertices x and y of a connected graph G is defined by $d_G(x,x) = 0$, and in case $x \neq y$,

$$d_G(x,y) = \min\{k : \text{ there is a path of length } k \text{ that connects } x \text{ and } y\}.$$

Theorem 5.16 *Let $G = (V,E)$ be a connected graph. Then (V, d_G) is a metric space, i.e. for all $x,y,z \in V$ the following holds:*

(D1) $d_G(x,y) \geqslant 0$*;*

(D2) $d_G(x,y) = 0$ *if and only if* $x = y$*;*

(D3) $d_G(x,y) = d_G(y,x)$*; and*

(D4) $d_G(x,z) \leqslant d_G(x,y) + d_G(y,z)$.

If G is obvious, instead of d_G se simply write d. The *diameter* $d(G)$ of a connected graph G is the maximum distance between two of its vertices:

$$d(G) = \max\{d(x,y) : x,y \in V(G)\}.$$

Example 5.17 (a) $d(G) = 1$ if and only if G is a complete graph.

(b) $d(P_n) = n-1$ and $d(C_n) = \lfloor \dfrac{n-1}{2} \rfloor$.

A graph G is *bipartite* if there is a partition $\{X,Y\}$ of $V(G)$ such that every edge in G has one end in X and the other in Y, i.e. $E(G) = E(X,Y)$. Therefore, X and Y are independent sets. A *complete bipartite graph* is a bipartite graph with partition $\{X,Y\}$ of vertices such that its edges are *all* pairs $\{x,y\}$ with $x \in X$ and $y \in Y$. If $|X| = p$ and $|Y| = q$, the complete bipartite graph with the partition $\{X,Y\}$ is denoted by $K_{p,q}$. A *star* with n vertices, denoted by S_n, is a complete bipartite graph $K_{1,n-1}$. A bipartite graph, a $K_{3,4}$ and a star S_{10} are depicted in Fig. 5.8.

Lemma 5.18 *A graph G with at least two vertices is a bipartite graph if and only if every connected component of G is either an isolated vertex or a bipartite graph.*

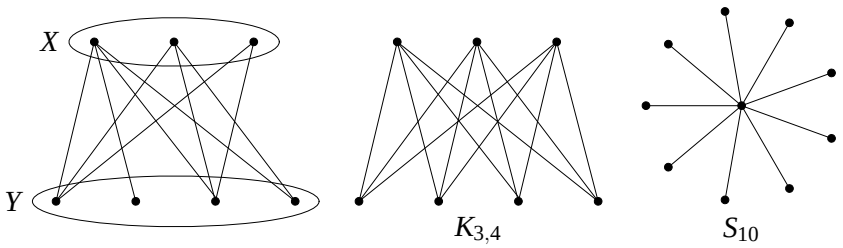


Figure 5.8: A bipartite graph, a $K_{3,4}$ and a star S_{10}

Theorem 5.19 *A graph G with at least two vertices is bipartite if and only if G does not contain an odd cycle.*

Proof. According to Lemma 5.18 it suffices to give the proof for connected graphs. So, let G be a connected graph and $n(G) \geqslant 2$.

($\Rightarrow$) Let G be a bipartite graph and suppose G contains an odd cycle whose vertices are v_1, v_2, ..., v_{2k+1}. So v_i is adjacent to v_{i+1} for all $i \in \{1,\ldots,2k\}$ and v_{2k+1} is adjacent to v_1. Let $\{X,Y\}$ be a partition of $V(G)$ showing that G is bipartite, i.e. such that $E(G[X]) = E(G[Y]) = \varnothing$. Now v_1 belongs to X or Y, so assume that $v_1 \in X$. Then $v_2 \in Y$ since v_2 is adjacent to v_1 and G is bipartite, and this forces $v_3 \in X$, $v_4 \in Y$ and so on. We see that vertices with odd indices belong to X, so $v_{2k+1} \in X$. But we have $x_1 \in X$ too, so $E(G[X])$ contains $\{x_1, x_{2k+1}\}$ which contradicts the assumption $E(G[X]) = \varnothing$.

($\Leftarrow$) Suppose G does not contain an odd cycle. Take any $v \in V(G)$ and define $A_0, A_1, \ldots \subseteq V(G)$ as follows:

$$A_n = \{x \in V(G) : d(v,x) = n\},$$

for $n \geqslant 0$. Since G is connected, there is a path connecting v to any other vertex of G, so each vertex of G appears in at least one of the A_i's. The A_i's are disjoint by the construction and the fact that $V(G)$ is finite now yields that there is an s such that $\{A_0, A_1, \ldots, A_s\}$ is a partition of $V(G)$ and $A_t = \varnothing$ for all $t > s$. Let

$$X = \bigcup_{j \text{ even}} A_j, \quad \text{and} \quad Y = \bigcup_{j \text{ odd}} A_j$$

and let us show that both X and Y are independent sets in G. First, let us note that $E(A_j, A_{j+2}) = \varnothing$, for if $x \in A_j$ and $y \in A_{j+2}$ were adjacent then $d(v,y) \leqslant d(v,x) + d(x,y) = j+1$ which contradicts $y \in A_{j+2}$. Next, let us show that each A_j is an independent set. Assume that there is a j and vertices $x, y \in A_j$ such that x and y are adjacent, Fig. 5.9. By the construction of A_i's there is a path $v \ldots x$ with j edges, and there is a path $y \ldots v$ with j edges. By chaining these two paths

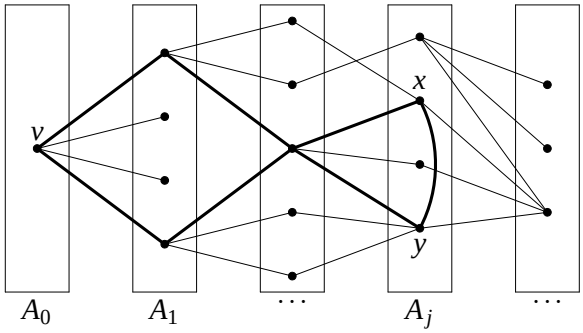
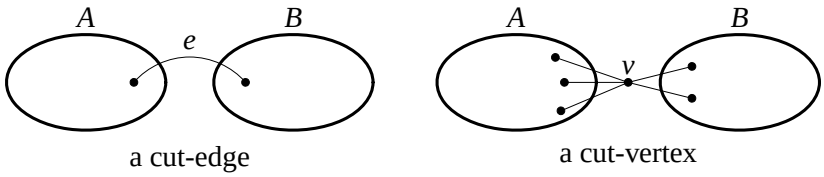


Figure 5.9: The proof of Theorem 5.19

together with the edge $e = \{x,y\}$ we obtain a walk $v \dots x\, e\, y \dots v$ of length $2j+1$, so by Lemma 5.12 G contains an odd cycle, which is impossible. Therefore, all A_j's are independent and consequently, both X and Y are sets of independent vertices. This shows that G is a bipartite graph and one possible partition of its vertices is $\{X,Y\}$. □

Note that this theorem does not imply that bipartite graphs have to have cycles. A graph with no cycles is a bipartite graph, and this follows from the theorem since it has *no odd cycles*.

Let e be an edge and v a vertex of a graph G. By $G-e$ we denote the graph obtained from G by removing the edge e, while $G-v$ denotes the graph obtained from G by removing v and all the edges of G incident to v. A *cut-vertex* of a graph G is a vertex $v \in V(G)$ such that $\omega(G-v) > \omega(G)$. A *cut-edge* of a graph G is an edge $e \in E(G)$ such that $\omega(G-e) > \omega(G)$. Cut-vertices and cut-edges are weak points in the graph since removing one of these makes the graph split. Intuitively, they look like this:



Theorem 5.20 *Let e be an edge of a graph G. The following are equivalent:*

(1) e is a cut-edge of G;

(2) there is a partition $\{A,B\}$ of $V(G)$ such that $E(A,B) = \{e\}$;

(3) e belongs to no cycle of G.

Proof. We give the proof in case G is connected. If G is not connected it suffices to consider the connected component of G that contains e.

(2) $\Rightarrow$ (1): If $E(A,B) = \{e\}$ in G then $E(A,B) = \varnothing$ in $G-e$, so $G-e$ is not connected by Theorem 5.14. Therefore, $2 = \omega(G-e) > \omega(G) = 1$.

(1) $\Rightarrow$ (3): Suppose that e appears in a cycle $C = v_0\, e\, v_1\, e_2\, v_2\, \ldots\, v_{k-1}\, e_k\, v_0$ of G. To show that $G-e$ is connected take any $x \neq y$. Since G is connected, there is a path P that connects x to y. If P does not contain e, it is also a path in $G-e$ that connects x to y. If, however, P contains e, say $P = x \ldots v_0\, e\, v_1 \ldots y$, then remove e from P and replace it with $C-e$ to obtain the following walk:

$$W = x \ldots \underbrace{v_0\, e_k\, v_{k-1} \ldots v_2\, e_2\, v_1}_{C-e} \ldots y,$$

Fig. 5.10. Since e appears once in P and once in C it follows that e does not appear in W, so W is a walk from x to y in $G-e$.

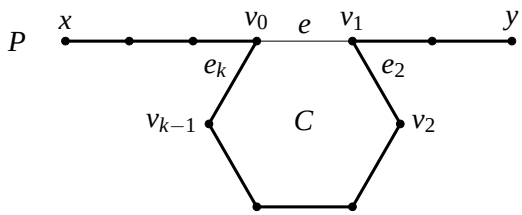


Figure 5.10: The walk W

(3) $\Rightarrow$ (2): Suppose that $e = \{a,b\}$ belongs to no cycle of G and define A and B as follows: $A = \{a\} \cup \{x \in V(G)$: there is a path from a to x that does not pass through $e\}$ and $B = V(G) \setminus A$. If $b \notin B$ then $b \in A$ and there is a path from a to b that does not pass through e. This path together with e forms a cycle that contains e. Since there are no such cycles we have $b \in B$. So, $\{A,B\}$ is a partition of $V(G)$ and $e \in E(A,B)$. Suppose now that there is an $e' \in E(A,B)$, $e' \neq e$, and let $e' = \{a',b'\}$, $a' \in A$, $b' \in B$, Fig. 5.11. We will assume further that $a \neq a'$ and $b \neq b'$ since these two cases follow by similar arguments. There is a path $P_A = a \ldots a'$ that does not pass through e and there is a path $P_B = b' \ldots b$ that does not pass through e. Now these two paths together with e and e' form a cycle $\underbrace{a \ldots a'}_{P_A}\, e'\, \underbrace{b' \ldots b}_{P_B}\, e\, a$ which contains e. This contradiction shows that $E(A,B) = \{e\}$. □

Theorem 5.21 *Let v be a vertex of G. Then v is a cut-vertex of G if and only if there is a partition $\{A,B\}$ of $V(G) \setminus \{v\}$ such that $E(A,B) = \varnothing$, $E(A,\{v\}) \neq \varnothing$ and $E(B,\{v\}) \neq \varnothing$.*

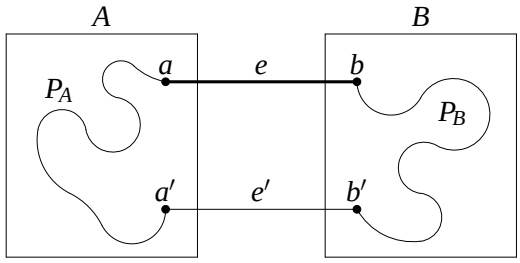


Figure 5.11: A cycle that contains e

Theorem 5.22 *If e is a cut-edge of G then $\omega(G-e) = \omega(G)+1$. If v is a cut-vertex of G then $\omega(G-v) < \omega(G)+\delta(v)$.*

Theorem 5.23 *If G is a connected graph with at least three vertices and if G has a cut-edge, then G has a cut-vertex.*

Proof. Let e be a cut-edge of G. Then there is a partition $\{A,B\}$ of $V(G)$ such that $E(A,B) = \{e\}$ (Theorem 5.20). Let $e = \{a,b\}$ and let $a \in A$ and $b \in B$. From $n(G) \geqslant 3$ it follows that $|A| \geqslant 2$ or $|B| \geqslant 2$, say $|A| \geqslant 2$. Since the graph is connected, a has a neighbour c in A, Fig. 5.12. Now let $A' = A \setminus \{a\}$ and note that

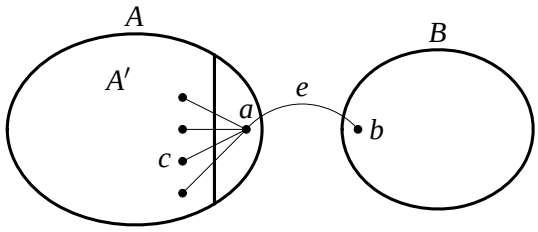


Figure 5.12: The proof of Theorem 5.23

$E(A',B) = \varnothing$, $E(A',\{a\}) \neq \varnothing$ and $E(B,\{a\}) \neq \varnothing$. Therefore, a is a cut-vertex according to Theorem 5.21. □

We have seen in Theorem 5.20 that a graph has no cut-edges if and only if every edge belongs to a cycle. The analogous statement for cut-vertices is the famous Whitney Theorem.

Theorem 5.24 (Whitney 1932) *Let G be a connected graph with at least three vertices. G has no cut-vertices if and only if any two vertices lie on a common cycle.*

Proof. ($\Leftarrow$) Since any two vertices u and v lie on a common cycle, removing one vertex from the graph cannot separate u from v, and hence $G-x$ is connected for all x.

($\Rightarrow$) For the converse, suppose that G has no cutvertices. We say that two paths $ux_1 \ldots x_k v$ and $uy_1 \ldots y_l v$ connecting u to v are internally disjoint if $\{x_1,\ldots,x_k\} \cap \{y_1,\ldots,y_l\} = \varnothing$. Now take any u and v in G, $u \neq v$, and let us show by induction on $d(u,v)$ that G has two intenally disjoint paths connecting u and v. Clearly, the two paths will then form a cycle containing both u and v.

Let $d(u,v) = 1$ and let $e = \{u,v\}$. The graph $G-e$ is connected by Theorem 5.23 so there is a path in $G-e$ from u to v. This is also a path in G and it is internally disjoint from the trivial path $u\,v$ consisting of the edge e itself.

For the induction step, let $d(u,v) = k > 1$ and assume that G has internally disjoint paths connecting every pair of vertices x, y such that $1 \leqslant d(x,y) < k$. Let $u\,x_1\,\ldots\,x_{k-1}\,v$ the a path of length k (i.e. one of the shortest paths that connect u to v). We have $d(u,x_{k-1}) = k-1$, and hence by the induction hypothesis G has internally disjoint paths P and Q joining u to x_{k-1}, Fig. 5.13. Since $G-x_{k-1}$ is

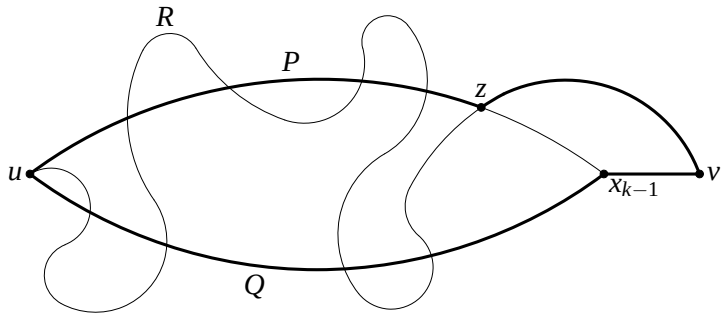


Figure 5.13: The proof of Whitney's theorem

connected, $G-x_{k-1}$ contains a path R that joins u and v. If this path is internally disjoint from P or Q we are done, so assume that R shares internal vertices with both P and Q. Let z be the last vertex of R belonging to $P \cup Q$. Without loss of generality we may assume that $z \in P$. We now combine the subpath of P joining u to z with the subpath of R joining z to v to obtain a path from u to v internally disjoint from the path $Q' = Q\,e'\,v$ where $e' = \{x_{k-1},v\}$. □

5.3 Trees

A *tree* is a connected graph with no cycles. By Theorem 5.20 we see that every edge of a tree is a cut-edge. Therefore, a tree is a minimal connected graph with the given set of vertices. The following theorem shows that in a way trees capture the essence of the property of being connected.

Recall that a spanning subgraph of a graph $G = (V,E)$ is a graph $H = (W,E')$ such that $W = V$ and $E' \subseteq E$. If H is a tree, we say that H is a *spanning tree of* G.

Theorem 5.25 *A graph with at least two vertices is connected if and only if it has a spanning tree.*

Proof. Clearly, if a graph G contains a connected subraph H then G is also connected. Therefore if a graph has a spanning tree, it is connected. For the converse, take any connected graph G and construct a sequence of graphs G_0, G_1, G_2, ... as follows: $G_0 = G$; if G_i has a cycle, take any edge e_i that lies on a cycle and let $G_{i+1} = G_i - e_i$, otherwise put $G_{i+1} = G_i$. Each G_i is a spanning subgraph of G and each G_i is connected since an edge that lies on a cycle cannot be a cut-edge (Theorem 5.20). Moreover, if $G_i = G_{i+1}$ then $G_i = G_j$ for all $j > i$. Let m be the number of edges of G. Since we cannot remove more than m edges from G, we conclude that $G_{m+1} = G_{m+2}$. By construction of the sequence this means that G_{m+1} has no cycles. Therefore, G_{m+1} is a spanning tree of G. □

We will now show that each tree with n vertices has $n-1$ edges and that each two of the three properties listed below implies the remaining one:

- being connected,
- having no cycles, and
- $m = n-1$.

Lemma 5.26 *Each tree with at least two vertices has at least two leaves.*

Proof. Let G be a tree with $n \geqslant 2$ vertices and let v_1, v_2, ..., v_k be the longest path in the tree. Then $k \geqslant 2$ since G is a connected graph with at least two vertices. If $\delta(v_1) > 1$ then v_1 has a neighbour x distinct from v_2. If x is a new vertex, i.e. $x \notin \{v_3, \ldots, v_k\}$, then the path x, v_1, v_2, ..., v_k is longer than the longest path in G, which is impossible. If, however, $x \in \{v_3, \ldots, v_k\}$ then G has a cycle, which contradicts the assumption that G is a tree. Therefore, v_1 is a leaf. The same argument shows that v_k is another leaf. □

Theorem 5.27 *Let $G = (V,E)$ be a tree with n vertices and m edges. Then $m = n-1$, and consequently $\sum_{v \in V} \delta(v) = 2(n-1)$.*

Proof. The second part of the theorem follows from the First Theorem of Graph Theory, so let us show that $m = n-1$. The proof is by induction on n. The cases $n = 1$ and $n = 2$ are trivial. Assume that the statement is true for all trees with less then n vertices and consider a tree G with n vertices. By Lemma 5.26 there is a leaf x in G. According Theorem 5.21 the degree of a cut-vertex is at least two, so x is not a cut-vertex and hence $G-x$ is connected. Clearly, $G-x$ does not have cycles (removing vertices and edges cannot introduce cycles), so $G-x$ is a tree with less than n vertices. By the induction hypothesis, $m' = n'-1$, where $m' = m(G-x)$ and $n' = n(G-x)$. But $m' = m-1$ and $n' = n-1$ since x is a leaf, whence $m = n-1$. □

Theorem 5.28 *Let G be a graph with n vertices and m edges. If $m = n-1$ and G has no cycles then G is connected (hence a tree).*

Proof. Suppose that $m = n-1$, G has no cycles, and G is not connected. Let S_1, ..., S_ω be the connected components of G, $\omega \geqslant 2$. Each connected component is a tree, so $m_i = n_i - 1$ for all i, where $m_i = m(S_i)$ and $n_i = n(S_i)$. Therefore $\sum_{i=1}^{\omega} m_i = \sum_{i=1}^{\omega} n_i - \omega$ i.e. $m = n - \omega$ (since $m = \sum_{i=1}^{\omega} m_i$ and $n = \sum_{i=1}^{\omega} n_i$). Now, $\omega \geqslant 2$ leads to contradiction: $m = n - \omega < n - 1 = m$. □

Theorem 5.29 *Let G be a connected graph with $n \geqslant 2$ vertices and m edges and let $m = n-1$. Then G has no cycles (and hence it is a tree).*

Proof. According to Theorem 5.25 the graph $G = (V,E)$ has a spanning tree $H = (V,E')$. Since H is a tree Theorem 5.27 yields $m(H) = n(H) - 1 = n - 1$. Assumption $m = n-1$ now implies $m(H) = m$ and thus from $E' \subseteq E$ we conclude $E' = E$. Therefore, $G = H$ and so G is a tree. □

Corollary 5.30 *A connected graph with n vertices and m edges is a tree if and only if $m = n-1$.*

We shall conclude the section by a result on the number of distinct trees. Let us first note that when counting structures we can count distinct structures and non-isomorphic structures. For example, there are 16 distinct trees on a four element set, but only two nonisomorphic, see Fig. 5.14. It is not surprising that counting nonisomorphic structures is more difficult.

Theorem 5.31 (Cayley 1889) *There are n^{n-2} distinct trees with n vertices.*

Proof. Let $V = \{1,\ldots,n\}$ be a finite set that serves as a set of vertices. The proof we are going to present is due to H. Prüfer[1]. The idea is to encode each tree on V by a sequence of integers $(a_1,\ldots,a_{n-2})$ and thus provide a bijection $\varphi : \mathscr{T}_n \to \{1,2,\ldots,n\}^{n-2}$, where $\mathscr{T}_n$ denotes the set of all trees on V.

[1] H. Prüfer, *Neuer Beweis eines Satzes über Permutationen*, Archiv der Math. und Phys. (3) 27(1918), 142–144

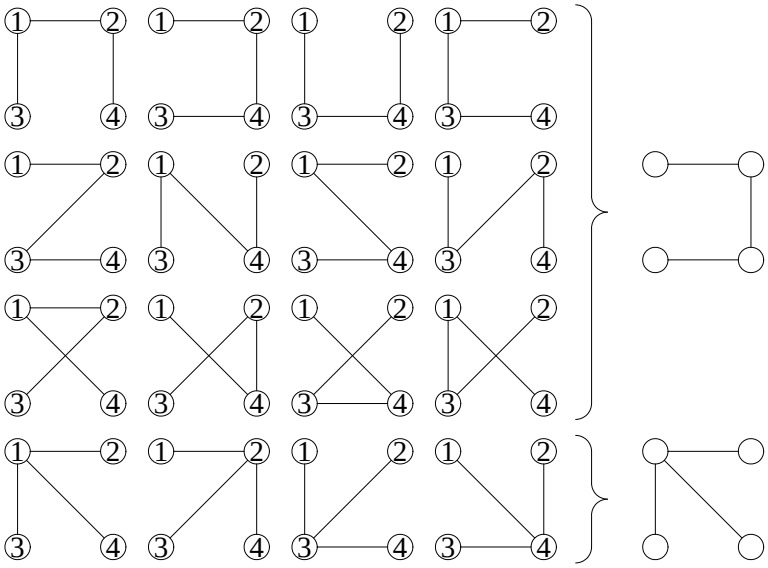


Figure 5.14: Sixteen distinct and only two nonisomorphic trees with four vertices

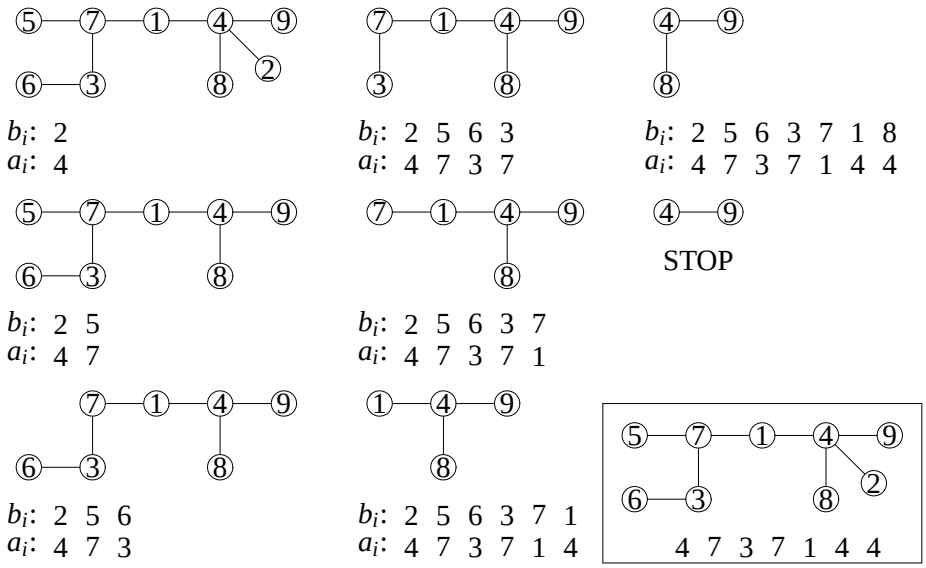


Figure 5.15: The Prüfer code of a tree

We first show how to construct the Prüfer code of a tree. Let T be a tree with the set of vertices V. We shall construct a sequence of trees (T_i) and two sequences of integers, the code (a_i) and an auxiliary sequence (b_i). Let $T_1 = T$. Given T_i, let b_i be the smallest leaf of the tree (vertices are integers, so out of all integers that appear as leaves we choose the smallest) and let a_i be its only neighbour. Now put $T_{i+1} = T_i - b_i$ and repeat until a tree with two vertices is obtained. The code of the tree is now $(a_1, a_2, \ldots, a_{n-2})$. An example is given in Fig. 5.15. Thus, we have a function $\varphi : \mathscr{T}_n \to \{1,\ldots,n\}^{n-2}$ that takes a tree onto its Prüfer code.

Conversely, given a sequence $(a_1,\ldots,a_{n-2})$ we can construct the tree as follows. For $S \subseteq \{1,\ldots,n\}$ let $\operatorname{mix} S = \min(\{1,\ldots,n\} \setminus S)$ denote the minimal number not in S (*mi*mimal e*x*cluded). Put $a_{n-1} = n$ and then construct b_1, b_2, …, b_{n-1} by

$$b_i = \operatorname{mix}\{a_i,\ldots,a_{n-1},b_1,\ldots,b_{i-1}\}$$

(for $i = 1$ there are no b_j's in the set). For example in case of $(4,7,3,4,1,4,4)$ we have $a_8 = 9$ and:

$$\begin{aligned}
b_1 &= \operatorname{mix}\{4,7,3,4,1,4,4,9\} = 2\\
b_2 &= \operatorname{mix}\{\ \ 7,3,4,1,4,4,9,2\} = 5\\
b_3 &= \operatorname{mix}\{\ \ \ \ 3,4,1,4,4,9,2,5\} = 6\\
b_4 &= \operatorname{mix}\{\ \ \ \ \ \ 4,1,4,4,9,2,5,6\} = 3\\
b_5 &= \operatorname{mix}\{\ \ \ \ \ \ \ \ 1,4,4,9,2,5,6,3\} = 7\\
b_6 &= \operatorname{mix}\{\ \ \ \ \ \ \ \ \ \ 4,4,9,2,5,6,3,7\} = 1\\
b_7 &= \operatorname{mix}\{\ \ \ \ \ \ \ \ \ \ \ \ 4,9,2,5,6,3,7,1\} = 8\\
b_8 &= \operatorname{mix}\{\ \ \ \ \ \ \ \ \ \ \ \ \ \ 9,2,5,6,3,7,1,8\} = 4
\end{aligned}$$

This process is called the *reconstruction procedure* since, as we shall see, it produces a tree whose Prüfer code is $(a_1,\ldots,a_{n-2})$.

Let us show that $\{\{b_i,a_i\} : 1 \leqslant i \leqslant n\}$ is the set of edges of a tree. If $i < j$ then, by construction, $b_j = \operatorname{mix}\{a_j,\ldots,a_{n-1},b_1,\ldots,b_i,\ldots,b_{j-1}\}$, so $b_j \neq b_i$. We see that all b_i's are distinct and smaller than $n = a_{n-1}$. Therefore, $\{b_1,\ldots,b_{n-1}\} = \{1,\ldots,n-1\}$ and hence $\{b_1,\ldots,b_{n-1},a_{n-1}\} = \{1,\ldots,n-1,n\}$. Moreover, if $i \leqslant j$ then $a_j \notin \{b_1,\ldots,b_j\}$ since $b_i = \operatorname{mix}\{a_i,\ldots,a_j,\ldots,a_{n-1},b_1,\ldots,b_{i-1}\}$, so from $\{b_1,\ldots,b_{n-1},a_{n-1}\} = \{1,\ldots,n-1,n\}$ it follows that $a_j \in \{b_{j+1},\ldots,b_{n-1},a_{n-1}\}$. To summarize,

$$\begin{aligned}
&a_j \in \{b_{j+1},b_{j+2},\ldots,b_{n-1},a_{n-1}\} \text{ and}\\
&b_j \notin \{a_{j+1},b_{j+1},a_{j+2},b_{j+2},\ldots,a_{n-1},b_{n-1}\},
\end{aligned} \qquad \text{for all } j. \qquad (\star)$$

To build the graph we start from $\{b_{n-1},a_{n-1}\}$ and then add edges $\{b_{n-2},a_{n-2}\}$, $\{b_{n-3},a_{n-3}\}$, …, $\{b_1,a_1\}$ one by one. From $(\star)$ it follows that at each step we

extend the graph by one new vertex b_i and one new edge $\{b_i, a_i\}$ that connects the new vertex to an existing one. Therefore, the graph we obtain at the end is connected, and a connected graph with n vertices and $n-1$ edges has to be a tree (Corollary 5.30). Thus, we have a function $\psi : \{1,\ldots,n\}^{n-2} \to \mathscr{T}_n$ that takes a code and produces a tree.

To complete the proof, we have to show that φ and ψ are inverses of one another, i.e. $\varphi \circ \psi = \mathrm{id}$ and $\psi \circ \varphi = \mathrm{id}$. We show only $\psi \circ \varphi = \mathrm{id}$ i.e. $\psi(\varphi(T)) = T$ for all $T \in \mathscr{T}_n$. For a tree T, a vertex $v \in V(T)$ is an *internal vertex* T if $\delta_T(v) > 1$. Let $\mathrm{int}(T)$ denote the set of all internal vertices of T.

Take any $T \in \mathscr{T}_n$, let $(a_1,\ldots,a_{n-2})$ be its Prüfer code and $(b_1,\ldots,b_{n-2})$ the auxiliary sequence. At the end of the procedure of constructing the Prüfer code two vertices remain the the graph, the vertex $a_{n-1} = n$ and its neighbour whom we denote by b_{n-1}. Starting from $(a_1,\ldots,a_{n-1})$ the reconstruction procedure produces a sequence of integers b'_1, …, b'_{n-1}. We will show that $b_i = b'_i$ for all i. Assume also that $n \geqslant 3$.

Since b_1 is adjacent to a_1 in T and $n \geqslant 3$, a_1 cannot be a leaf of T so $a_1 \in \mathrm{int}(T)$. The same argument shows that $a_2 \in \mathrm{int}(T - b_1)$, $a_3 \in \mathrm{int}(T - b_1 - b_2)$, and in general, $a_{i+1} \in \mathrm{int}(T - b_1 - \ldots - b_i)$. Since $\mathrm{int}(T - v) \subseteq \mathrm{int}(T)$ whenever v is a leaf of T and $n(T) \geqslant 2$, it follows that $\mathrm{int}(T - b_1 - \ldots - b_i) = \{a_{i+1},\ldots,a_{n-2}\}$. In particular, $\mathrm{int}(T) = \{a_1,\ldots,a_{n-2}\}$. Since each vertex of a tree with at least two vertices is either a leaf or an internal vertex we obtain that

$$V(T - b_1 - \ldots - b_i) \setminus \mathrm{int}(T - b_1 - \ldots - b_i)$$

is the set of leaves of $T - b_1 - \ldots - b_i$. Now $V(T - b_1 - \ldots - b_i) = \{1,\ldots,n\} \setminus \{b_1,\ldots,b_i\}$ and $\mathrm{int}(T - b_1 - \ldots - b_i) = \{a_{i+1},\ldots,a_{n-2}\}$, so the set of leaves of $T - b_1 - \ldots - b_i$ is

$$\Big(\{1,\ldots,n\} \setminus \{b_1,\ldots,b_i\}\Big) \setminus \{a_{i+1},\ldots,a_{n-2}\} =$$
$$= \{1,\ldots,n\} \setminus \{a_{i+1},\ldots,a_{n-2},b_1,\ldots,b_i\}.$$

It is now easy to show that $b_i = b'_i$ by induction on i. As we have seen, b_1 is a leaf of T, so $b_1 \in \{1,\ldots,n\} \setminus \{a_1,\ldots,a_{n-2}\}$. But b_1 is the smallest such integer, whence $b_1 = \min(\{1,\ldots,n\} \setminus \{a_1,\ldots,a_{n-2}\}) = \mathrm{mix}\{a_1,\ldots,a_{n-2}\} = b'_1$. Assume that $b_j = b'_j$ for all $j \in \{1,\ldots,i\}$ and consider b_{i+1}. It is the smallest leaf in $T - b_1 - \ldots - b_i$ so, with the help of induction hypothesis

$$\begin{aligned} b_{i+1} &= \min(\{1,\ldots,n\} \setminus \{a_{i+1},\ldots,a_{n-2},b_1,\ldots,b_i\}) \\ &= \mathrm{mix}\{a_{i+1},\ldots,a_{n-2},b_1,\ldots,b_i\} = \mathrm{mix}\{a_{i+1},\ldots,a_{n-2},b'_1,\ldots,b'_i\} = b'_{i+1} \end{aligned}$$

Therefore, $\{a_i,b_i\} = \{a_i,b'_i\}$ for all i and the tree produced by the reconstruction procedure is T, the tree we started with. □

5.4 Digraphs

A *digraph* is an ordered pair $D = (V,E)$ where V is a nonempty finite set and E is an arbitrary subset of V^2 such that $(x,x) \notin E$ for all $x \in V$. Elements of V are called *vertices* of D, while elements of E are called *edges* of D. We shall often write $V(D)$ and $E(D)$ to denote the set of vertices and the set of edges of D, and $n(D)$ and $m(D)$ to denote the number of vertices and the number of edges of D. Instead of $(x,y) \in E$ we often write $x \to y$ or $x \underset{D}{\to} y$. If $x \to y$ we say that x is a *predecessor* of y and y is a *successor* of x. The number of edges that go out of v iz called the *out-degree* of v and will be denoted by $\delta_D^+(v)$. The number of edges that go into v is called the *indegree* of v and will be denoted by $\delta_D^-(v)$. Further, let,

$$I_D(v) = \{x \in V : x \to v\}, \qquad O_D(v) = \{x \in V : v \to x\},$$

denote the set of predecessors and the set of successors of v. Clearly, $\delta_D^-(v) = |I_D(v)|$ and $\delta_D^+(v) = |O_D(v)|$. The *total degree* of a vertex v is $\delta_D(v) = \delta_D^-(v) + \delta_D^+(v)$. If D is clear from the context, we simply write $\delta^-(v)$, $\delta^+(v)$, $I(v)$, $O(v)$ and $\delta(v)$.

A *source* of a digraph D is a vertex $v \in V(D)$ such that $\delta^-(v) = 0$ and $\delta^+(v) > 0$. A *sink* of a digraph D is a vertex $v \in V(D)$ such that $\delta^-(v) > 0$ and $\delta^+(v) = 0$. A *back-edge* in a digraph D is an edge $(x,y) \in E(D)$ such that $(y,x) \in E(D)$. If D has no back-edges then $I(v) \cap O(v) = \varnothing$ for every $v \in V(D)$.

If v is a vertex and e an edge of a digraph D then $D - e$ denotes the digraph obtained from D by removing the edge e, while $D - v$ denotes the digraph obtained from D by removing v, the edges that go into v and the edges that go out of v.

Digraphs also have a very natural graphical representation. Vertices are represented as points in a plane, while an edge $x \to y$ is represented as a directed curve (usually an arrow) going from x to y. Fig. 5.16 (a) depicts a digraf with 10 vertices.

Theorem 5.32 (The First Theorem for Digraphs) *Let $D = (V,E)$ be a digraph with m edges. Then $\sum_{v \in V} \delta^-(v) = \sum_{v \in V} \delta^+(v) = m$.*

Digraphs $D_1 = (V_1, E_1)$ and $D_2 = (V_2, E_2)$ are *isomorphic* if there exists a bijection $\varphi : V_1 \to V_2$ such that $(x,y) \in E_1$ if and only if $(\varphi(x), \varphi(y)) \in E_2$. The bijection φ is referred to as as *isomorphism* and we write $D_1 \cong D_2$.

The notions of the oriented path, oriented cycle and oriented walk in a digraph are straightforward generalizations of their "unoriented" versions. An *oriented walk* is a sequence of vertices and edges $x_0\ e_1\ x_1 \ldots x_{k-1}\ e_k\ x_k$ such that $e_i = (x_{i-1}, x_i)$. We say that k is the length of the walk. An *oriented path* is an oriented walk where all vertices and all edges are distinct. An *oriented cycle* is an oriented walk where all edges and vertices are distinct, with the exception of $x_0 = x_k$.

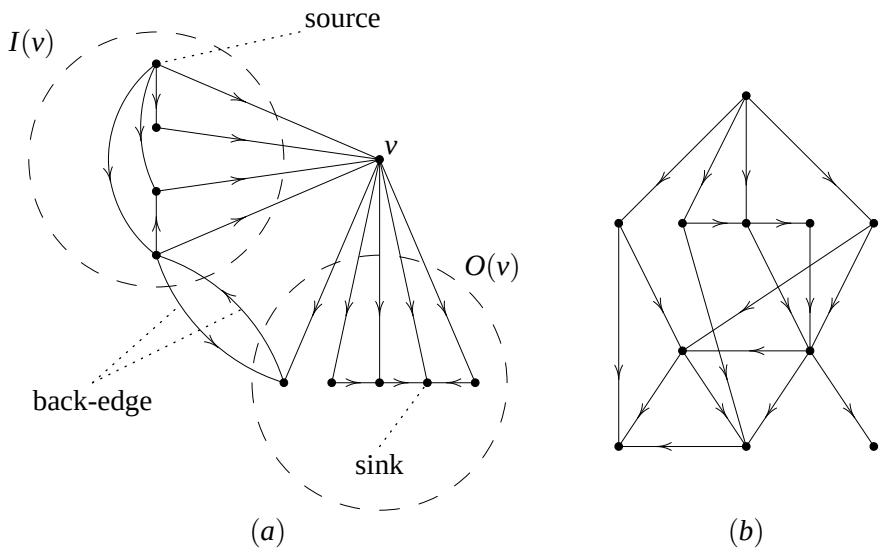


Figure 5.16: Two digraphs

Theorem 5.33 *Let D be a digraph with at least one edge. If D has no sinks, then it has an oriented cycle. Dually, if D has no sources, it has an oriented cycle.*

A digraph is *acyclic* if it has no oriented cycles. Fig. 5.16 (b) is an example of an acyclic digraph.

Corollary 5.34 *Each acyclic digraph with at least one edge has both a source and a sink.*

Theorem 5.35 *A digraph D with n vertices is acyclic if and only if it is possible to arrange its vertices as $(v_1,\ldots,v_n)$ in such a way that $v_i \to v_j$ implies $i < j$.*

Proof. $(\Leftarrow)$ If such an arrangement of vertices exists then clearly G has no oriented cycles.

$(\Rightarrow)$ We use induction on n. Cases $n = 1$ and $n = 2$ are easy. Assume that such an arrangement of vertices exists for all acyclic digraphs with less than n vertices and let D be an acyclic digraph with n vertices. If there is a vertex x such that $\delta(v) = 0$ put $v_1 = x$. Otherwise, D has at least one edge, so it has a source. Let v_1 be any source of D. Now, $D - v_1$ is again an acyclic digraph and by induction hypothesis its vertices can be arranged into a sequence $(v_2,\ldots,v_n)$ in such a way that $v_i \to v_j$ implies $i < j$ for all $i, j \geqslant 2$. Since $I(v_1) = \varnothing$ and $O(v_1) \subseteq \{v_2,\ldots,v_n\}$, it is easy to see that $(v_1, v_2,\ldots,v_n)$ is the required arrangement of vertices of D. □

A digraph $D' = (V', E')$ is a *subdigraph* of a digraph $D = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. We write $D' \leqslant D$. For $S \subseteq V$, the *subdigraph induced by* S is the digraph $D[S] = (S, S^2 \cap E)$.

We say that $S \subseteq V(D)$ *dominates* D if $D[S]$ has no edges and the following holds: for every $x \in V(D) \setminus S$ there is an $s \in S$ such that either $s \to x$ or $s \to y \to x$ for some $y \in V(D)$.

Theorem 5.36 (Chvátal, Lovász 1974) *For every digraph D there is a set of vertices $S \subseteq V(D)$ which dominates D.*

Proof. We use induction on $n = n(D)$. For $n = 1$ or $n = 2$ the claim is obvious. Suppose the claim is true for all digraphs with less than n vertices and let D be a digraph with $n \geqslant 3$ vertices. Take any $x \in V(D)$ and let $A = V(D) \setminus (\{x\} \cup O(x))$. If $A = \varnothing$ then $S = \{x\}$ dominates D. If, however, $A \neq \varnothing$, by the induction hypothesis the digraph $D[A]$ has a set of vertices $S' \subseteq A$ that dominates $D[A]$. If there are no edges in $D[S' \cup \{x\}]$ then $S = S' \cup \{x\}$ dominates D. Otherwise, there is a $z \in S'$ such that $x \to z$ or $z \to x$ in D. From $z \notin O(x)$ we conclude that $z \to x$ in D, so $S = S'$ dominates D. □

There are two natural notions of connectedness for digraphs. It seems natural to be able to go from any vertex to any other vertex respecting the orientation of the edges, but sometimes we might wish to be able to do the same thing regardless of the orientation of edges.

A *base* of a digraph $D = (V, E)$ is a graph $G = (V, E')$ where $E' = \{\{x, y\} : (x, y) \in D\}$. A base of a digraph is obtained by replacing oriented edges of the digraph by nonoriented edges, see Fig. 5.17. A digraph D is *weakly connected* if

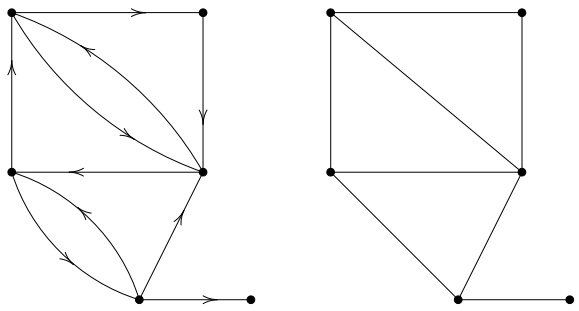


Figure 5.17: A digraph and its base

its base is a connected graph. A digraph D is *strongly connected* if for every pair of vertices $x, y \in V$, $x \neq y$, there is an oriented path going from x to y, see Fig. 5.18.

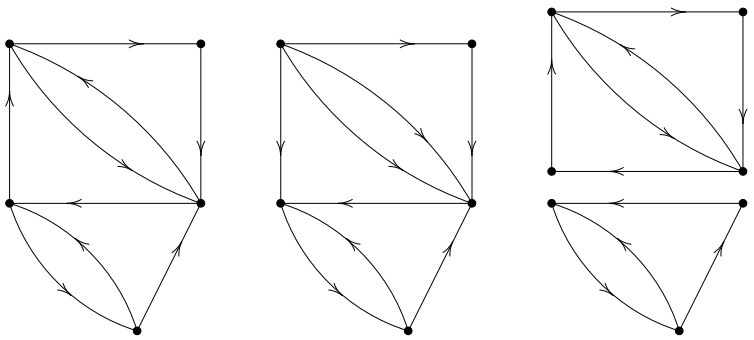


Figure 5.18: Two types of connectedness for digraphs

For disjoint $A,B \subseteq V(D)$ let $E(A,B) = \{(x,y) \in E(D) : x \in A, y \in B\}$ be the set of all edges of D that go from a vertex in A to a vertex in B.

Theorem 5.37 *A digraph D is weakly connected if and only if $E(A,B) \neq \varnothing$* ***or*** *$E(B,A) \neq \varnothing$ for every partition $\{A,B\}$ of $V(D)$.*

A digraph D is strongly connected if and only if $E(A,B) \neq \varnothing$ ***and*** *$E(B,A) \neq \varnothing$ for every partition $\{A,B\}$ of $V(D)$.*

Proof. We shall prove the second part of the theorem.

($\Rightarrow$) Let D be a strongly connected digraph and let $\{A,B\}$ be an arbitrary partition of $V(D)$. Take any $a \in A$ and any $b \in B$. The digraph D is strongly connected, so there exists an oriented path from a to b. Since $a \in A$ and $b \in B$, the path has to cross from A into B at some point, so there exists an edge (x,y) along this path such that $x \in A$ and $y \in B$. Therefore, $E(A,B) \neq \varnothing$. Similarly, $E(B,A) \neq \varnothing$.

($\Leftarrow$) Take any $x,y \in V(D)$, $x \neq y$, and let us show that there is an oriented path from x to y. Let $A = \{x\} \cup \{v \in V(D) : \text{there is an oriented path from } x \text{ to } v\}$. We wish to show that $y \in A$. Suppose this is not the case and let $B = V(D) \setminus A$. Then $y \in B$ and so $B \neq \varnothing$. Now, $\{A,B\}$ is a partition of $V(D)$ and by the assumption $E(A,B) \neq \varnothing$. This means that there is a $v \in A$ and a $w \in B$ such that $v \to w$. But $v \in A$ means that there is an oriented path from x to v, so $v \to w$ implies that there is an oriented path from x to $w \notin A$. This contradiction shows that $y \in A$ and hence there is an oriented path from x to y. □

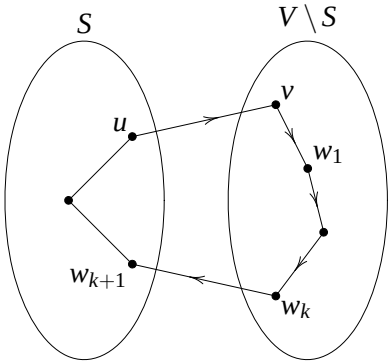
Every connected graph $G = (V,E)$ can be turned into a strongly connected digraph $D(G) = (V,E')$ where $E' = \{(x,y) : \{x,y\} \in E\}$, that is, by replacing each

edge $\{x,y\}$ of G by a pair of edges (x,y), (y,x). Therefore, each connected graph is a base of some strongly connected digraph, possibly with back-edges. The following theorem shows that this is not the case if we forbid back-edges.

Theorem 5.38 *A connected graph G with at least two vertices is a base of a strongly connected digraph with no back-edges if and only if G has no cut-edges.*

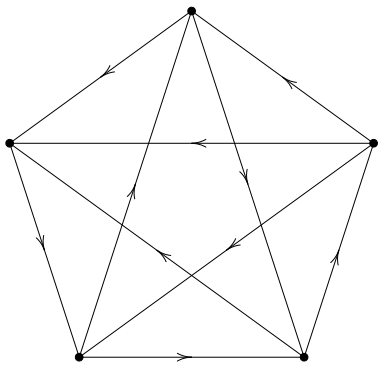
Proof. ($\Rightarrow$) Let $G = (V, E_G)$ be a base of a digraph $D = (V, E_D)$ and suppose that G has a cut-edge $e = \{u, v\}$. Then by Theorem 5.20 there is a partition $\{A, B\}$ of V such that $E_G(A, B) = \{e\}$. Since D has no back-edges then either $(u, v) \in E_D$ or $(v, u) \in E_D$, but not both. Therefore, either $E_D(A, B) = \varnothing$ or $E_D(B, A) = \varnothing$. In any case, D is not strongly connected by Theorem 5.37.

($\Leftarrow$) Let $G = (V, E)$ be a graph with no cut-edges and let $S \subseteq V$ be a maximal set of vertices such that $G[S]$ is a base of a strongly connected digraph $D(S)$ with no back-edges. Let us show that $S \neq \varnothing$. Note first that G contains a cycle (G has no cut-edges, so by Theorem 5.20 every edge of G belongs to a cycle; hence there is at least one cycle in G). Take any cycle C in G, orient its edges to obtain an oriented cycle and orient the remaining edges in $G[V(C)]$ arbitrarily. We thus obtain a strongly connected digraph $D(C)$ with no back-edges whose base is $G[V(C)]$. Therefore, there exists a set $S' \subseteq V$ with at least three vertices such that $G[S']$ is a base of a strongly connected digraph with no back-edges, so the maximal such set cannot be empty.



Let us show that $S = V$. Suppose to the contrary that $S \subset V$, i.e. $V \setminus S \neq \varnothing$. Since G is connected we have $E(S, V \setminus S) \neq \varnothing$, so take any $e = \{u, v\}$ such that $u \in S$ and $v \in V \setminus S$. There are no cut-edges in G so according to Theorem 5.20 the edge e belongs to a cycle in G. Let $v\, w_1 \ldots w_k$ be a part of the cycle that belongs to $V \setminus S$ and let w_{k+1} be the vertex that follows w_k on the cycle. By assumption, $w_{k+1} \in S$. Now orient the edges on the path $u\, v\, w_1 \ldots w_k\, w_{k+1}$ to obtain an oriented path that goes from u to w_{k+1} and attach the path to the digraph $D(S)$. Orient the remaining edges in $G[S \cup \{v, w_1, \ldots w_k\}]$ arbitrarily. The digraph D' obtained this way is strongly connected, has no back-edges and its base is $G[S \cup \{v, w_1, \ldots w_k\}]$ whose set of vertices is a proper superset of S. This contradiction shows that $S = V$, i.e. that G is a base of a strongly connected digraph with no back-edges. □

5.5 Tournaments



A *tournament* is a digraph $T = (V, E)$ with the property that for each pair $x, y \in V$, $x \neq y$, either $(x, y) \in T$ or $(y, x) \in T$. Equivalently, a tournament is a digraph with no back-edges whose base is a complete graph. Tournaments (as digraphs) appear as models of tournaments (as sprot events) where no match ends in a draw; each arrow then represents one match and goes from the vertex representing the winner to the vertex representing the loser.

A tournament with n vertices has $\binom{n}{2}$ edges and $\delta^+(v) + \delta^-(v) = n-1$ for each vertex v. Therefore, it has become customary to consider only $\delta^+(v)$. When working with tournaments, $\delta^+(v)$ is called the *score* of v and denoted by $s(v)$. A tournament is *transitive* if $x \to y$ and $y \to z$ implies $x \to z$ whenever x, y and z are three distinct vertices of the tournament.

Theorem 5.39 *Let T be a tournament with n vertices. Then the following are equivalent:*

(1) T is an acyclic tournament;

(2) T is a transitive tournament;

(3) the scores of vertices in T are $0, 1, \ldots, n-1$.

Proof. $(1) \Rightarrow (2)$: Suppose T is not a transitive tournament. Then there exist distinct vertices x, y and z such that $x \to y$ and $y \to z$ but $x \not\to z$. Since T is a tournament, $x \not\to z$ means that $z \to x$ and we obtain a cycle $x \to y \to z \to x$.

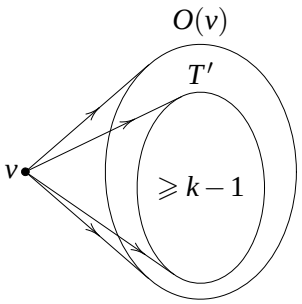
$(2) \Rightarrow (3)$: The proof is by induction on n. Cases $n = 2$ and $n = 3$ are trivial. Suppose that in each transitive tournament with $k < n$ vertices the scores of vertices are $0, 1, \ldots, k-1$ and let T be a transitive tournament with n vertices. Let v_1 be the vertex of T with maximal score and let us show that $s(v_1) = n-1$. Suppose that there is a vertex x such that $x \to v_1$. Then due to transitivity $v_1 \to z$ implies $x \to z$ and hence $s(x) \geqslant 1 + s(v_1) > s(v_1)$, which is impossible. Therefore, $v_1 \to x$ for all $x \neq v_1$ and hence $s(v_1) = n-1$. It is easy to see that $T - v_1$ is again a transitive tournament and by the induction hypothesis the scores of its vertices are $0, 1, \ldots, n-2$. Therefore, the scores of vertices in T are $0, 1, \ldots, n-2, n-1$.

$(3) \Rightarrow (1)$: The proof is again by induction on n and the cases $n = 2$ and $n = 3$ are trivial. Suppose that each tournament with $k < n$ vertices and with scores $0, 1,$

$\ldots, k-1$ is acyclic and let T be a tournament with n vertices and scores 0, 1, $\ldots$, $n-1$. Let v be the vertex of T whose score is $n-1$ and let C be an oriented cycle in T. Since $T-v$ is a tournament with scores 0, 1, $\ldots$, $n-2$, it is acyclic according to the induction hypothesis so $V(C) \not\subseteq V(T-v)$. Therefore, C has to pass through v. On the other hand, $s(v) = n-1$ means that $v \to x$ for every $x \neq v$ so no cycle in T can pass through v. Contradiction. □

Corollary 5.40 *Two transitive tournaments are isomorphic if and only if they have the same number of vertices.*

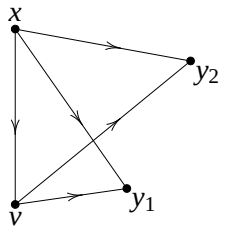
Theorem 5.41 *Every tournament with at least 2^{k-1} vertices, $k \geqslant 2$, has a transitive subtournament with at least k vertices.*



Proof. The proof is by induction on k. If $k = 2$ the tournament has at least two vertices and hence at least one edge, so each edge $x \to y$ is a transitive subtournament with two vertices. Assume the claim is true for all integers less than k and consider a tournament T with at least 2^{k-1} vertices. Take any $v \in V(T)$. Then $V(T) = I(v) \cup \{v\} \cup O(v)$, so one of the sets $I(v)$, $O(v)$ has at least 2^{k-2} vertices. Without loss of generality we can assume that $|O(v)| \geqslant 2^{k-2}$. Inductioin hypothesis now yields that there is a transitive subtournament T' of $T[O(v)]$ with at least $k-1$ vertices. Then T' together with v induces a transitive subtournament of T with at least k vertices. □

A *king* in a tournament T is a vertex $v \in V(T)$ such that $\{v\}$ dominates T. This means that for every $x \neq v$ either $v \to x$ or $v \to y \to x$ for some $y \in V(T)$.

Theorem 5.42 *Each tournament with at least two vertices has a king.*



Proof. Let v be a vertex of T whose score is maximal and let us show that v is a king. Suppose to the contrary that v is not a king. Then there is an $x \neq v$ such that $v \not\to x$ and no $y \in V(T)$ satisfies $v \to y \to x$. Since T is a tournament, $v \not\to x$ means $x \to v$, while the other condition means that if $v \to y$ then $x \to y$. But then $s(x) \geqslant 1 + s(v) > s(v)$, which contradicts the maximality of $s(v)$. □

Homework

5.1. An *automorphsm* of a graph G is every isomorhism $\varphi : V(G) \to V(G)$ from the graph onto itself. By $\mathrm{Aut}(G)$ we denote the set of all the auto-

morphisms of G.

(a) Show that $(\mathrm{Aut}(G), \circ)$ is a group.

(b) Describe $\mathrm{Aut}(K_n)$, $\mathrm{Aut}(S_n)$ and $\mathrm{Aut}(P_n)$ for $n \geqslant 3$.

(c) Show that $\mathrm{Aut}(G) = \mathrm{Aut}(\overline{G})$.

5.2. (a) Show that for every $n \geqslant 6$ there exists a graph G with n vertices such that $|\mathrm{Aut}(G)| = 1$.

(b) Show that for every $k \geqslant 2$ and every $n \geqslant k+3$ there exists a graph G with n vertices such that $|\mathrm{Aut}(G)| = k!$.

5.3. Prove Lemma 5.12.

5.4. Prove Theorem 5.16.

5.5. If G is not connected show that $d(\overline{G}) \leqslant 2$. (We know that $\overline{G}$ is connected).

5.6. Prove Theorem 5.21.

5.7. Prove Theorem 5.22.

5.8. Show that a graph is a tree if and only if each pair of distinct vertices of of the graph is connected by a unique path.

5.9. Find the number of distinct spanning trees of K_n.

5.10. Complete the proof of Theorem 5.31 by showing that $\varphi \circ \psi = \mathrm{id}$.

5.11. Prove Theorem 5.33.

5.12. In the distant land of Xÿç there are n cities some of which are connected by roads, but still it is possible to reach each city from every other city by traveling along the roads (and possibly passing through some other cities). The Evil Magician who rules the Xÿç would like to terrorize his people by making each road a one-way road in such a way that after leaving a city it is impossible to get back. Show that it is possible to do such a thing.

5.13. Prove the first part of Theorem 5.37 (the characterization of weak connectedness).

5.14. Prove Corollary 5.40.

5.15. A tournament is *regular* if $s(x) = s(y)$ for all x and y. Show that in a regular tournament each vertex is a king.

Exercises

5.16. Let G be a graph with n vertices and m edges. Show that $\Delta(G) \geqslant \frac{2m}{n}$.

5.17. Which of the following integer sequences can be a sequence of degrees of vertices of a graph?

(a) $(1,2,2,4,5,6,7)$;

(b) $(1,1,2,2,2,3,3)$;

(c) $(1,1,3,3,3,3,5,6,8,9)$.

†**5.18.** Show that there are

(a) $2^{\binom{n}{2}}$ distinct graphs with n vertices;

(b) $2^{\binom{n-1}{2}}$ distinct graphs with n vertices such that the degree of each vertex in the graph is even.

5.19. Let G be a graph with $\delta(G) \geqslant 2$. Then G contains a path of length $\geqslant \delta(G)$ and a cycle of length $\geqslant \delta(G)+1$.

5.20. Let G be a bipartite graph (not necessarily a complete bipartite graph!) with n vertices and m edges. Show that $m \leqslant \frac{1}{4}n^2$.

5.21. Show that a graph G is bipartite if and only if every subgraph H of G satisfies $\alpha(H) \geqslant \frac{1}{2}n(H)$.

5.22. A *k-dimensional hypercube* is a graph $Q_k = (V_k, E_k)$ where V_k is the set of all 01-words of length k and $a_1 \ldots a_k$, $b_1 \ldots b_k \in V_k$ are adjacent if and only if the two words differ at exactly one place. For example, if $k = 4$ then 0101 and 0001 are adjacent in Q_4 while 0101 and 0000 are not.

(a) Find the number of vertices and the number of edges of Q_k.

(b) Show that Q_k is bipartite.

(c) Find $d(Q_k)$.

5.23. Show that for every even $n \geqslant 6$ there exists a connected regular graph of degree 3 with n vertices and with no trianlges.

5.24. Show that if $\delta(G) \geqslant \frac{1}{2}n(G)$ then G is connected and $d(G) \leqslant 2$.

5.25. Show that for every graph G there exists a regular graph H such that G is an induced subgraph of H and $\Delta(G) = \Delta(H)$.

5.26. Show that $\delta(\overline{G}) = (n(G)-1) - \Delta(G)$ and $\Delta(\overline{G}) = (n(G)-1) - \delta(G)$.

5.27. Show the following:

(a) If G is connected and $d(G) \geqslant 3$ then $\overline{G}$ is connected and $d(\overline{G}) \leqslant 3$.

(b) Every selfcomplementary graph G with at least two vertices is connected and $2 \leqslant d(G) \leqslant 3$.

5.28. Suppose that the degree of every vertex in a connected graph G is even. Show that $\omega(G-v) \leqslant \frac{1}{2}\delta(v)$ for all $v \in V(G)$.

5.29. Let $G = (V,E)$ be a connected graph with n vertices and let u be an arbitrary vertex of G. Show that $\sum_{x \in V} d(u,x) \leqslant \binom{n}{2}$.

5.30. Let G be a connected graph with at least two vertices. Show that G has at least two vertices that are not cut-vertices.

5.31. Show that if v is a cut-vertex of G, then v is not a cut-vertex of $\overline{G}$.

5.32. Show that each tree G has at least $\Delta(G)$ leafs.

5.33. Let T be a tree, $\Delta = \Delta(T)$ and f_k the number of vertices in T of degree k. Show that $f_1 = 2 + \sum_{k=3}^{\Delta} (k-2) f_k$.

5.34. Find all trees G such that $\overline{G}$ is a tree.

5.35. For every $n \geqslant 4$ find a graph G with n vertices such that for each $k \in \{2,\ldots,n-2\}$ there is a spanning tree of G whose diameter is k.

5.36. Note first that each tree is a bipartite graph since no cycles means no odd cycles. Let $\{X,Y\}$ be a partition of the vertices of a tree T which demonstrates that T is a bipartite graph and assume that $|X| = |Y| + p$ for some $p > 0$. Show that X contains at least $p+1$ leaves of T.

5.37. A *forest* is a graph whose connected components are trees. Show that G is a forest if and only if $\delta(H) \leqslant 1$ for all induced subgraphs H of G.

†**5.38.** How many nonisomorphic spanning trees does $K_{2,n}$ have?

†**5.39.** Show that each spanning tree of a connected graph contains all cut-edges of the graph.

†**5.40.** A block of a connected graph G is a maximal set of vertices $S \subseteq V(G)$ such that $G[S]$ has no cut-vertices (that is, if $S' \supseteq S$ and $G[S']$ has no cut-vertices then $S' = S$).

(a) Show that any two blocks of a graph have at most one vertex in common.

(b) Let $B_1, \ldots, B_k$ be blocks of G and let $\mathscr{B}_G$ be the graph with vertices

$\{1,\ldots,k\}$ where i is adjacent to j if and only if $i \neq j$ and B_i and B_j have a nonempty intersection. Show that $\mathscr{B}_G$ is a tree.

†**5.41.** Let $D = (V,E)$ be a weakly connected digraph. A strongly connected component of D is a maximal set of vertices $S \subseteq V$ such that $D[S]$ is strongly connected (that is, if $S' \supseteq S$ and $D[S']$ is strongly connected then $S' = S$).

(a) Show that $S \cap S' = \varnothing$ whenever S and S' are distinct strongly connected components of D.

(b) Let S and S' be distinct strongly connected components of D. Show that if $E(S,S') \neq \varnothing$ then $E(S',S) = \varnothing$.

(c) Let S_1, ..., S_k be strongly connected components of D and let $\mathscr{S}_D$ be the graph with vertices $\{1,\ldots,k\}$ where $i \to j$ if and only if $i \neq j$ and $E(S_i,S_j) \neq \varnothing$. Show that $\mathscr{S}_D$ has no back-edges and its base is a tree.

5.42. Show that $\sum_{v\in V} (\delta^+(v))^2 = \sum_{v\in V} (\delta^-(v))^2$ in every tournament $T = (V,E)$.

5.43. A tournament is *regular* if $s(x) = s(y)$ for all x and y. Show that for each odd integer $n \geqslant 3$ there exists a regular tournament with n vertices.

5.44. Scores $s_1 \leqslant s_2 \leqslant \ldots \leqslant s_n$ of a tournament T satisfy $\sum_{i=1}^{k} s_i = \binom{k}{2}$ for every $k \in \{1,\ldots,n\}$. Show that T is an acyclic tournament.

Chapter 6

Eulerian and Hamiltonian graphs

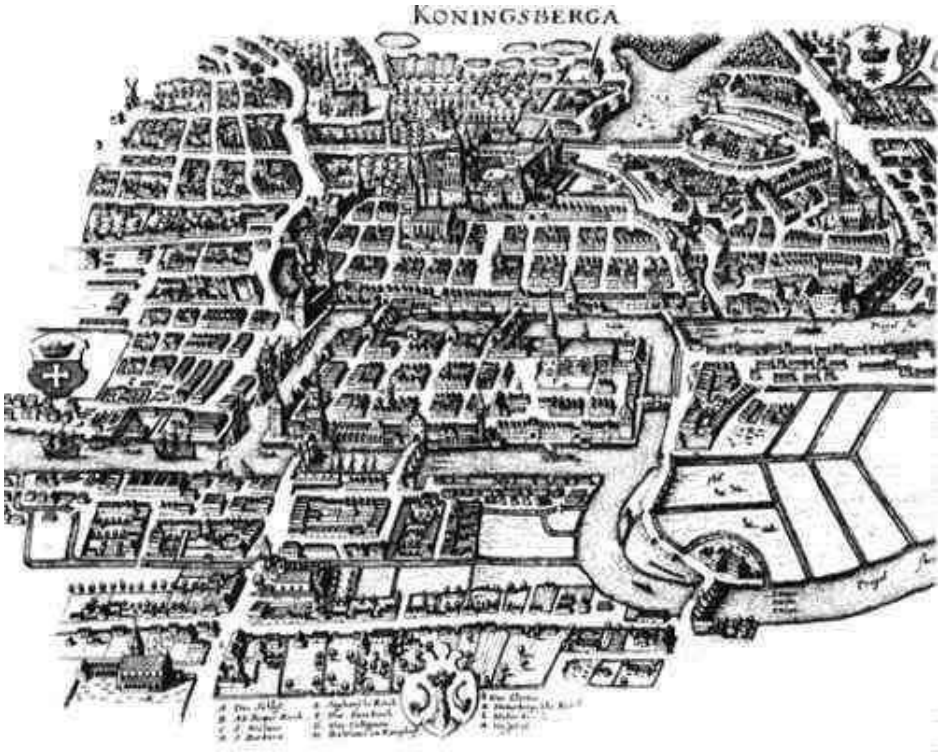
In this chapter we deal with two important classes of graphs:

- Eulerian graphs, which are graphs with the closed walk in which each edge occurs precisely once; and
- Hamiltonian graphs, which are graphs with the cycle in which every vertex occurs precisely once.

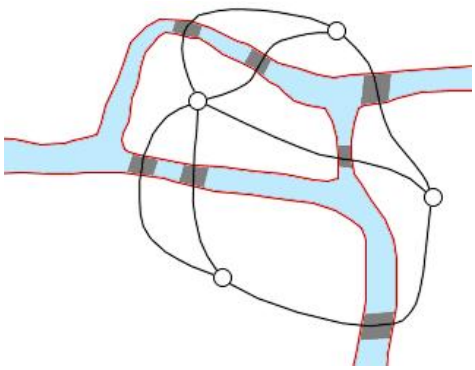
We present an easy characterisation of Eulerian graphs and discuss several necessary and sufficient conditions for a graph to be Hamiltonian. The fact that there is no "easy" and "useful" characterisation of Hamiltonian graphs is justified by the discussion at the end of the chapter where we argue that checking for a Hamiltonian cycle in a graph is an NP-complete problem.

6.1 Eulerian graphs

The famous Swiss mathematician Leonhard Euler was visiting the city of Königsberg in the year 1735. Königsberg was a city in Prussia situated on the Pregel River, which served as the residence of the dukes of Prussia in the 16th century. (Today, the city is named Kaliningrad, and is a major industrial and commercial center of western Russia.) The river Pregel flowed through the city such that in its center was an island, and after passing the island, the river broke into two parts. Seven bridges were built so that the people of the city could get from one part to another. A map of the center of Königsberg in 1735 looked like this:



A favorite pastime for visitors to the city was to try to cross each of the bridges of Königsberg exactly once. Euler was told by some people that it was impossible and by others that they doubted whether or not it could be done. No one believed it was possible. Eventually, Euler realized that all problems of this form could be represented by replacing areas of land by vertices, and the bridges to and from them by edges of a graph such as:



The problem now becomes to draw this picture without tracing any line twice and without picking the pencil up off the paper. All four of the vertices in the above picture have an odd degree. Take one of these vertices, say one of the ones of degree three. We could start at that vertex, and then arrive and leave later. But then we can't come back. So, every vertex with an odd degree has to be either the beginning or the end of the pencil-path and thus we can have at most two odd vertices. Therefore it is impossible to draw the above picture in one pencil stroke without tracing some line twice.

This is the first recorded problem in graph theory, and W. Tutte, himself a prominent graph-theorist, decided to celebrate the problem with a poem:

From Königsberg to König's book
by William T. Tutte

Some citizens of Koenigsberg
Were walking on the strand
Beside the river Pregel
With its seven bridges spanned.

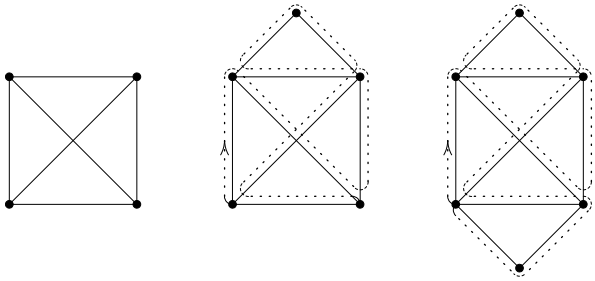
O, Euler, come and walk with us
Those burghers did beseech
We'll walk the seven bridges o'er
And pass but once by each.

"It can't be done" then Euler cried
"Here comes the Q.E.D.
Your islands are but vertices,
And all of odd degree."

We shall now go for a more formal treatment of this and similar problems. We shall first solve the general problem in case of oriented graphs, and then infer the solution in case of undirected graphs.

Definition 6.1 A *trail* in a graph is a walk in which edges are not allowed to repeat. An *Eulerian trail* in a graph is a trail that contains each edge of the graph precisely once. A graph is said to be *Eulerian* if it contains a closed Eulerian trail, Fig. 6.1.

Definition 6.2 Analogously, an *oriented trail* in a digraph is an oriented walk in which edges are not allowed to repeat. An *Eulerian trail* in a digraph is an oriented trail in the digraph that contains each edge of the digraph precisely once. A digraph is said to be *Eulerian* if it contains a closed Eulerian trail.

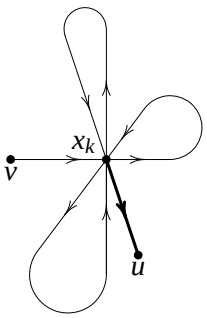


(a) (b) (c)

Figure 6.1: (a) A graph with no Eulerian trail; (b) a non-eulerian graph with an Eulerian trail; (c) an Eulerian graph

Lemma 6.3 *Let D be a digraph with no isolated vertices and with the property that $\delta^-(v) = \delta^+(v)$ for every $v \in V(D)$. Then every vertex of D belongs to a closed oriented trail in D.*

Proof. Let $W = v\, e_1\, x_1 \ldots e_k\, x_k$ be the longest trail in D that starts with v and let us show that $x_k = v$. Suppose to the contrary that $x_k \neq v$ and assume that x_k appears $l \geqslant 1$ times on the trail W. Each appearance of x_k on W engages one edge that leads into x_k and one edge that leads out of x_k, except for the last appearance of x_k that engages one edge leading into x_k. Therefore, W contains l edges leading into x_k and $l-1$ edges leading out of x_k. Since $\delta^-(x_k) = \delta^+(x_k)$, there exists an edge $e' = (x_k, u) \in E(D)$ that does not appear in W. Now, $v\, e_1\, x_1 \ldots e_k\, x_k\, e'\, u$ is a trail that starts from v longer that W. Contradiction. □



Theorem 6.4 *Let D be a digraph with no isolated vertices. Then D is an Eulerian digraph if and only if D is weakly connected and $\delta^-(v) = \delta^+(v)$ for every $v \in V(D)$.*

Proof. $(\Rightarrow)$ Let D be an Eulerian digraph with no isolated vertices and consider a closed Eulerian trail W in D. Walking along W we can start from any vertex in D and reach any other vertex in D which shows that D is strongly, and hence also weakly connected. The trail W can be partitioned into oriented cycles C_1, $\ldots$, C_k in such a way that every edge in D belongs to exactly one of the cycles (Homework 6.1). Each vertex of D appears on W, so each vertex belongs to at least one of the cycles. Now, if $v \in V(D)$ lies on exactly l of these cycles, then $\delta^-(v) = l = \delta^+(v)$ since every edge in W belongs to precisely one of the cycles

$C_1, \ldots, C_k$, and each of the cycles "absorbs" one edge that goes into v and one edge that goes out of v.

($\Leftarrow$) Take any $v \in V(D)$. According to Lemma 6.3, v belongs to some closed oriented trail in D. Let W be the longest closed oriented trail in D that contains v and let us show that W is an Eulerian trail in D.

Suppose that W is not an Eulerian trail in D, i.e. $E(W) \subset E(D)$. If $V(W) = V(D)$, take any $e = (u,v) \in E(D) \setminus E(W)$. If $V(W) \subset V(D)$ then $\{V(W), V(D) \setminus V(W)\}$ is a partition of $V(D)$ and since D is weakly connected there is an edge $e = (u,v) \in E(D) \setminus E(W)$ such that $u \in V(W)$ and $v \in V(D) \setminus V(W)$ (or the other way around; the proof is analogous). In any case, let S be the weak connected component of $D - E(W)$ that contains e. Since W is a closed trail, it is easy to see that $\delta_S^-(v) = \delta_S^+(v)$ for every $v \in V(S)$. Hence, by Lemma 6.3 there exists a closed trail W' in S that contains u. Since $E(W') \subseteq E(S) \subseteq E(D) \setminus E(W)$, it follows that $E(W') \cap E(W) = \varnothing$, so glueing W and W' at u provides a trail that contains v and which is longer than W. Contradiction. □

The characterisation of Eulerian graphs is similar, and the proof goes along the same guidelines as in case of digraphs.

Theorem 6.5 *Let G be a graph with no isolated vertices. Then G is an Eulerian graph if and only if G is connected and each vertex of G is even.*

Proof. Analogous to the proof of Theorem 6.4. □

It is now easy to characterize noneulerian graphs that contain an Eulerian trail (which therefore cannot be a closed Eulerian trail).

Theorem 6.6 *Let G be a noneulerian graph with no isolated vertices. Then G has an Eulerian trail if and only if it is connected and has precisely two odd vertices.*

Proof. ($\Rightarrow$) Let W be an Eulerian trail in G. Since G is not Eulerian, W is not closed. Denote the vertices it starts and ends with by u and v. Introduce a new vertex $x \notin V(G)$ and two new edges $\{x,u\}$, $\{x,v\}$, and apply Theorem 6.5.

($\Leftarrow$) Let u and v be the odd vertices in G. Introduce a new vertex $x \notin V(G)$ and two new edges $\{x,u\}$, $\{x,v\}$, and apply Theorem 6.5. □

Finally, we conclude the section with another characterization of Eulerian graphs.

Theorem 6.7 *Let G be a connected graph. Then G is Eulerian if and only if every edge of G belongs to an odd number of cycles in G.*

Proof. We start by proving an auxiliary statement.

Claim. Let G be a connected noneulerian graph with an Eulerian trail and let u and v be the only two odd vertices in G. Then the number of trails that start at u, end in v and where v appears only once (i.e. at the end of the trail) is odd.

Proof. The proof is by induction on $m(G)$. The claim is true for connected noneulerian graphs with an Eulerian trail that have 1, 2 and 3 edges. Suppose the claim holds for all such graphs with $< m$ edges, and let G be such a graph with m edges. Furthermore, let u and v be the two odd vertices in G, let $k = \delta(u)$ and let $x_1, \ldots, x_k$ be the neighbours of u. For $j \in \{1,\ldots,k\}$ let $e_j = \{u, x_j\}$ and let T_j be the set of all the trails $u\, e_j\, x_j \ldots\ v$ with the property that v appears only at the end of the trail. Then $T_1 \cup \ldots \cup T_k$ is the set of all the trails we are considering and we have to show that $|T_1| + \ldots + |T_k|$ is odd. Since k is odd, it suffices to show that every $|T_j|$ is odd.

Take any $j \in \{1,\ldots,k\}$ and let $G_j = G - e_j$. The degree of u in G_j is even, so x_j and v are the only odd vertices in G_j. This is why they have to belong to the same connected component of G_j. The number of edges in this connected component is strictly less than m, so by the induction hypothesis the number of trails that start at x_j, end in v and contain v only once is odd. It is easily seen that the number of such trails equals $|T_j|$, and hence $|T_j|$ is also odd. This completes the proof of the claim.

Let us now go back to the proof of the theorem.

($\Leftarrow$) Let G be a connected graph that is not Eulerian. Then G has an odd vertex v. For an edge e incident to v let $c(e)$ denote the number of cycles in G that contain e. Since each such cycle contains two edges that are adjacent to v, the sum $\sum_{v \in e} c(e)$ is even (= twice the number of cycles that pass through v). But $\delta(v)$ is odd, so this sum consists of an odd number of summands. Therefore, one of the summands has to be even, and thus there exists an edge e adjacent to v such that $c(e)$ is even.

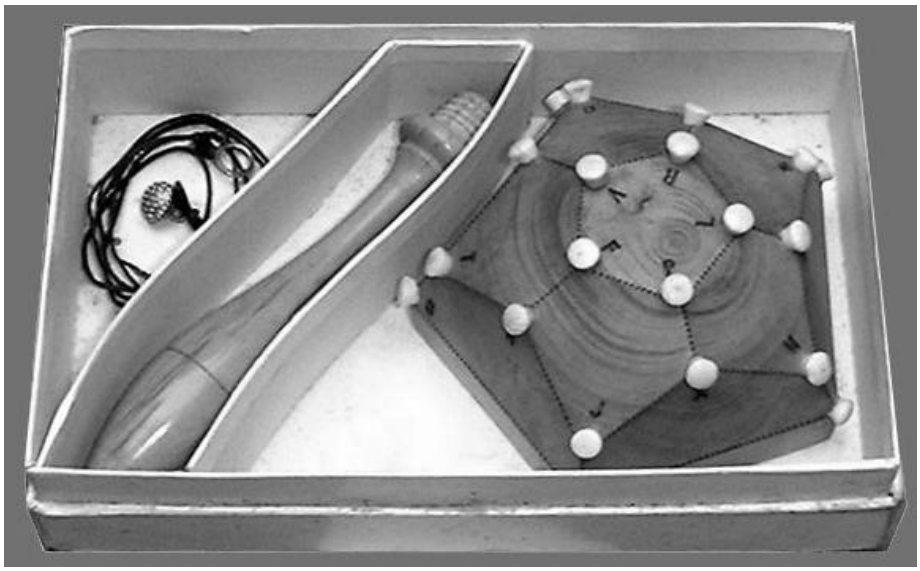
($\Rightarrow$) Let G be an Eulerian graph and let $e = \{u, v\} \in E(G)$ be arbitrary. According to Exercise 6.13, e is not a cut-edge, so $G - e$ is connected. Hence, $G - e$ is not Eulerian, but has an Eularian trail. Let this trail start at u and end in v. The Claim now yields that there is an odd number of trails that start at u, end in v and contain v only once. If S is one such trail which is not a path, then S contains some vertex more than once (for otherwise S would be a path). Let w_i be the first vertex in S that appears more than once in S and let $w_i e_{i+1} w_{i+1} \ldots e_j w_j = w_i$ be the shortest cycle in S that contains w_i. "Mirroring" the cycle within S produces a new trail S' having the same properties as S:

$$\begin{array}{rccccccccc} S: & u\, e_1\, w_1 \ldots w_i & e_{i+1} & w_{i+1} & \ldots & e_j & w_j \ldots w_{s-1} & e_s & v \\ & & \| & & & & \| & & \\ S': & u\, e_1\, w_1 \ldots w_j & e_j & \ldots & w_{i+1} & e_{i+1} & w_i \ldots w_{s-1} & e_s & v \end{array}$$

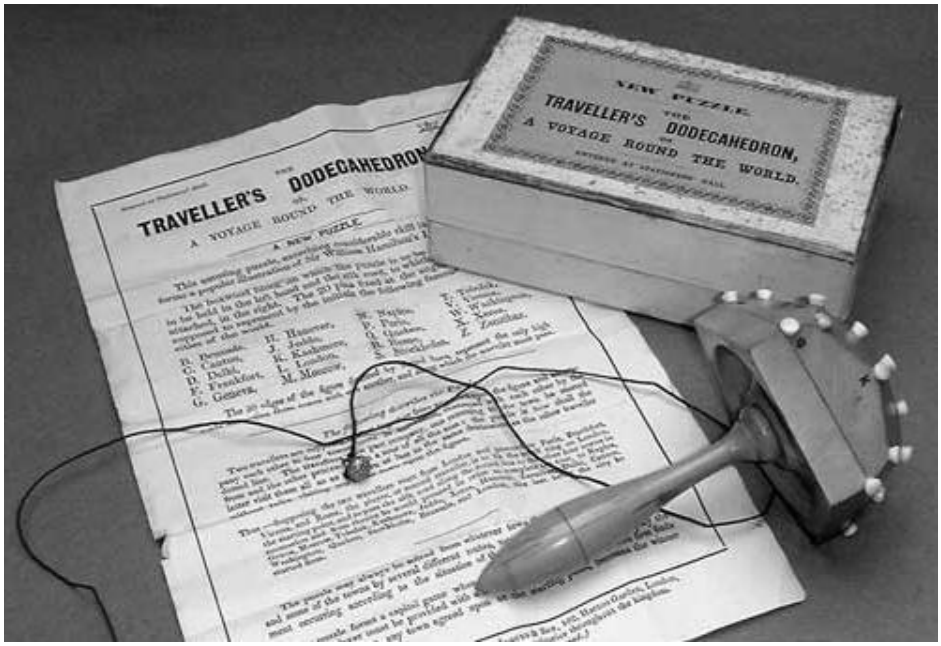
Therefore, trails that start at u, end in v, contain v only once and are not paths appear in pairs. Hence, the number of such trails which are not paths is even. But, we know that there is an odd number of trails with these properties, whence follows that the number of paths connecting u and v in $G - e$ is odd. Each of the paths together with e builds a cycle in G that contains e. Therefore, e belongs to an odd number of cycles. □

6.2 Hamiltonian graphs

Sir William Rowan Hamilton, who was Astronomer Royal of Ireland, invented in 1857 a puzzle called *The Travellers Dodecahedron or A Voyage Around the World*. It is not a true dodecahedron but is a "schematic" of a dodecahedron on a wooden "mushroom".



The 30 edges represent the only roads that one is allowed to pass along as one visits the 20 vertices that represent cities. Two travellers were supposed to set off visiting the cities: the first was supposed to pose a problem and start the tour by visiting four cities that belong to the same face of the dodecahedron. The player posing the problem then returns home and the other continues to travel around the world trying to visit all the remaining cities only once, and eventaully return home. The silk cord that accompanied the puzzle was used to mark the voyage and thus prevent the voyager from visiting a city more than once.



Until recently, only information we had on *The Travellers Dodecahedron* was its description in a chapter on Hamilton's Game in volume 2 of Édouard Lucas' *Récréations Mathématiques* and another mention in the 3rd edition of Ahrens' German work on Recreational Mathematics. But then an example was recovered, complete and in almost new condition.

In graph-theoretic terms the puzzle boils down to finding a spanning cycle of the incidence graph of a dodecahedron. The graph shown in Fig. 6.2 is a plane projection of a dodecahedron and we outlined a spanning cycle in this graph.

Definition 6.8 A *Hamiltonian path* in a graph is a path that contains all vertices of the graph. A *Hamiltonian cycle* in a graph is a cycle that contains all vertices of the graph. A graph is called *Hamiltonian* if it has a Hamiltonian cycle.

In comparison with Eulerian graphs, Hamiltonian graphs are much more hard to grasp. There is no "useful" characterisation of Hamiltonian graphs and we shall see in the next section that there is a justification for this: deciding whether a graph is Hamiltonian is one of the most complicated computational problems. We will actually show that this decision problem is NP-complete (for the moment, think of this as "extremely hard").

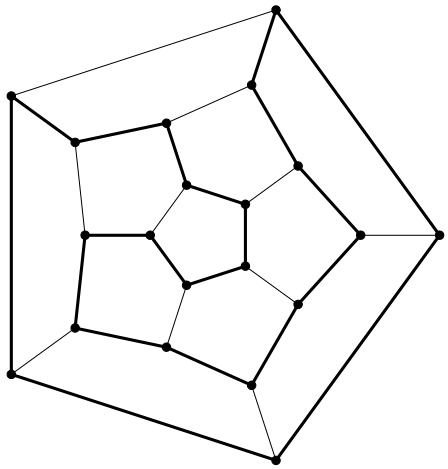


Figure 6.2: A solution to The Travellers Dodecahedron is a spaning cycle of the incidence graph of the dodecahedron

Theorem 6.9 *Let G be a Hamiltonian graph and $\varnothing \neq S \subset V(G)$ a nonempty set of vertices of G. Then $\omega(G-S) \leqslant |S|$.*

Proof. Let C be a Hamiltonian cycle of G. Then $\omega(C-S) \geqslant \omega(G-S)$ since $G-S$ has more edges than $C-S$, and they might connect some of the connected components of $C-S$ together. On the other hand, it is easy to see that $\omega(C-S) \leqslant |S|$. Therefore, $\omega(G-S) \leqslant |S|$. □

Theorem 6.9 is useful when it comes to showing that a graph is *not* Hamiltonian.

Corollary 6.10 *Hamiltonian graphs have no cut-vertices and no cut-edges.*

Proof. If v is a cut-vertex of a graph G then $\omega(G-v) \geqslant 2 > |\{v\}|$. Theorem 6.9 now implies that G is not Hamiltonian. We leave the cut-edges as Homework 6.5. □

We have already mentioned that there is no "useful" characterisation of Hamiltonian graphs. However, it is generally accepted that the best characterization of Hamiltonian graphs was given in 1972 by Bondy and Chvátal who generalized earlier results by G. A. Dirac and O. Ore. The idea behind their result is that a graph is Hamiltonian if enough edges exist.

If u,v are nonadjacent vertices in G and $e = \{u,v\}$, then by $G+e$ we denote the graph obtained by adding the edge e to G.

The closure of a graph G is a graph on the same set of vertices constructed as follows. Define a sequence of graphs $G_0, G_1, \ldots$, by $G_0 = G$ and

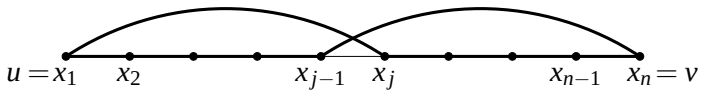
$$G_{i+1} = \begin{cases} G_i + e, & \text{where } e \notin E(G_i) \text{ joins two nonadjacent vertices} \\ & u, v \in V(G_i) \text{ such that } \delta_{G_i}(u) + \delta_{G_i}(v) \geqslant n(G_i), \\ G_i, & \text{if no such pair of vertices exists.} \end{cases}$$

Since we leave the set of vertices fixed and add new edges whenever possible, there exists a k such that $G_k = G_{k+j}$ for all $j \geqslant 1$. Then the graph G_k is called the *closure* of G and denoted by $\mathrm{cl}(G)$.

Theorem 6.11 (Bondy, Chvátal 1972) *A graph G is Hamiltonian if and only if $\mathrm{cl}(G)$ is Hamiltonian.*

Proof. If G is Hamiltonian, then so is $\mathrm{cl}(G)$ since $E(G) \subseteq E(\mathrm{cl}(G))$. For the converse, suppose that G is not Hamiltonian but that $\mathrm{cl}(G)$ is Hamiltonian. Then there exists a graph G_i in the sequence $G = G_0, G_1, \ldots, G_k = \mathrm{cl}(G)$ defining $\mathrm{cl}(G)$ such that G_i is not Hamiltonian and G_{i+1} is Hamiltonian. Let $G_{i+1} = G_i + e$ where $e = \{u, v\}$. Then by the construction, u and v are not adjacent and $\delta_{G_i}(u) + \delta_{G_i}(v) \geqslant n$.

Since $G_i + e$ is Hamiltonian and G_i is not, it follows that each Hamiltonian cycle in $G_i + e$ passes through e. Take any Hamiltonian cycle C in $G_i + e$. Then $e \in E(C)$ and hence $C - e$ is a Hamiltonian path $u = x_1\, x_2\, \ldots\, x_{n-1}\, x_n = v$ in G_i. Now it is easy to see that if u is adjacent to x_j for some $j > 1$ then v is *not* adjacent to x_{j-1} for otherwise we would have a Hamiltonian cycle in G_i:



Therefore, if $\delta_{G_i}(u) = k$ then $\delta_{G_i}(v) \leqslant n - (1 + k)$ since v is not adjacent to itself, nor is it adjacent to predecessors of the k neighbours of u. Hence $\delta_{G_i}(u) + \delta_{G_i}(v) \leqslant n - 1$. Contradiction. □

Corollary 6.12 *Let G be a graph with n vertices.*

(a) *If $\delta(u)+\delta(v) \geqslant n$ whenever u and v are distinct, nonadjacent vertices of G then G is Hamiltonian. (O. Ore 1960)*

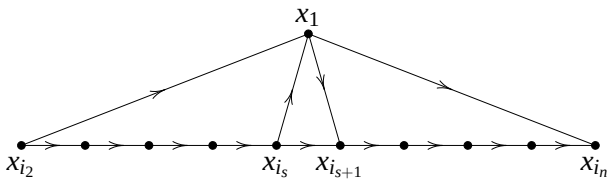
(b) *If $\delta(u) \geqslant \frac{n}{2}$ for all $u \in V(G)$ then G is Hamiltonian. (G. A. Dirac 1952)*

All these statements have their analogues for digraphs. We shall, however, treat only tournaments to show how very special digraphs they are.

Definition 6.13 A *Hamiltonian path* in a digraph is an oriented path that contains all vertices of the digraph. A *Hamiltonian cycle* in a digraph is an oriented cycle that contains all vertices of the digraph. A digraph is called *Hamiltonian* if it has a Hamiltonian cycle.

Theorem 6.14 (Rédei) *Every tournament has a Hamiltonian path.*

Proof. The proof is by induction on the number of vertices in the tournament. The statement is easily seen to be true in case of tournaments with 2 and 3 vertices. Assume now that every tournament with less than n vertices has a Hamiltonian path, and let T be a tournament on n vertices, $V(T)=\{x_1,\ldots,x_n\}$. By the induction hypothesis $T'=T-x_1$ has a Hamiltonian path $x_{i_2}\,x_{i_3}\,\ldots\,x_{i_n}$. If $x_1 \to x_{i_2}$ or $x_{i_n} \to x_1$, the Hamiltonian path of T' easily extends to a Hamiltonian path of T. If, however, $x_1 \not\to x_{i_2}$ and $x_{i_n} \not\to x_1$ then $x_{i_2} \to x_1$ and $x_1 \to x_{i_n}$. It is easy to see that there exists an s such that $x_{i_s} \to x_1 \to x_{i_{s+1}}$:



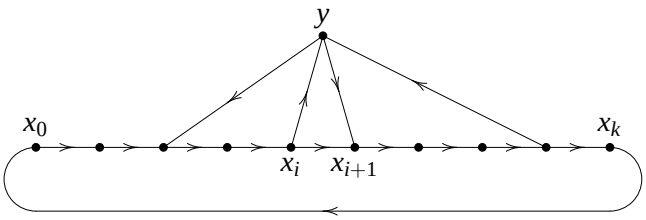
so $x_{i_2}\,\ldots\,x_{i_s}\,x_1\,x_{i_{s+1}}\,\ldots\,x_{i_n}$ is a Hamiltonian path for T. □

Theorem 6.15 *A tournament is Hamiltonian if and only if it is strongly connected.*

Proof. ($\Rightarrow$) If a tournament is Hamiltonian, then walking along the Hamiltonian cycle we can get from every vertex of the tournament to every other vertex. Hence, the tournament is strongly connected.

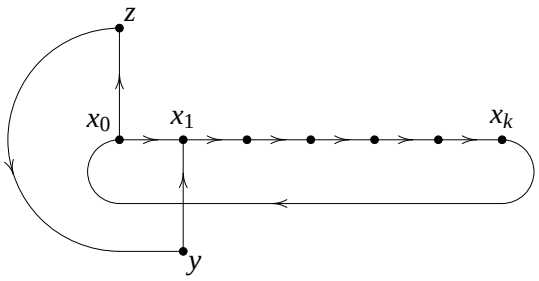
($\Leftarrow$) Let T be a strongly connected tournament. Then T is not transitive and hence contains an oriented cycle. Let $C = x_0 \to x_1 \to \ldots \to x_k \to x_0$ be the longest

oriented cycle in T and let us show that $V(C) = V(T)$. Suppose to the contrary that $V(C) \subset V(T)$. Then $\{V(C), B\}$ is a partition of $V(T)$, where $B = V(T) \setminus V(C)$. If there exists a $y \in B$ such that $E(V(C), \{y\}) \neq \varnothing$ and $E(\{y\}, V(C)) \neq \varnothing$ then there exists an index i such that $x_i \to y \to x_{i+1}$:



and $x_0 \to \ldots \to x_i \to y \to x_{i+1} \to \ldots \to x_k \to x_0$ is an oriented cycle in T which is longer than C. Contradiction.

Therefore, for each $y \in B$ either $E(V(C), \{y\}) = \varnothing$ or $E(\{y\}, V(C)) = \varnothing$. Let $Y = \{y \in B : E(V(C), \{y\}) = \varnothing\}$ and $Z = \{z \in B : E(\{z\}, V(C)) = \varnothing\}$. Since T is strongly connected it follows that $Y \neq \varnothing$, $Z \neq \varnothing$ and $E(Z, Y) \neq 0$. Take $z \in Z$ and $y \in Y$ such that $z \to y$. From $E(V(C), \{y\}) = \varnothing$ it follows that $y \to x_i$ for all i.



Similarly, $x_i \to z$ for all i, so $x_0 \to z \to y \to x_1 \to \ldots \to x_k \to x_0$ is an oriented cycle in T and it is longer than C. Contradiction. Therefore, $V(C) = V(T)$, so T is a Hamiltonian tournament. □

A careful analysis of the previous proof reveals that we can actually prove much more.

Theorem 6.16 (Camion 1959) *Let T be a Hamiltonian tournament with n vertices. For every vertex $v \in V(T)$ and every $k \in \{3, \ldots, n\}$ there exists an oriented cycle of length k that contains v.*

6.3 Complexity issues

In this section we consider the computational complexity of deciding whether a graph has a Hamiltonian cycle. We show that this decision problem not only falls into the NP complexity class, but that it is an NP-complete problem, i.e. a paradigm of an NP-hard problem.

The notion of an algorithm (= "effective procedure") was recognised as one of the essential notions in mathematics as early as 1928 when D. Hilbert and W. Ackermann published their influential booklet "Grundzüge der theoretischen Logik" in which they posed a problem of finding an algorithm (whatever that might mean) which decides whether a first-order sentence is a consequence of the axioms of arithmetic. At that time there was no formal notion of an algorithm, so the problem was actually twofold: on the "philosophical" level it was required to introduce the precise definition of an algorithm, while on the mathematical level the definition should have been used in solving the particular problem of mathematical logic. The problem (both on the philosophical and the mathematical level) was independently solved in 1936 by A. Church and A. Turing. Although Church's solution was published a few months ahead of Turing's, the approach taken by A. Turing is more intuitive, and constitues a basis of what is today known as Computability Theory.

We shall not present a formal definition of a Turing machine. For our purposes it suffices to say that a *Turing machine* is a mathematical model of a computer program written for a modern computer with infinite memory. Since computers actually operate on finite 01-words we shall take $\Sigma = \{0,1\}$ as the alphabet in which to carry out our considerations. Let Σ^* denote the set of all finite 01-words, together with the empty word ε. By $|w|$ we denote the length of $w \in \Sigma^*$. A *language* is any set $\mathscr{L} \subseteq \Sigma^*$ of 01-words. In particular, for every graph G there is a 01-word $\langle G \rangle$ representing the graph, so we also have the language $\mathscr{G} = \{\langle G \rangle : G \text{ is a graph}\}$.

A computer program A can take any 01-word w as its input, but may fail to produce an output. Hence, each computer program A corresponds to a function $\widehat{A} : \Sigma^* \to \Sigma^* \cup \{\infty\}$ such that

$$\widehat{A}(w) = \begin{cases} u, & A \text{ takes } w \text{ as its input and after a finite number of computation} \\ & \text{steps stops and prints } u \text{ as a result;} \\ \infty, & A \text{ never stops on input } w. \end{cases}$$

For a computer program A and a word $w \in \Sigma^*$ let

$$t_A(w) = \begin{cases} n, & A \text{ takes } w \text{ as its input and stops after } n \text{ computation steps;} \\ \infty, & A \text{ never stops on input } w. \end{cases}$$

A computer program A *runs in polynomial time* if there exists a positive integer k such that $t_A(w) = O(|w|^k)$ whenever $\widehat{A}(w) \neq \infty$.

The complexity class P. A language $\mathscr{L} \subseteq \Sigma^*$ is *decidable* if there exists a computer program A such that $\widehat{A} : \Sigma^* \to \{0,1\}$ and

$$\mathscr{L} = \{w \in \Sigma^* : \widehat{A}(w) = 1\}.$$

(Note that the computer program which decides a language stops on all inputs and outputs 0 or 1.) The language $\mathscr{L} \subseteq \Sigma^*$ is *decidable in polynomial time* if there exists a computer program A wich runs in polynomial time such that $\widehat{A} : \Sigma^* \to \{0,1\}$ and $\mathscr{L} = \{w \in \Sigma^* : \widehat{A}(w) = 1\}$.

Definition 6.17 The complexity class P consists of all languages over $\Sigma = \{0,1\}$ that are decidable in polynomial time:

$$\mathsf{P} = \{\mathscr{L} \subseteq \Sigma^* : \mathscr{L} \text{ is decidable in polynomial time}\}.$$

Equivalently, the complexity class P consists of all problems that can be solved in polynomial time. Indeed, given an problem Q it suffices to encode each instance I of the problem by a 01-word $\langle I \rangle$ and consider the language $\mathscr{L}_Q = \{\langle I \rangle : I \text{ is an instance of } Q\}$. Then each instance I of the problem can be solved in polynomial time (where the degree of the polynomial does not depend on the instance) if and only if $\mathscr{L}_Q$ is decidable in polynomial time. For example, the problem of deciding in polynomial time whether a graph is connected corresponds to polynomial decidability of the language $\mathscr{L}_{conn} = \{\langle G \rangle : G \text{ is a connected graph}\}$. For some other problems the transformation $\boxed{\text{problem}} \to \boxed{\text{language}}$ may not be so obvious.

The complexity class NP. Instead of requiring a computer program to solve a problem, we might only wish to pull a solution out of a sleeve and verify that then solution is indeed a solution to a problem. A *verification algorithm* is a computer program A with two inputs such that $\widehat{A} : \Sigma^* \times \Sigma^* \to \{0,1\}$. If there exists a positive integer k such that $t_A(p,s) = O((|p| + |s|)^k)$ for all $p, s \in \Sigma^*$ we say that A *is a polynomial verification algorithm.* A language $\mathscr{L}$ is *verified* by a verification algorithm A if

$$\mathscr{L} = \{p \in \Sigma^* : \exists s \in \Sigma^* \ (\widehat{A}(p,s) = 1)\}.$$

A *language* $\mathscr{L} \subseteq \Sigma^*$ *is verifiable in polynomial time* if there exists a positive integer c and a polynomial verification algorithm A such that

$$\mathscr{L} = \{p \in \Sigma^* : \exists s \in \Sigma^* \ (|s| \leqslant |p|^c \text{ and } \widehat{A}(p,s) = 1)\}.$$

Definition 6.18 The complexity class NP consists of all languages over $\Sigma = \{0,1\}$ that are verifiable in polynomial time:

$$\mathsf{NP} = \{\mathscr{L} \subseteq \Sigma^* : \mathscr{L} \text{ is verifiable in polynomial time}\}.$$

Equivalently, the complexity class NP consists of problems for which it is easy to check whether what we claim to be a solution is indeed a solution. For example, $\mathscr{L}_{Ham} = \{\langle G\rangle : G \text{ is a Hamiltonian graph}\}$ is in NP since given a graph G and a sequence of vertices $x_1,\ldots,x_n$ it is easy to check whether $x_1,\ldots,x_n$ is a Hamiltonian cycle of G.

Theorem 6.19 $\mathsf{P} \subseteq \mathsf{NP}$.

Proof. Take any $\mathscr{L} \in \mathsf{P}$. Then $\mathscr{L} = \{w \in \Sigma^* : \widehat{A}(w) = 1\}$ for some computer program A that decides $\mathscr{L}$ in polynomial time. Now take a verification algorithm $B : \Sigma^* \times \Sigma^* \to \{0,1\}$ so that $\widehat{B}(p,s) = \widehat{A}(p)$. Then B clearly verifies $\mathscr{L}$ in polynomial time, so $\mathscr{L} \in \mathsf{NP}$. □

The exact relationship between P and NP is still unknown. It is strongly believed that $\mathsf{P} \neq \mathsf{NP}$, but we still haven't got a proof. The problem is actually so important that the Clay Mathematics Institute is offering a USD 1,000,000 prize for the correct solution.[1] Apart from the prize, the importance of the problem is also reflected by the fact that the security of RSA, the most widely used crypto-system, depends on $\mathsf{P} \neq \mathsf{NP}$. If it turns out that $\mathsf{P} = \mathsf{NP}$ the security of all transactions based on RSA, PGP and the such will be broken and many aspects of our everyday life would have to change.

Polynomial reducibility and NP-completeness. We say that a language $\mathscr{L}_1 \subseteq \Sigma^*$ is *polynomially reducible* to a language $\mathscr{L}_2 \subseteq \Sigma^*$ and write and write $\mathscr{L}_1 \preccurlyeq_p \mathscr{L}_2$ if there exists a computer program A which runs in polynomial time such that $\widehat{A} : \Sigma^* \to \Sigma^*$ and

$$w \in \mathscr{L}_1 \text{ if and only if } \widehat{A}(w) \in \mathscr{L}_2.$$

Intuitively, regarding polynomial-time as "easy", this means: if there is a polynomial reduction from $\mathscr{L}_1$ to $\mathscr{L}_2$, then $\mathscr{L}_1$ cannot be harder than $\mathscr{L}_2$.

Theorem 6.20 *If* $\mathscr{L} \in \mathsf{P}$ *and* $\mathscr{L}' \preccurlyeq_p \mathscr{L}$ *then* $\mathscr{L}' \in \mathsf{P}$.

[1] http://www.claymath.org/millennium/P_vs_NP/

Proof. If A is a computer program that decides $\mathscr{L}$ in polynomial time, and if B is a computer program that reduces $\mathscr{L}'$ to $\mathscr{L}$ in polynomial time, then $B \circ A$ is a computer program that decides $\mathscr{L}'$ in polynomial time, so $\mathscr{L}' \in \mathsf{P}$. □

Definition 6.21 A language $\mathscr{L} \subseteq \Sigma^*$ is NP-*hard* if $\mathscr{L}' \preccurlyeq_p \mathscr{L}$ for every $\mathscr{L}' \in \mathsf{NP}$. A language $\mathscr{L} \subseteq \Sigma^*$ is NP-*complete* if it is NP-hard and belongs to NP.

An NP-complete problem is a paradigm of an NP-problem. Moreover, if one of them happens to be in P then $\mathsf{P} = \mathsf{NP}$:

Theorem 6.22 *Let $\mathscr{L}$ be an* NP*-complete language. If $\mathscr{L} \in \mathsf{P}$ then* $\mathsf{P} = \mathsf{NP}$.

Proof. Suppose that $\mathscr{L}$ is an NP-complete language such that $\mathscr{L} \in \mathsf{P}$. Take any $\mathscr{L}' \in \mathsf{NP}$. Since $\mathscr{L}$ is NP-hard, it follows that $\mathscr{L}' \preccurlyeq_p \mathscr{L}$ and thus $\mathscr{L}' \in \mathsf{P}$ by Theorem 6.20. This shows that $\mathsf{NP} \subseteq \mathsf{P}$. □

The first hands-on NP-complete problem was discovered in 1971 by S. Cook. A *Boolean formula* is a formula built up from Boolean variables $x_1, \ldots, x_n$ (each of which can take the values *true* or *false*) and Boolean connectives $\neg$, $\wedge$ and $\vee$. A Boolean formula $F(x_1, \ldots, x_n)$ is said to be in a *conjunctive form* (CF for short) if it has the form

$$F(x_1, \ldots, x_n) = C_1(x_1, \ldots, x_n) \wedge C_2(x_1, \ldots, x_n) \wedge \ldots \wedge C_k(x_1, \ldots, x_n)$$

where each clause $C_i(x_1, \ldots, x_n)$ is a disjuction of literals

$$C_i(x_1, \ldots, x_n) = (l_{i1} \vee l_{i2} \vee \ldots \vee l_{im_i})$$

and each literal l_{ij} is a variable x_{ij} or a negated variable $\neg x_{ij}$. It is a well known fact from Boolean logic that every Boolean formula is equivalent to a CF Boolean formula.

A Boolean formula $F(x_1, \ldots, x_n)$ is *satisfiable* if there exists an assignment $\tau : \{x_1, \ldots, x_n\} \to \{true, false\}$ of truth values to variables such that $\tau(F) = true$, that is, F evaluates to *true* under the assignment τ. Let us fix a systematic way of encoding CF Boolean formulas by 01-words and let $\langle F \rangle$ denote an encoding of F. Let us denote the language that corresponds to satisfiable Boolean formulas by *SAT*:

$$SAT = \{\langle F \rangle : F \text{ is a satisfiable CF Boolean formula}\}.$$

Theorem 6.23 (Cook 1971) *SAT is* NP*-complete.*

Now that we have an explicite NP-complete problem, it gives us a strategy to show that other problems are also NP-complete: if an NP-complete problem is polynomially reducible to some other problem, this new problem also has to be NP-complete.

Theorem 6.24 *If $\mathscr{L}$ is an* NP*-complete language and if $\mathscr{L}' \in$ NP has the property that $\mathscr{L} \preccurlyeq_p \mathscr{L}'$ then $\mathscr{L}'$ is also* NP*-complete.*

Proof. This is an immediate consequence of the fact that $\preccurlyeq_p$ is transitive. □

Therefore, in order to show that finding a Hamiltonian cycle in a graph is an NP-complete problem, it suffices to show that *SAT* is polynomially reducible to it. In this particular case, working with digraphs turns out to be easier than working with graphs, so we introduce the two languages:

- $HAMG = \{\langle G\rangle : G \text{ is a Hamiltonian graph}\}$, which is a 01-language that encodes Hamiltonian graphs, and
- $HAMD = \{\langle D\rangle : D \text{ is a Hamiltonian digraph}\}$, which is a 01-language that encodes Hamiltonian digraphs,

and carry out the proof in two steps:

- we first show that $HAMG \preccurlyeq_p HAMD$ and $HAMD \preccurlyeq_p HAMG$; and then
- we show that $SAT \preccurlyeq_p HAMD$.

Lemma 6.25 *$HAMG \preccurlyeq_p HAMD$ and $HAMD \preccurlyeq_p HAMG$.*

Proof. For every graph $G = (V,E)$ let $D_G = (V,E')$ denote the digraph with the same set of vertices whose set of edges is

$$E' = \{(u,v) \in V^2 : \{u,v\} \in E\}.$$

Clearly, there exists a polynomial algorithm that converts $\langle G\rangle$ to $\langle D_G\rangle$ and it is easy to see that G is a Hamiltonian graph if and only if D_G is a Hamiltonian digraph (Homework 6.11). Therefore, $HAMG \preccurlyeq_p HAMD$.

Now, let $D = (V,E)$ be a digraph and let $G_D = (V',E')$ be a graph constructed from G as follows. For each $v \in V$ we add three vertices v^0, v^1, v^2 to V' and two edges $\{v^0, v^1\}$ and $\{v^1, v^2\}$ to E' replacing thus each vertex of D by a path of length 2 in G_D. Moreover, for each edge (u,v) in E we add an edge $\{u^2, v^0\}$ to E'. An illustration of this proces is given in Fig. 6.3. Clearly, $|V'| = 3|V|$ and $|E'| = |E| + 2|V|$, so the reduction is polynomial. It is also easy to see that D is a

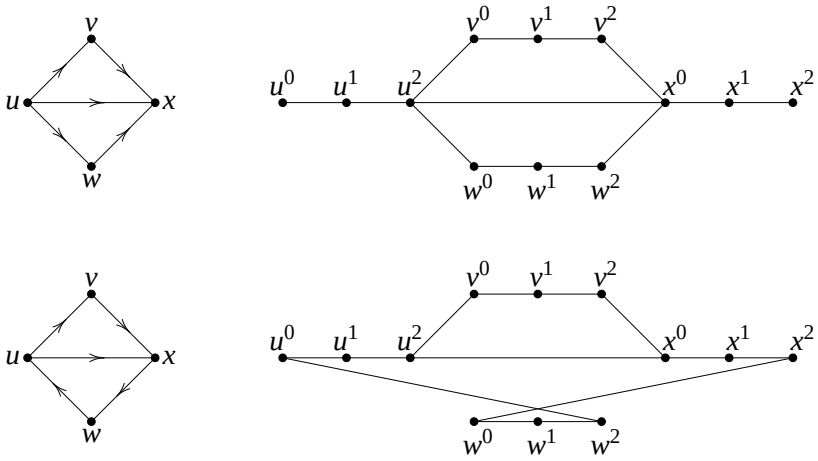


Figure 6.3: Two digraphs and their associated graphs

Hamiltonian digraph if and only if G_D is a Hamiltonian graph (Homework 6.11). Therefore, $HAMD \preccurlyeq_p HAMG$. □

Theorem 6.26 *HAMG is* NP*-complete.*

Proof. According to Theorem 6.24 it suffices to show that $SAT \preccurlyeq_p HAMG$. We shall actually show that $SAT \preccurlyeq_p HAMD$ and then use $HAMD \preccurlyeq_p HAMG$ established in Lemma 6.25. Therefore, for every Boolean formula $F(x_1,\ldots,x_n)$ in CF we have to construct a not too complicated digraph D_F such that F is satisfiable if and only if D_F has an oriented Hamiltonian cycle.

Let $F(x_1,\ldots,x_n)$ be a Boolean formula given in its conjunctive form:

$$F(x_1,\ldots,x_n) = C_1(x_1,\ldots,x_n) \wedge C_2(x_1,\ldots,x_n) \wedge \ldots \wedge C_k(x_1,\ldots,x_n).$$

Recall that each clause $C_i(x_1,\ldots,x_n)$ is a disjuction of literals

$$C_i(x_1,\ldots,x_n) = (l_{i1} \vee l_{i2} \vee \ldots \vee l_{im_i})$$

and each literal l_{ij} is a variable x_{ij} or a negated variable $\neg x_{ij}$. We construct a digraph D_F with $2nk+k$ vertices as follows. For each variable x_i we have $2k$ vertices u_{i1}, v_{i1}, u_{i2}, v_{i2}, …, u_{ik}, v_{ik}, and for each clause C_i we have a vertex c_i. The vertices u_{ij}, v_{ij} are connected by edges as in Fig 6.4 (a). We choose a direction, say from left to right, and say that that x_i evaluates to *true* if we traverse vertices that correspond to x_i in that direction, while it evaluates to false if we traverse the vertices that correspond to x_i in the oposite direction.

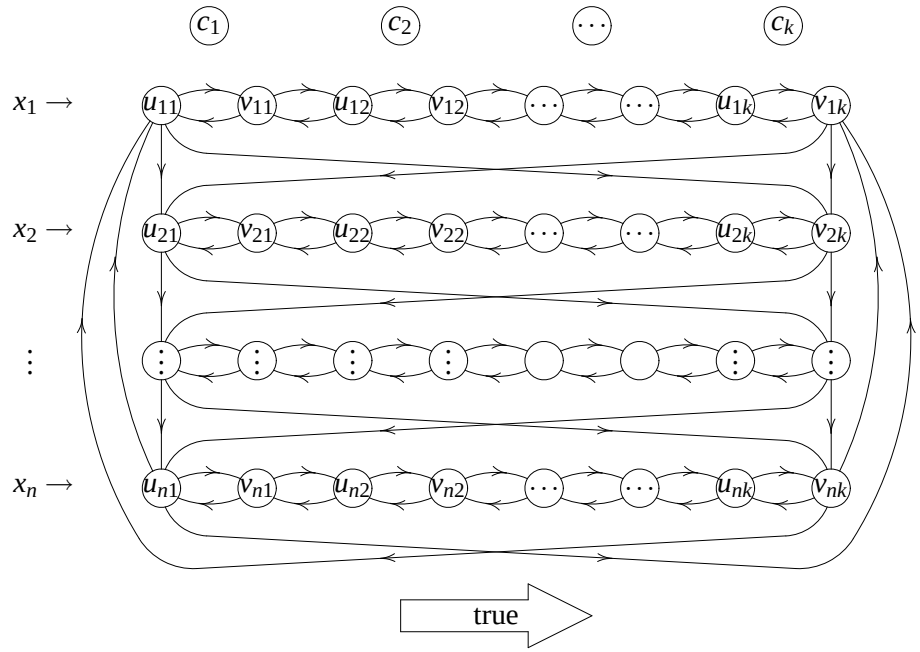


Figure 6.4: The construction of the digraph D_F, Part I: vertices that correspond to variables

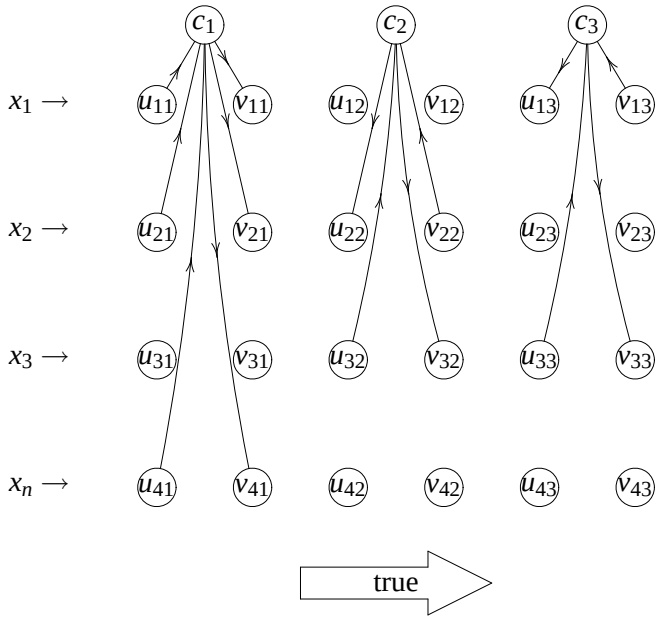


Figure 6.5: The construction of the digraph D_F, Part II: vertices that correspond to clauses

Next, we describe how to connect vertices that correspont to clauses to vertices that correspond to variables. If a variable x_i appears in a clause C_j and it is not negated in C_j, we add the edges $u_{ij} \rightarrow c_j$ and $c_j \rightarrow v_{ij}$. If, however, x_i is negated in C_j we add the edges $v_{ij} \rightarrow c_j$ and $c_j \rightarrow u_{ij}$. So, if a variable x_i is not negated in a clause C_j we add edges that go "in the direction of truth". If x_i is negated in C_j, we add edges that go "in the direction oposite of truth". An example is given in Fig. 6.5 (for clarity, the figure indicates only the edges incident to vertices that represent clauses; edges connecting u_{ij}'s to v_{ij}'s have been omitted). The digraph in Fig. 6.5 corresponds to the boolean formula $F(x_1,x_2,x_3,x_4) = C_1 \wedge C_2 \wedge C_3$ where $C_1 = x_1 \vee x_2 \vee x_4$, $C_2 = \neg x_2 \vee x_3$ and $C_3 = \neg x_1 \vee x_3$. The full graph that represents F is given in Fig. 6.6.

It is easy to see that this construction can be carried out in polynomial time. Let us finally show that F is satisfiable if and only if D_F has an oriented Hamiltonian cycle. Recall that traversing a row of vertices that corresponds to x_i from left to right means $\tau(x_i) = \mathit{true}$ while traversing from right to left means $\tau(x_i) = \mathit{false}$. The idea is that an oriented Hamiltonian cycle through the digraph represents an assignment of truth values to the variables $x_1, \ldots, x_n$.

Assume the formula F is satisfiable by some truth assignment τ. Choose one true literal in each clause, traverse the graph moving across each variableŠs path

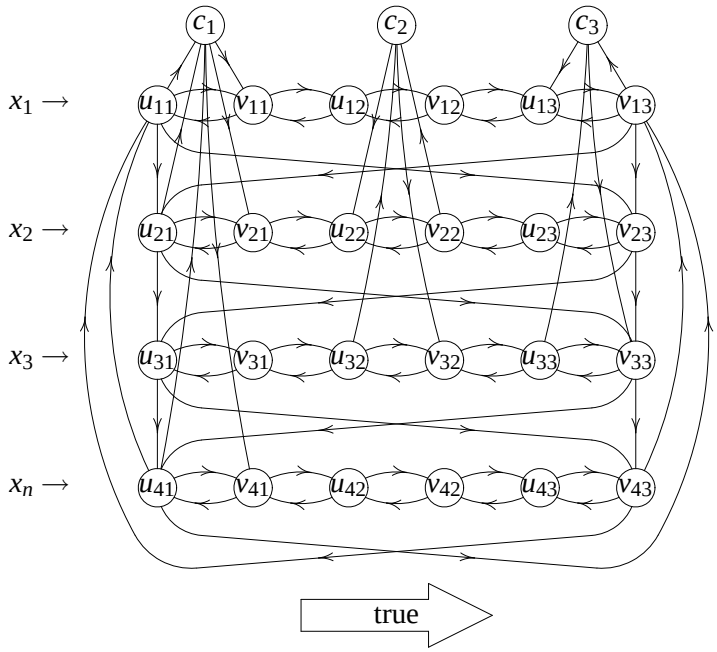


Figure 6.6: The digraph D_F for $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_4) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_1 \vee x_3)$

in the appropriate direction, and take a diversion to a clause-node for each literal chosen above. This oriented path is a Hamiltonian cycle.

Conversely, suppose there exists an oriented Hamiltonian cycle H in D_F. Then H traverses each variableŠs row either from left to right or from right to left and thus determines an assignment of truth values τ to variables. Each clause-node is visited by a side-trip from a variable row. This variable corresponds to a true literal in the clause. Hence, each clause evaluates to *true* under τ and hence $\tau(F) = true$, i.e. F is a satisfiable formula. □

Homework

6.1. Let D be an Eulerian digraph. Prove that each closed Eulerian walk in D can be partitioned into oriented cycles in such a way that every edge of D belongs to exactly one of the cycles. (Hint: use induction on the length of the walk.)

6.2. Prove Theorem 6.5.

6.3. Complete the proof of Theorem 6.6.

6.4. There are five regular polyhedra: tetrahedron, hexahedron, octahedron, dodecahedron and icosahedron (Fig. 6.7). Which of them could have been used instead of the dodecahedron in the Hamilton's Voyage Around the World puzzle?

6.5. Complete the proof of Corollary 6.10.

6.6. Prove Corollary 6.12. (Hint: for (a) show that $\mathrm{cl}(G)$ is a complete graph and use the Bondy-Chvátal Theorem; (b) follows from (a).)

6.7. (Ore 1960) Let G be a graph with n vertices. If $\delta(u) + \delta(v) \geqslant n-1$ whenever u and v are distinct, nonadjacent vertices of G then G has a Hamiltonian path. (Hint: add a new vertex to G and connect it by an edge to every vertex of G; show that the new graph is Hamiltonian using a similar result for Hamiltonian graphs.)

6.8. Show that a transitive tournament has exactly one Hamiltonian path.

6.9. Show that each tournament which is not strongly connected can be turned into a strongly connected tournament by changing the orientation of only one edge.

6.10. Prove Theorem 6.16. (Hint: induction on k using the fact that a Hamiltonian tournament is strongly connected; for $k = 3$ show that $E(O(v), I(v)) \neq$

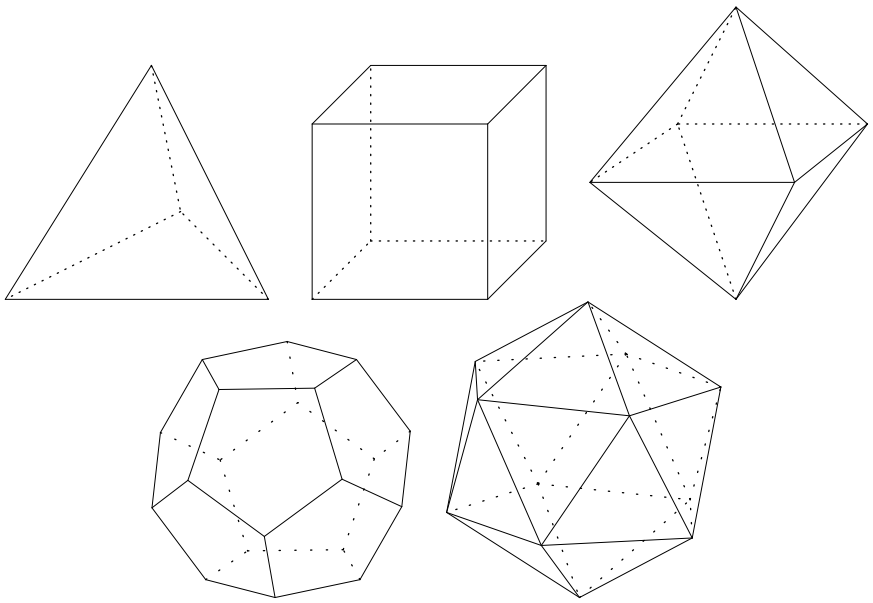


Figure 6.7: The five regular polyhedra

$\varnothing$; for the induction step modify slightly the idea used in the proof of Theorem 6.15.)

6.11. Complete the proof of Lemma 6.25 by showing that

- G is a Hamiltonian graph if and only if D_G is a Hamiltonian digraph; and
- D is a Hamiltonian digraph if and only if G_D is a Hamiltonian graph.

Exercises

6.12. (a) For each $n \geqslant 2$ give an example of a graph with n vertices which is neither Eulerian nor Hamiltonian.

(b) For each $n \geqslant 3$ give an example of a graph with n vertices which is both Eulerian and Hamiltonian.

(c) For each $n \geqslant 4$ give an example of a Hamiltonian graph with n vertices which is not Eulerian.

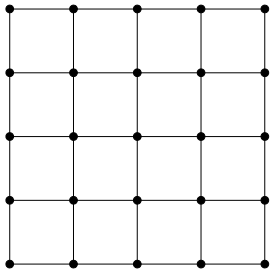


Figure 6.8: Exercise 6.19

(d) For each $n \geqslant 5$ give an example of an Eulerian graph with n vertices which is not Hamiltonian.

6.13. Prove that an Eulerian graph with no isolated vertices has no cut-edges.

6.14. For a digraph D and a set of edges $F \subseteq E(D)$ let W be the set of all vertices of D incident to an edge in F and let $D[F] = (W, F)$ denote the *subdigraph of D induced by F*.

Let D be a weakly connected digraph. Prove that D is Eulerian if and only if there exists a partition $\{F_1, \ldots, F_k\}$ of $E(D)$ such that each $D[F_i]$ is an oriented cycle.

6.15. Let A be a finite set with at least three elements. On $V = \mathscr{P}(A) \setminus \{\varnothing, A\}$ as a set of vertices we define a graph G as follows: two proper subsets X and Y of A are adjacent if and only if $X \subset Y$ or $Y \subset X$ (i.e., if and only if one of them is a proper subset of the other one). Show that G is an Eulerian graph.

6.16. Let G be an Eulerian graph with no isolated vertices and with $n(G)$ odd. If $\Delta(G) \leqslant \lfloor \frac{n}{2} \rfloor$ show that $\overline{G}$ is an Eulerian graph.

6.17. Let G ne an Eulerian graph with no isolated vertices and with $n(G)$ odd. If $d(G) \geqslant 3$ show that $\overline{G}$ is an Eulerian graph.

6.18. Let G be a connected graph with $2k$ odd vertices. Show that $E(G)$ can be partitioned into k edge-disjoint trails.

6.19. Is it possible to partition the edge-set of the graph in Fig. 6.8 into five edge-disjoint paths of legth 8?

6.20. Which of the graphs in Fig. 6.9 are Hamiltonian?

6.21. (a) Let G be a bipartite Hamiltonian graph and let $\{X, Y\}$ be a partition

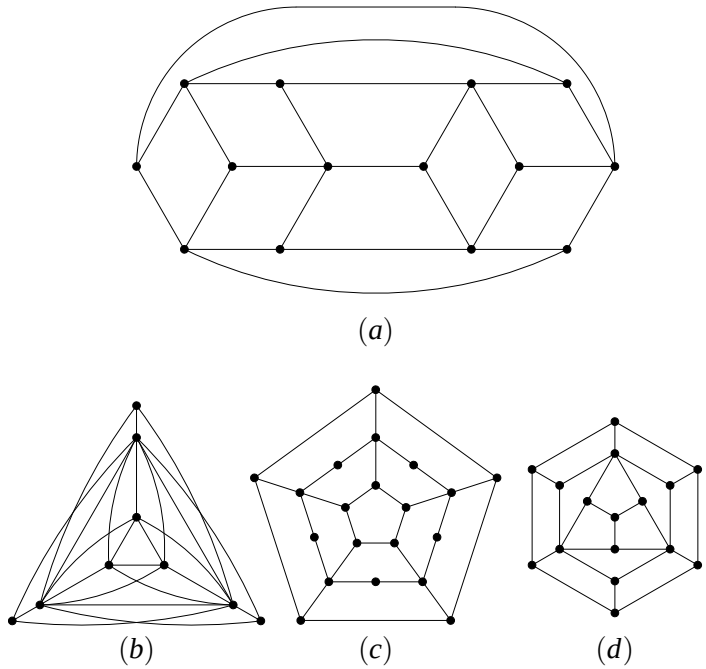


Figure 6.9: Exercise 6.20

of the set of its vertices that demonstrates that G is bipartite. Show that $|X| = |Y|$.

(b) Is the graph in Fig. 6.10 Hamiltonian?

6.22. A vertex cover of a graph G is a set of vertices $W \subseteq V(G)$ such that every edge in G is incident to a vertex from W. Show that if G has a vertex cover W such that $|W| < \frac{1}{2}n(G)$ then G is not Hamiltonian.

6.23. Let G be a graph with n vertices and m edges such that $m \geqslant \binom{n-1}{2} + 2$. Show that G is a Hamiltonian graph.

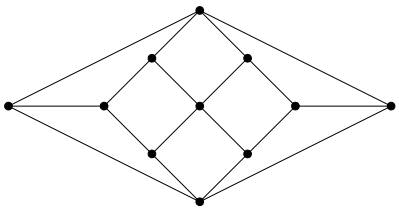


Figure 6.10: Exercise 6.21

6.24. Show that the complement of a regular disconnected graph is a Hamiltonian graph.

6.25. Show that a hypercube of dimension $k \geqslant 2$ is a Hamiltonian graph.

6.26. Show that every strongly connected tournament with $n \geqslant 4$ vertices contains a vertex v such that after changing the orientation of all the edges incident to v we again obtain a strongly connected tournament.

6.27. Show that a strongly oriented tournament with $n \geqslant 3$ vertices has at least $n-2$ oriented triangles. (An oriented triange is an oriented cycle of length 3.)

†**6.28.** Let $s_1 \leqslant s_2 \leqslant \ldots \leqslant s_n$ be the scores in a tournament T with n vertices. If $s_n - s_1 < \frac{n}{2}$, show that T is a Hamiltonian tournament. (Hint: show that $s_j - s_i < \frac{n}{2}$ whenever $i < j$ and conclude that T is strongly connected.)

Chapter 7

Planar graphs

The main point about graphs is that we can draw them. Speaking about graphs as abstract objects without the appropriate accompanying drawing is unheard of. Because drawing graphs is so important, graphs with particularly nice drawings are particularly important.

Planar graphs are graphs that can be drawn in the Euclidean plane in such a way that no two edges have a common internal point. Planarity is therefore introduced as a geometric concept. However, one of the deepest results in graph theory tells us that there is a combinatorial characterisation of planar graphs, showing that, although introduced as a geometric concept, planarity is a combinatorial property of graphs.

In this chapter we first introduce planar graphs as graphs which have "nice" drawings and then present the results of Kuratowski (1930) and Wagner (1937) that characterise planar graphs in a purely combinatorial fashion. At the end of the chapter we discuss regular polyhedra. Euler's proof that there are only five regular polyhedra is in the heart of Euler's formula for planar graphs.

7.1 Planarity as a geometric concept

An *arc* (or a *Jordan curve*) in $\mathbb{R}^2$ is any injective continuous mapping $\gamma : [0,1] \to \mathbb{R}^2$, where the real interval $[0,1]$ and $\mathbb{R}^2$ are endowed with the usual topologies. Points $\gamma(0)$ and $\gamma(1)$ are called the *end-points of* γ. Arcs $\gamma, \gamma' : [0,1] \to \mathbb{R}^2$ are *internally disjoint* if $\{\gamma(x) : 0 < x < 1\} \cap \{\gamma'(x) : 0 < x < 1\} = \varnothing$, Fig. 7.1. Let $\mathrm{Arc}(\mathbb{R}^2)$ denote the set of all arcs in $\mathbb{R}^2$. A *drawing of a graph* $G = (V,E)$ is a pair of mappings (ν, ε) where $\nu : V \to \mathbb{R}^2$ is injective, $\varepsilon : E \to \mathrm{Arc}(\mathbb{R}^2)$ is injective, and the following compatibility requirement is satisfied (see Fig 7.2):

$$\text{if } e = \{u,v\} \text{ and } \varepsilon(e) = \gamma \text{ then } \{\gamma(0), \gamma(1)\} = \{\nu(u), \nu(v)\}.$$

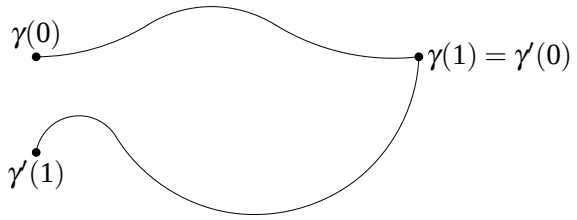


Figure 7.1: Two internally disjoint arcs

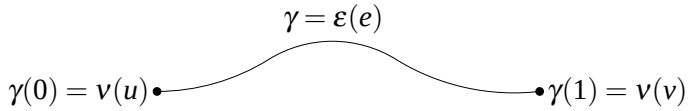


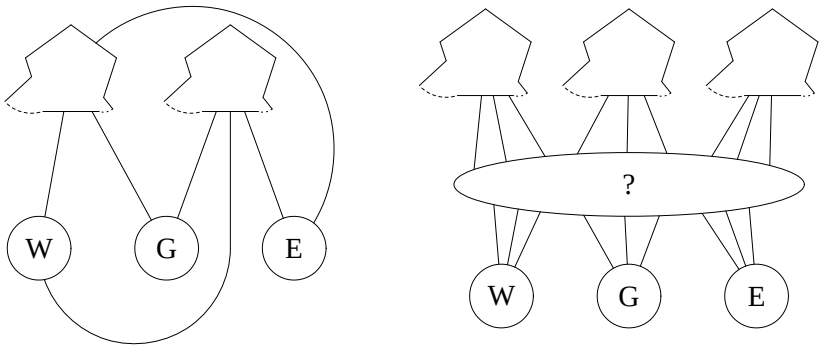
Figure 7.2: The compatibility of ε and v

A *planar representation of a graph G* is a drawing (v, ε) of G such that $\varepsilon(e_1)$ and $\varepsilon(e_2)$ are internally disjoint arcs whenever e_1 and e_2 are distinct edges of G. A graph G is *planar* if there is a planar representation of G.

Example 7.1 In 1884 Edwin A. Abbott wrote a fascinating novel *Flatland: A romance of many dimensions* in which two-dimensional beings live in a two-dimensional universe. Here is an excerpt from the introduction:

> "I call our world Flatland, not because we call it so, but to make its nature clearer to you, my happy readers, who are privileged to live in Space.
>
> Imagine a vast sheet of paper on which straight Lines, Triangles, Squares, Pentagons, Hexagons, and other figures, instead of remaining fixed in their places, move freely about, on or in the surface, but without the power of rising above or sinking below it, very much like shadows – only hard with luminous edges – and you will then have a pretty correct notion of my country and countrymen. [...]
>
> The most common form for the construction of a house is five-sided or pentagonal [...]. The two Northern sides [...] constitute the roof, and for the most part have no doors; on the East is a small door for the Women; on the West a much larger one for the Men; the South side or floor is usually doorless. [...]"

In the twnetieth century, the Flatlanders were faced with the Water-Gas-Electricity



(a) (b)

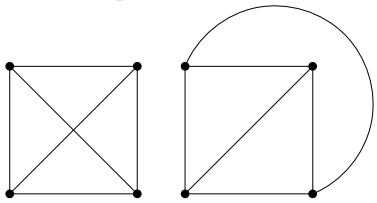
Figure 7.3: The Water-Gas-Electricity problem in Flatland

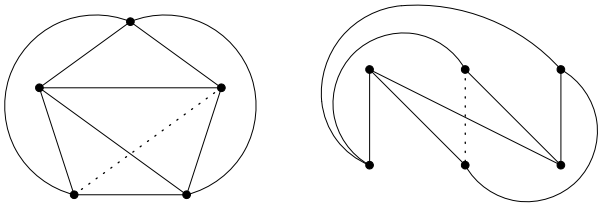
problem: provide each house with water, gas and electricity. It is easy to provide two houses with all of the three resources (Fig. 7.3 (a)), but is it possible to do it with three houses (Fig. 7.3 (b))? Of course, water pipes, gas pipes and electric wires are not allowed to intersect. [Answer: No, see below.]

Example 7.2 (a) K_4 is a planar graph; see the adjacent figure for a nonplanar and a planar representation of K_4.

(b) $K_5 - e$ is a planar graph, see Fig. 7.4 (a).

(c) $K_{3,3} - e$ is a planar graph, see Fig. 7.4 (b).





(a) (b)

Figure 7.4: Planar representations of $K_5 - e$ and $K_{3,3} - e$

Lemma 7.3 *Every subgraph of a planar graph is a planar graph.*

A subset $\Omega \subseteq \mathbb{R}^2$ of the real plane is *arcwise connected* if for every $a, b \in \Omega$, $a \neq b$, there is an arc $\gamma : [0,1] \to \Omega$ such that $\gamma(0) = a$ and $\gamma(1) = b$. A *region* is an open, arcwise connected subset of $\mathbb{R}^2$, Fig. 7.5 (a). A *closed Jordan curve* in $\mathbb{R}^2$ is

any continuous mapping $\gamma : [0,1] \to \mathbb{R}^2$ such that the restriction $\gamma|_{[0,1)} : [0,1) \to \mathbb{R}^2$ is injective and $\gamma(0) = \gamma(1)$.

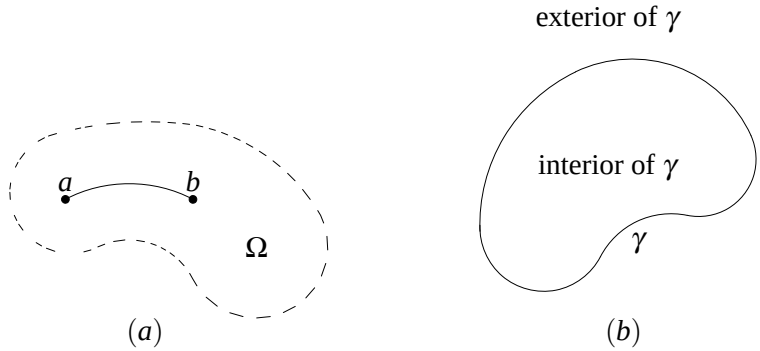


Figure 7.5: (a) A region; (b) Jordan's Theorem

Theorem 7.4 (Jordan) *Every closed Jordan curve γ splits the plane into two regions. One of them is bounded and is called the* interior of γ. *The other is unbounded and is called the* exterior of γ. *(Fig. 7.5 (b))*

A planar graph has many planar representations and from Jordan's Theorem it follows that every planar representation of a planar graph splits the plane into regions called *faces* of the representation. The famous result of Euler shows that although the geometry of planar representations may differ significantly, the number of faces does not depend on the representation.

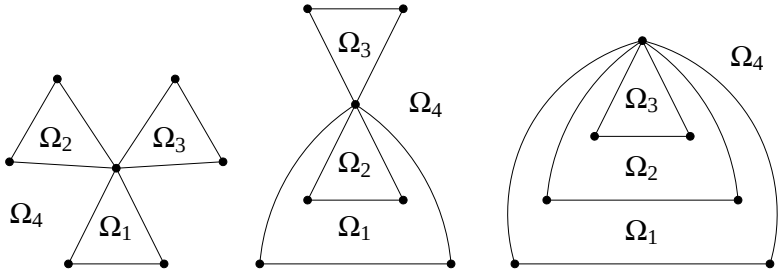
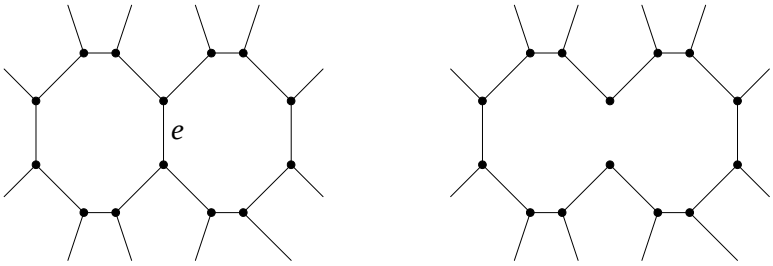


Figure 7.6: Three planar representations of the same graph

Example 7.5 In Fig. 7.6 we have three distinct planar representations of the same graph with 7 vertices and 9 edges. Each of the three representations has 4 faces (three bounded regions, and one unbounded region).

Fact 7.6 Let (v,ε) be a planar representation of a planar graph G which is not a tree and let e be an edge of G that belongs to a cycle of G. According to Lemma 7.3, $G-e$ is a planar graph. Let (v,ε') be a representation of $G-e$ obtained by restricting ε to $E\setminus\{e\}$. This is clearly a planar representation of $G-e$. If f is the number of faces of (v,ε) and f' the number of faces of (v,ε') then $f=f'+1$, see Fig. 7.7.



this representation of G has f faces

this representation of $G-e$ has $f-1$ faces

Figure 7.7: Removing an edge on a cycle increases the number of faces of the representaiton

Theorem 7.7 (Euler 1792) *Let (v,ε) be a planar representation of a connected, planar graph G with n vertices and m edges. If f is the number of faces of the representation, then $n-m+f=2$.*

Proof. The proof is by induction on m. If $m=0$ then $n=1$ and $f=1$, so the claim is obviously true. Suppose that the claim of the theorem is true for all planar representations of all connected, planar graphs with $<m$ edges and let G be a connected planar graph with m edges. Let (v,ε) be any planar representation of G and let f be the number of faces of this representation. If G is a tree then $f=1$ and $m=n-1$ so $n-m+f=2$. Assume now that G is not a tree. Then there is an edge e that belongs to a cycle of G. According to Lemma 7.3, $G-e$ is a planar graph. Let (v,ε') be a representation of $G-e$ obtained by restricting ε to $E\setminus\{e\}$. This is clearly a planar representation of $G-e$. Let m' be the number of edges of $G-e$ and let f' be the number of faces of (v,ε'). According to the induction hypothesis, $n-m'+f'=2$. Fact 7.6 implies $f=f'+1$ which together with $m=m'+1$ gives $n-m+f=2$. □

Theorem 7.8 *Let (v,ε) be a planar representation of a planar graph G with n vertices, m edges and ω connected components. If f is the number of faces of the representation, then $n-m+f-\omega=1$.*

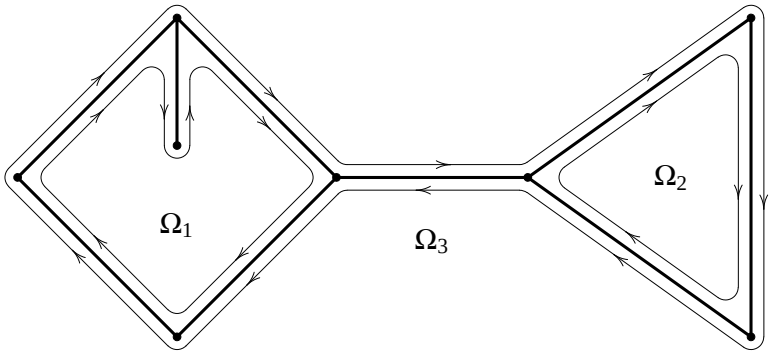


Figure 7.8: Boundaries of faces

As we have just seen, the number of faces f depends on the planar graph itself, rather than on the planar representation of the graph. In the sequel we therefore refer to f as the *number of faces of the planar graph G* and denote the number by $f(G)$.

The boundary $b(\Omega)$ of a face Ω of a planar representation of a planar graph G consists of arcs representing some edges of G. If $b(\Omega)$ consists of p edges and if q of them are cut-edges, then $p+q$ is the *length of* Ω and it is denoted by $l(\Omega)$. Note that cut-edges count twice!

Example 7.9 (a) Consider the plane representation of a planar graph given in Fig. 7.8. It has three faces whose lengths are $l(\Omega_1) = 6$, $l(\Omega_2) = 3$ and $l(\Omega_3) = 9$.

(b) The graph  has three faces; two of them have length 3, and the length of the third face is 6.

(c) An extreme example is the graph  with only one face Ω. Here $l(\Omega) = 4$.

Fact 7.10 Let G be a planar graph with at least two edges, let (v,ε) be any planar representation of G and let Ω be a face of (v,ε). Then $l(\Omega) \geqslant 3$. If G is bipartite then $l(\Omega) \geqslant 4$ since G has no odd cycles.

As a main corollary of Theorem 7.7 we shall now show that a planar graph cannot have "too many edges".

Corollary 7.11 *Let G be a planar graph with $n \geqslant 2$ vertices and m edges such that $G \ncong P_2$. Then $m \leqslant 3n-6$. Moreover, if G is bipartite, then $m \leqslant 2n-4$.*

Proof. Let $G = (V,E)$ be a planar graph with n vertices, m edges, ω connected components and f faces. If $m = 1$ then $G \ncong P_2$ implies $n \geqslant 3$ and $m \leqslant 2n-4 \leqslant$

$3n-6$ trivially holds.

Assume now that $m \geqslant 2$. First, from Theorem 7.8 we know that $n-m+f = \omega+1 \geqslant 2$. Let us estimate the number of faces f in terms of n and m. Let (v,ε) be a planar representation of G and let $\Omega_1, \ldots, \Omega_f$ be the faces of this representation. Since each edge which is not a cut-edge belongs to boundaries of exactly two faces, it follows that

$$l(\Omega_1)+\ldots+l(\Omega_f) = 2m.$$

On the other hand, $l(\Omega_i) \geqslant 3$ for all i (Fact 7.10), so

$$l(\Omega_1)+\ldots+l(\Omega_f) \geqslant 3f.$$

Therefore, $2m \geqslant 3f$ i.e. $\frac{2}{3}m \geqslant f$. Now

$$n-m+\frac{2}{3}m \geqslant n-m+f \geqslant 2$$

whence $m \leqslant 3n-6$ as required. The proof of the other part of the theorem is analogous. Just use the fact that $l(\Omega_i) \geqslant 4$. □

Corollary 7.11 is our main tool for showing that graphs are *not* planar. The idea behind all such proofs is that if a graph has "too many edges" it cannot be a planar graph (see Lemma 7.15).

Corollary 7.12 *If G is a planar graph, then $\delta(G) \leqslant 5$.*

Proof. Let $G=(V,E)$ be a planar graph with n vertices and m edges. If $n=1$ or $G \cong P_2$ then $\delta(G) \leqslant 1$. Otherwise, we have that $m \leqslant 3n-6$, so the assumption $\delta(G) \geqslant 6$ yields $2m = \sum_{v\in V}\delta(v) \geqslant 6n$, which contradicts $m \leqslant 3n-6$. □

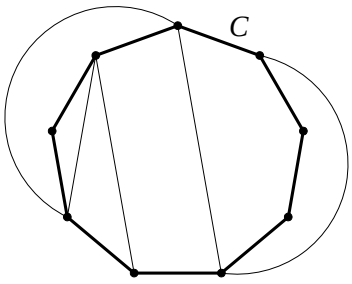


Figure 7.9: A planar Hamiltonian graph

We close this section by a discussion of planar Hamiltonian graphs. Let G be a planar Hamiltonian graph with the Hamiltonian cycle C and let (ν,ε) be a planar representation of G. Then the representation of C under (ν,ε) is a closed Jordan curve which splits the plane into two regions, the interior of C and the exterior of C, Fig. 7.9. Let $\mathrm{int}_C(l)$ denote the number of faces of (ν,ε) of length l which are in the interior of C, and let $\mathrm{ext}_C(l)$ denote the number of faces of (ν,ε) of length l which are in the exterior of C.

Example 7.13 For the planar representation of the Hamiltonian graph in Fig. 7.9 we have $\mathrm{int}_C(3)=2$, $\mathrm{int}_C(4)=1$, $\mathrm{int}_C(5)=1$, $\mathrm{ext}_C(4)=2$, $\mathrm{ext}_C(5)=1$, and all other int_C and ext_C values are 0.

Theorem 7.14 (Grinberg 1968) *Let C be a Hamiltonian cycle of a planar Hamiltonian graph G with $n\geqslant 3$ vertices. Take any planar representation of G. Then with respect to this represention,*

$$\sum_{l=3}^{n}(l-2)(\mathrm{int}_C(l)-\mathrm{ext}_C(l))=0.$$

Proof. Let (ν,ε) be a planar representation of G and let $\varepsilon(C)$ denote the closed Jordan curve that represents C. The cycle C contains all the vertices of G. Some of the edges from $E(G)\setminus E(C)$ belong to the interior of $\varepsilon(C)$, and the others belong to the exterior of $\varepsilon(C)$. Assume that s edges from $E(G)\setminus E(C)$ belong to the interior of $\varepsilon(C)$. These s edges divide the interior of $\varepsilon(C)$ into $s+1$ regions, whence

$$\sum_{l=3}^{n}\mathrm{int}_C(l)=s+1. \tag{7.1}$$

Each of these s edges belongs to the boundary of two faces in the interior of $\varepsilon(C)$ while each of the edges from $E(C)$ belongs to the boundary of exactly one face in the interior of $\varepsilon(C)$. Therefore,

$$\sum_{l=3}^{n}l\cdot\mathrm{int}_C(l)=2s+n. \tag{7.2}$$

Multiplying (7.1) by 2 and subtracting from (7.2) yields

$$\sum_{l=3}^{n}(l-2)\cdot\mathrm{int}_C(l)=n-2.$$

By the same argument,

$$\sum_{l=3}^{n}(l-2)\cdot\mathrm{ext}_C(l)=n-2,$$

and the theorem follows by subtracting the last two equalities. □

7.2 Combinatorial characterization of planar graphs

We are now going to present two deep results that show that, although defined in geometric terms, planarity is a purely combinatorial property of graphs. We start by showing that two small graphs are nonplanar.

Lemma 7.15 *The graphs K_5 and $K_{3,3}$ are not planar graphs.*

Proof. For K_5 we have $m = 10$ and $n = 5$. Since $m > 3n - 6$ it follows from Corollary 7.11 that K_5 is not a planar graph. Similarly, for $K_{3,3}$ we have $m = 9$ and $n = 6$, so $m > 2n - 4$ and $K_{3,3}$ is not a planar graph. □

The graphs K_5 and $K_{3,3}$ are paradigms of nonplanar graphs: theorems of Kuratowski and Wagner (see below) show that a graph is nonplanar if and only if it contains a sort of a copy of K_5 or $K_{3,3}$. In order to make this notion more precise, we introduce two graph editing operations: edge splitting, and edge contraction, Fig. 7.10.

Edge splitting: Let $e \in E(G)$, $e = \{u,v\}$, and let $x \notin V(G)$ be a new vertex. Let $G * e$ be a new graph obtained from G by replacing the edge e by the path $u\, e_1\, x\, e_2\, v$ where $e_1 = \{u,x\}$ and $e_2 = \{x,v\}$. More precisely,

$$V(G*e) = V(G-e) \cup \{x\}$$
$$E(G*e) = E(G-e) \cup \{e_1, e_2\}.$$

The we say that $G * e$ is obtained from G by *splitting the edge e*.

Edge contraction: Let $e \in E(G)$, $e = \{u,v\}$, and let $x \notin V(G)$ be a new vertex. Let G/e be a new graph obtained from G by replacing vertices u and v by the vertex x and joining the neighbours of u and v to x. More precisely,

$$V(G/e) = V(G-u-v) \cup \{x\}$$
$$E(G/e) = E(G-u-v) \cup \{\{x,w\} : w \in (N_G(u) \cup N_G(v)) \setminus \{u,v\}\}.$$

The we say that G/e is obtained from G by *contracting the edge e*.

A graph H is a *subdivision of a graph G* if H can be obtained from G by a finite sequence of edge spilttings. A graph H is a *contraction of a graph G* if H can be obtained from G by a finite sequence of edge contractions. A graph, one of its subdivisions and one of its contractions is depicted in Fig. 7.11.

Planarity is invariant under edge splittings and edge contractions, as the following lemma shows.

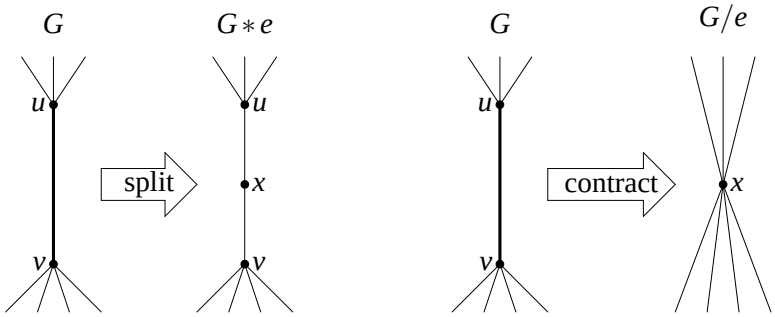


Figure 7.10: Edge splitting and edge contraction

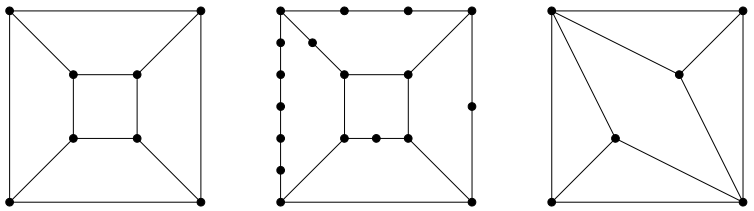


Figure 7.11: A subdivision and a contraction of a graph

Fact 7.16 If H is a subdivision of G, then G is planar if and only if H is planar. A contraction of a planar graph is a planar graph. The converse of the last statement does not hold: just take K_5 and any of its contractions.

Finally, the following pair of theorems shows that the presence of K_5 or $K_{3,3}$ is the main reason for nonplanarity of a graph.

Theorem 7.17 (Kuratowski 1930) *A graph is planar if and only if it does not have a subgraph that is a subdivision of K_5 or $K_{3,3}$.*

Theorem 7.18 (Wagner 1937) *A graph is planar if and only if it does not have a subgraph that is a contraction of K_5 or $K_{3,3}$.*

Example 7.19 Finding subdivisions of K_5 or $K_{3,3}$ in a nonplanar graph can be tricky. For example, the Petersen graph is easily seen to have a K_5 as its contraction, while it is not so easy to find a subgraph that is a subdivision of $K_{3,3}$, Fig. 7.12. Surprisingly, no subgraph of the Petersen graph is a subdivision of K_5.

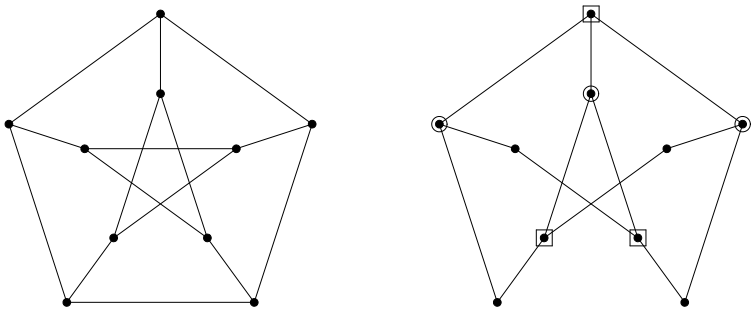


Figure 7.12: The Petersen graph and a subdivision of $K_{3,3}$ as its subgraph

7.3 Regular polyhedra

The five regular polyhedra, or Platonic bodies: tetrahedron, hexahedron, octahedron, dodecahedron and icosahedron (see Fig. 7.13), had been known to geometers of Ancient Greece, but there was no proof that these are the only ones until L. Euler proved a version of Theorem 7.7. We shall now demonstrate the application of the graph-theoretic version of the Euler's result to show that these are indeed the only regular polyhedra.

To each regular polyhedron P we can in an obvious way assign a planar graph $G(P)$: vertices of P correspond to vertices of $G(P)$, and edges of P correspond to the edges of $G(P)$. Clearly, faces of every planar representation of $G(P)$ correspond to faces of P. Now, regularity of the polyhedron translates to the requirement that $G(P)$ be a regular graph and that all faces of a planar representation of $G(P)$ be of the same length. Graphs of the five regular polyhedra are given in Fig. 7.14.

Lemma 7.20 *Let P be a regular polyhedron. Then*
(a) $3 \leqslant \delta(v) \leqslant 5$ for every vertex v of $G(P)$, and
(b) $3 \leqslant l(\Omega) \leqslant 5$ for every face Ω of every planar representation of $G(P)$.

Proof. (a) Take any vertex v of $G(P)$. For geometric reasons we have $\delta(v) \geqslant 3$, while $\delta(v) \leqslant 5$ follows from Corollary 7.12 since $G(P)$ is a regular graph.

(b) Take any face Ω of a planar representation of $G(P)$. Then $l(\Omega) \geqslant 3$ for geometric reasons. All faces of P have the same length since P is a regular polyhedron, so in order to prove $l(\Omega) \leqslant 5$ it suffices to show that there exists a face whose length is $\leqslant 5$. Using the ideas from proofs of Corollaries 7.11 and 7.12 it is easy to show that $m \leqslant 3f - 6$ whence there exists a face Ω' such that $l(\Omega') \leqslant 5$ (Homework 7.10). □

Let P be a regular polyhedron with n vertices, m edges and f faces. Then $G(P)$

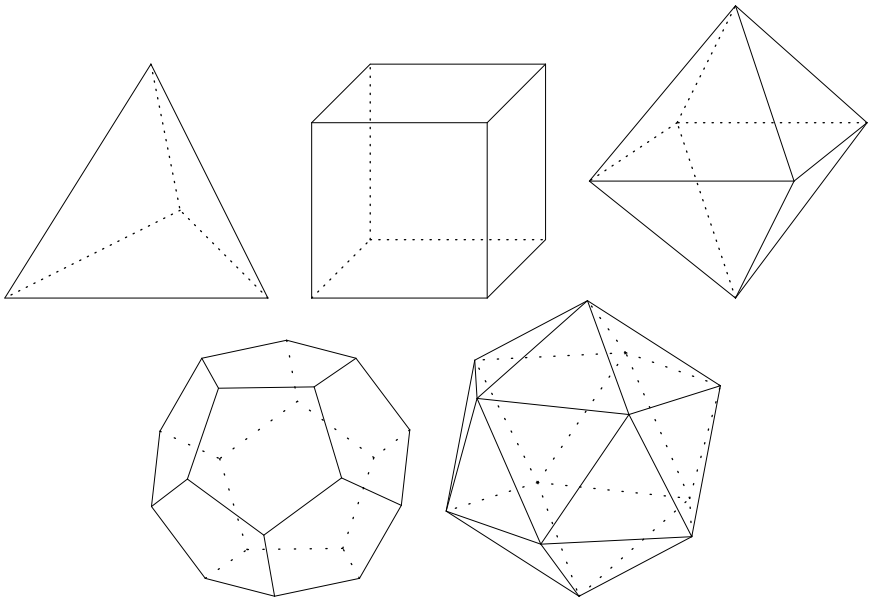


Figure 7.13: The five regular polyhedra

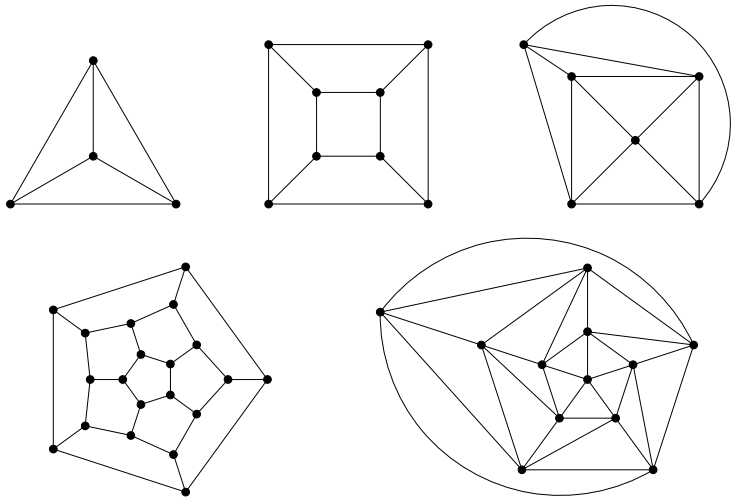


Figure 7.14: Graphs of the five regular polyhedra

is a planar connected graph, so $n-m+f=2$ by Theorem 7.7. Moreover, $G(P)$ is a regular graph and every face of every planar representation of G has the same length. Let δ be the common degree of vertices of $G(P)$ and let l be the common length of faces of planar representations of $G(P)$. Then from the First Theorem of Graph Theory and by the counting argument we used to prove Corollary 7.11,

$$n\cdot\delta = l\cdot f = 2m.$$

Now, $4n-2m-2m+4f=8$ together with $n\cdot\delta = l\cdot f = 2m$ yields $4n-n\cdot\delta-l\cdot f+4f=8$ whence

$$n(4-\delta)+f(4-l)=8.$$

From Lemma 7.20 we know that $\delta\in\{3,4,5\}$ and $l\in\{3,4,5\}$, so there are nine cases to discuss.

(1) $\delta=l=3$: then $n+f=8$ and $3n=3f$, whence $n=f=4$ and P is the **tetrahedron**.

(2) $\delta=3$, $l=4$: then $n=8$ and $3n=4f$, whence $f=6$ and P is the **hexahedron**.

(3) $\delta=3$, $l=5$: then $n-f=8$ and $3n=5f$, whence $n=20$, $f=12$ and P is the **dodecahedron**.

(4) $\delta=4$, $l=3$: then $f=8$ and $4n=3f$, whence $n=6$ and P is the **octahedron**.

(5) $\delta=l=4$: then $0=8$ – impossible.

(6) $\delta=4$, $l=5$: then $-f=8$ – impossible.

(7) $\delta=5$, $l=3$: then $-n+f=8$ and $5n=3f$, whence $n=12$, $f=20$ and P is the **icosahedron**.

(8) $\delta=5$, $l=4$: then $-n=8$ – impossible.

(9) $\delta=l=5$: then $-n-f=8$ – impossible.

Therefore, there are only five regular polyhedra.

Homework

7.1. Prove Lemma 7.3.

7.2. Prove Theorem 7.8.

7.3. Complete the proof of Corollary 7.11 by showing that $m \leqslant 2n-4$ if G is a bipartite planar graph with n vertices and m edges.

7.4. A *maximal planar graph* is a graph $G=(V,E)$ such that

- G is a planar graph, and
- if $G'=(V,E')$ is a planar graph with the same set of vertices and with $E' \supseteq E$ then $E'=E$.

Show that every maximal planar graph G is connected and $l(\Omega)=3$ for every face Ω of every planar representation of G.

7.5. Let G be a planar graph with $n \geqslant 3$ vertices and m edges. Show that G is a maximal planar graph if and only if $m=3n-6$.

7.6. Show that among nonplanar graphs K_5 has the smallest number of vertices and $K_{3,3}$ has the smallest number of edges.

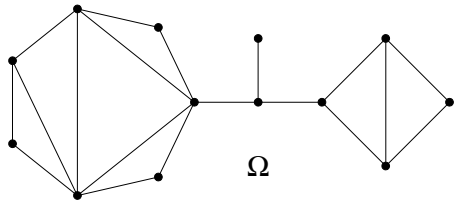


Figure 7.15: An outer planar graph

7.7. A graph G is called *outer planar* if there exists a planar representation (v,ε) of G and a face Ω of this representation such that all vertices of G belong to the boundary of Ω. An example of an outer planar graph is given in Fig. 7.15.

For a graph G let G^* denote the graph obtained by adding a new vertex to G and joining the new vertex to every vertex of G. More precisely, if $x \notin V(G)$, let

$$V(G^*) = V(G) \cup \{x\}$$
$$E(G^*) = E(G) \cup \{\{x,v\} : v \in V(G)\}.$$

Show that G is an outer planar graph if and only if G^* is a planar graph.

7.8. Show that a graph is outer planar if and only if it does not have a subgraph that is a subdivision of K_4 or $K_{2,3}$.

7.9. Show that a graph is outer planar if and only if it does not have a subgraph that is a contraction of K_4 or $K_{2,3}$.

7.10. Let G be a connected planar graph with m edges and f faces such that $\delta(G) \geqslant 3$.

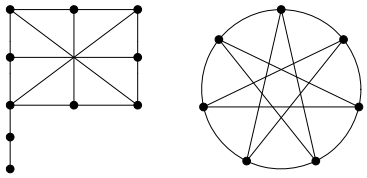
(a) Show that $m \leqslant 3f - 6$. (Hint: Use the idea of the proof of Corollary 7.11 together with the fact that $2m \geqslant 3n$.)

(b) Show that every planar representation of G has a face Ω such that $l(\Omega) \leqslant 5$. (Hint: Use (a) and the idea of the proof of Corollary 7.12.)

Exercises

7.11. (a) Show that the two graphs in the adjacent figure are not planar.

(b) Show that $\overline{Q}_3$ is not a planar graph.



7.12. Find two 3-regular graphs with the same number of vertices such that one of them is planar and the other is not.

7.13. (a) Let G be a planar graph such that $\delta(G) \geqslant 5$. Show that G has at least 12 vertices whose degree is exactly 5.

(b) Find an example of a planar graph with 12 vertices such that the degree of every vertex of the graph is exactly 5.

7.14. Let G be a graph with $n > 10$ vertices. Show that at least one of the graphs G, $\overline{G}$ is not planar.

7.15. Find all trees T such that $\overline{T}$ is a planar graph.

7.16. Let G be a connected planar graph with $\delta(G) \geqslant 4$. Show that every planar representation of G has a face of length 3.

7.17. Let G be a planar graph with n vertices and m edges and let $g > 0$ be the minimal length of a cycle in G. Show that

$$m \leqslant \frac{g}{g-2}(n-2).$$

7.18. (a) Show that Q_n contains a subdivision of K_{n+1}.

(b) Find all n such that Q_n is a planar graph.

7.19. Let $G=(V,E)$ be a nonplanar graph. The *thickness of* G, denoted by $\theta(G)$, is the smallest positive integer k with the property that there is a partition $\{E_1,\ldots,E_k\}$ of E such that (V,E_i) is a planar graph for all i.

(a) Let G be a graph with $n \geqslant 3$ vertices and m edges. Show that

$$\theta(G) \geqslant \frac{m}{3n-6}.$$

(b) Show that $\theta(K_{6s-1}) > s$.

7.20. Let T be a tree with at least 4 vertices and let e_1, e_2, $e_3 \in E(\overline{T})$ be three edges not in T. Show that $T+e_1+e_2+e_3$ is a planar graph.

7.21. (a) Show that a graph with three edge-disjoint spanning trees cannot be planar.

(b) Show that a bipartite graph with two edge-disjoint spanning trees cannot be planar.

7.22. Is there a convex polyhedron (not necessarily a regular one) whose faces are hexagons?

7.23. The graph G in Fig. 7.16 is clearly a Hamiltonian graph. Is there a Hamiltonian cycle of G which contains edges e_1 and e_2?

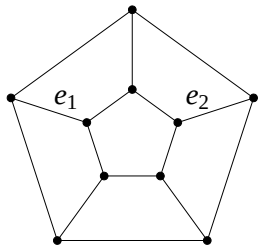


Figure 7.16: A planar Hamiltonian graph

7.24. For a connected graph G let $\tau(G)$ denote the number of distinct (possibly isomorphic) spanning trees of G. Show that $\tau(G)=\tau(G-e)+\tau(G/e)$ for any $e \in E(G)$.

Chapter 8

Graph colourings

Colouring vertices and edges of graphs is one of the most popular topics in graph theory. Popularity aside, applications of graph colourings range from scheduling meetings of committees to compiler optimizations.

In this chapter we address some basic problems concerning colouring vertices and colouring edges of graphs. We discuss the famous Four Colour Problem which states that every planar graph is 4-colourable.

8.1 Colouring vertices

Suppose p senators $x_1, \ldots, x_p$ are members of q committees $M_1, \ldots, M_q$, of the University Senate and suppose that a senator can be a member of more than one committee. Then a schedule of meetings of the q committees has to be made in such a way that committees that share members cannot meet at the same time. When planning committee meetings, one of the fundamental parameters of the schedule is the number of time slots that have to be allocated.

A graph-theoretic interpretation of this problem can be made as follows. Consider a graph G with vertices $M_1, \ldots, M_q$ where M_i is adjacent to M_j if $i \neq j$ and committees M_i and M_j share a member. If we enumerate time slots by $1, \ldots, k$ and assign a committee M_i a time slot s_i then clearly $s_i \neq s_j$ whenever M_i is adjacent to M_j. This is because adjacent vertices in G correspond to committees that share members and hence the meetings are not allowed to be scheduled at the same time.

Let B be a finite nonempty set that we think of as the *set of colours*. A *vertex colouring of a graph* $G = (V,E)$ is any mapping $f : V \to B$. A vertex colouring $f : V \to B$ is called *proper* if adjacent vertices are coloured by distinct colours, that is, $\{u,v\} \in E$ implies $f(u) \neq f(v)$, for all $u,v \in V$. A graph $G = (V,E)$ is *k-colourable* if there exists a proper colouring $f : V \to \{1,\ldots,k\}$. The *chromatic number of G*

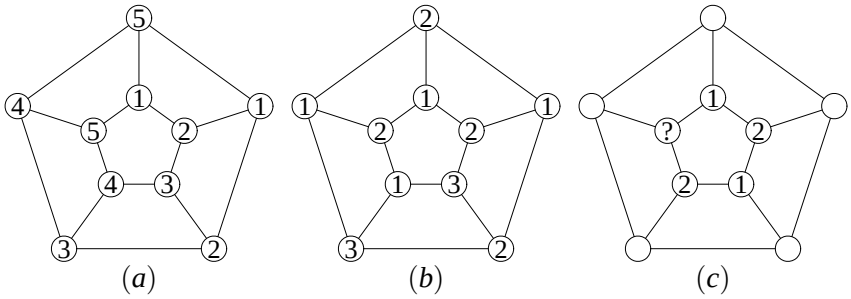


Figure 8.1: A graph with (a) a proper vertex colouring with 5 colours; (b) a proper vertex colouring with 3 colours; (c) no proper vertex colouring with 2 colours

denoted by $\chi(G)$ is the least positive integer k such that G is k-colourable.

We see that if a graph G is k-colourable, then it is l-colourable for every $l \geqslant k$. This is because we are *not* obliged to use all the colours from B. It is also easy to see that if $H \leqslant G$ then $\chi(H) \leqslant \chi(G)$.

Example 8.1 Consider a graph G in Fig. 8.1 (a). As we can see, the graph is 5-colourable. It is also easy to see that the same graph in 3-colourable, Fig. 8.1 (b), and that it is *not* 2-colourable, Fig. 8.1 (c). Therefore, $\chi(G) = 3$.

Note that $\chi(G) \leqslant k$ is equivalent to the fact that G is k-colourable. On the other hand, to show that $\chi(G) = k$ we have to show

- that k colours suffice for proper vertex colouring of G, that is $\chi(G) \leqslant k$; we usually show this by exibiting an explicite proper vertex colouring of G with k colours; and

- that k colours are necessary for proper vertex colouring of G, that is $\chi(G) \geqslant k$; we usually show this by assuming that k colours suffice and then deriving a contradiction, or by showing that K_k is a subgraph of G.

Lemma 8.2 (a) $\chi(G) = 1$ *if and only if* $E(G) = \varnothing$.
(b) $\chi(G) = 2$ *if and only if* $E(G) \neq \varnothing$ *and* G *is a bipartite graph.*
(c) $\chi(K_n) = n$.
(d) *If* $K_s \leqslant G$ *then* $\chi(G) \geqslant s$.
(e) $\chi(C_{2s}) = 2$ *and* $\chi(C_{2s+1}) = 3$.

Proof. We shall prove (c) and the second part of (e).

To show (c) note first that n colours suffice to colour vertices of K_n, so $\chi(K_n) \leqslant n$. On the other hand, we need at least n colours to colour the vertices of K_n prop-

erly since every pair of vertices of K_n is adjacent, whence $\chi(K_n) \geqslant n$. Therefore, $\chi(K_n) = n$.

To show the second part of (e), note first that from (a) and (b) it follows that $\chi(C_{2s+1}) \neq 1,2$, so $\chi(C_{2s+1}) \geqslant 3$. On the other hand, it is easy to see that three colours suffice to colour $\chi(C_{2s+1})$, whence $\chi(C_{2s+1}) = 3$. □

Finding the exact value of $\chi(G)$ is usually a very complicated task. We shall therefore present two standard upper bounds. For a graph G let

$$\inf(G) = \max\{\delta(H) : H \leqslant G\}.$$

Theorem 8.3 $\chi(G) \leqslant \inf(G) + 1$.

Proof. The proof is by induction on $n = n(G)$. The claim is trivially true for $n = 1,2,3$. Assume that $\chi(H) \leqslant \inf(H) + 1$ for all graphs H with less than n vertices and let G be a graph with n vertices. Let $v \in V(G)$ be a vertex of G such that $\delta(v) = \delta(G)$, let $d = \delta(v)$ and $N(v) = \{w_1, \ldots, w_d\}$. Put $H = G - v$. By the induction hypothesis we have $\chi(H) \leqslant \inf(H) + 1$, so there is a proper colouring $f : V(H) \to B$ of H with $|B| = \inf(H) + 1$ colours. In order to complete the proof we are going to extend the colouring f to a proper colouring of G.

From $H \leqslant G$ it follows that $\inf(H) \leqslant \inf(G)$ (Homework 8.2) so let $B' \supseteq B$ be a superset of B such that $|B'| = \inf(G) + 1$ and let $f' : V(H) \to B'$ be a mapping such that $f'(v) = f(v)$ for all $v \in V(H)$. Clearly, f' is a proper vertex colouring of H with $\inf(G) + 1$ colours. The colouring f' uses at most d colours for colouring the neighbours $w_1, \ldots, w_d$ of v. Now $d = \delta(G) \leqslant \inf(G)$, so at least one of the $\inf(G) + 1$ colours from B' is not used for colouring of neighbours of v. Therefore, $B' \setminus \{f'(w_1), \ldots, f'(w_d)\} \neq \varnothing$. Take any $c \in B' \setminus \{f'(w_1), \ldots, f'(w_d)\}$ and define $f^* : V(G) \to B'$ by

$$f^*(x) = \begin{cases} c, & x = v \\ f'(x), & \text{otherwise.} \end{cases}$$

It is easy to see that f^* is a proper vertex colouring of G by $\inf(G) + 1$ colours, so $\chi(G) \leqslant \inf(G) + 1$. □

The parameter $\inf(G)$ is not one of the "standard" parameters. Although rather easy to compute (see Homework 8.3), we prefer to replace it by more convenient ones.

Corollary 8.4 $\chi(G) \leqslant \Delta(G) + 1$.

Proof. Since $\delta(H) \leqslant \Delta(H) \leqslant \Delta(G)$ for every $H \leqslant G$, it is easy to see that $\inf(G) \leqslant \Delta(G)$. □

The following theorem shows that odd cycles and complete graphs are the only classes of graphs where the upper bound in the above corollary is reached. The proof of this fact is surprisingly complicated and we shall omit it.

Theorem 8.5 (Brooks, 1941) *Let G be a graph which is neither the complete graph, nor the odd cycle. Then $\chi(G) \leqslant \Delta(G)$.*

Finally, we shall present an upper bound on the chromatic number of $\chi(G)$ which depends on the number of edges of the graph.

Theorem 8.6 *Let G be a graph with m edges. Then $\chi(G) \leqslant \frac{1}{2} + \sqrt{2m + \frac{1}{4}}$.*

Proof. Let $k = \chi(G)$ and let $f : V(G) \to \{1, \ldots, k\}$ be a proper vertex colouring of G. For $i \in \{1, \ldots, k\}$ let $V_i = \{v \in V(G) : f(v) = i\}$ be the set of all vertices of G coloured by the colour i. Then $E(V_i, V_j) \neq \varnothing$ whenever $i \neq j$ (Homework 8.5), whence $m \geqslant \binom{k}{2} = \frac{1}{2}k(k-1)$. Solving for k we obtain $\chi(G) = k \leqslant \frac{1}{2} + \sqrt{2m + \frac{1}{4}}$. □

8.2 The Four Colour Problem

Around 1850, Francis Guthrie (1831–1899) showed how to colour a map of all the counties in England using only four colours so that any two neighboring regions have different colours. He became interested in the general problem and conjectured that the smallest number of colours needed to colour any planar map so that any two neighboring regions have different colours is four. Guthrie talked about his conjecture with his brother, Frederick. Frederick talked about it with his mathematics teacher, Augustus DeMorgan (from the DeMorgan's laws in logic), who sent the problem to William Hamilton (for whom Hamiltonian mechanics is named). Hamilton was evidently too interested in other things to work on the four colour problem, and it seemed to have been forgotten for about 25 years. In 1878, Arthur Cayley made the scientific community aware of the problem again, and shortly thereafter, British mathematician Sir Alfred Kempe devised a "proof" that was unquestioned for over ten years. However, in 1890, another British mathematician, Percy John Heawood, found a mistake in Kempe's work. The Four Colour Problem remained unsolved until 1976, when Kenneth Appel and Wolfgang Haken produced a proof involving an intricate computer analysis of 1936 different configurations. Although some mathematicians have expressed dissatisfaction with

Figure 8.2: A political map of Europe and its associated planar graph

Appel's and Haken's proof, there is still no proof of the Four colour Problem that does not involve computer analysis.

Colouring maps is closely related to colouring vertices of planar graphs. To each region of a planar map we assign a vertex of a graph and join two vertices by an edge if the two regions have a common border. The main observation is that if a map is not too weird (i.e. no country has a hole, which is the case with Italy where Vatican makes a hole, or no country consists of two or more parts, which is the case with Russia where the region around Kaliningrad is detached from the rest of the country) then the graph associated to the map is planar and every proper colouring of the map uniquely determines a proper vertex colouring of the graph. For example, a political map of Europe (as of year 2001, and ignoring Vatican and the region around Kaliningrad to make Italy and Russia simply connected) together with the associated planar graph is given in Fig. 8.2. The Four Colour Problem now becomes a statement about planar graphs: $\chi(G) \leqslant 4$ *for every planar graph* G.

As an easy consequence of Theorem 8.3 we immediately obtain the solution to the "Six Colour Problem":

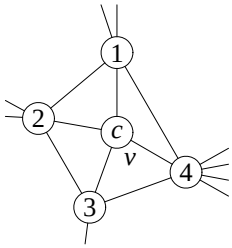
Corollary 8.7 $\chi(G) \leqslant 6$ *for every planar graph* G.

Solving the "Five Colour Problem" requires a little bit more work:

Theorem 8.8 (Heawood 1890) $\chi(G) \leqslant 5$ *for every planar graph* G.

Proof. The proof is by induction on the number of vertices of G. For planar graphs with 1, 2, 3, 4 and 5 vertices the statement is obviously true. Assume that every planar graph with $< n$ vertices satisfies $\chi \leqslant 5$ and let G be a planar graph with n vertices. As we have seen, $\delta(G) \leqslant 5$ so there is a vertex $v \in V(G)$ such that $\delta_G(v) \leqslant 5$.

If $\delta_G(v) \leqslant 4$, let $G' = G - v$. Then G' is a planar graph with $< n$ vertices and by the induction hypothesis there is a proper vertex colouring $f' : V(G') \to \{1, \ldots, 5\}$ of G' with 5 colours. Since v has at most four neighbours, at least one colour c does *not* appear as a colour of one of the neighbours of v, so



$$f(x) = \begin{cases} c, & x = v \\ f'(x), & \text{otherwise.} \end{cases}$$

is a proper vertex colouring of G with 5 colours.

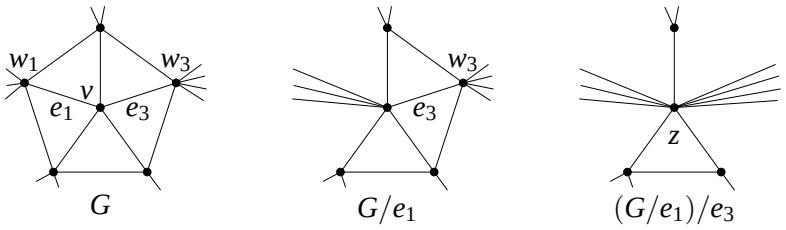


Figure 8.3: The "Five Colour Problem" for planar graphs

Assume now that $\delta_G(v) = 5$ and let $w_1, \ldots, w_5$ be all the neighbours of v. There exist $i \neq j$ such that w_i and w_j are not adjacent, for otherwise $G[w_1, \ldots, w_5]$ would be isomorphic to K_5. Suppose that w_1 and w_3 are not adjacent and let $e_1 = \{v, w_1\}$ and $e_3 = \{v, w_3\}$. Consider $G' = (G/e_1)/e_3$ and denote the new vertex obtained by contracting edges e_1 and e_3 by z, Fig. 8.3. Then G' is a planar graph (since planarity is invariant under contracting edges) with $< n$ vertices and by the induction hypothesis there is a proper vertex colouring $f' : V(G') \to \{1, \ldots, 5\}$ of G' with 5 colours. Take a $c \in \{1, \ldots, 5\} \setminus \{f'(z), f'(w_2), f'(w_4), f'(w_5)\}$ and define

$f^* : V(G) \to \{1,\ldots,5\}$ by

$$f^*(x) = \begin{cases} c, & x = v \\ f'(z), & x = w_1 \text{ or } x = w_3 \\ f'(x), & \text{otherwise.} \end{cases}$$

It is easy to see that f^* is a proper vertex colouring of G with five colours, whence $\chi(G) \leqslant 5$. □

However, even more is true:

Theorem 8.9 (Appel, Haken 1976) *$\chi(G) \leqslant 4$ for every planar graph G.*

Theorem 8.10 (Grötzsch 1959) *Let G be a planar graph such that $C_3 \not\leqslant G$. Then $\chi(G) \leqslant 3$.*

8.3 Colouring edges

Let B be a finite nonempty set of colours. An *edge colouring of a graph $G = (V,E)$* is any mapping $f : E \to B$. A colouring $f : E \to B$ is called *proper* if adjacent edges are coloured by distinct colours, that is, $|e_1 \cap e_2| = 1$ implies $f(e_1) \neq f(e_2)$, for all $e_1, e_2 \in E$. A graph $G = (V,E)$ is *k-edge-colourable* if there exists a proper colouring $f : E \to \{1,\ldots,k\}$. The *chromatic index of G* (sometimes also called the *edge chromatic number of G*) denoted by $\chi'(G)$ is the least positive integer k such that G is k-edge-colourable.

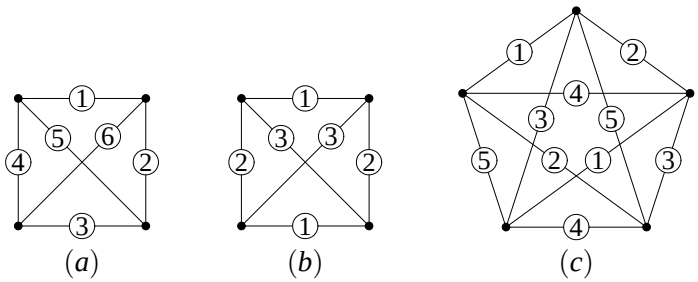


Figure 8.4: (a) K_4 is 6-edge-colourable; (b) K_4 is 3-edge-colourable; (c) K_5 is 5-edge-colourable

Example 8.11 The graph K_4 is 6-edge-colourable (Fig. 8.4 (a)), but also 3-edge-colourable (Fig. 8.4 (b)). K_4 is not 2-edge-colourable since it has a vertex of degree 3, so $\chi'(K_4) = 3$. The graph K_5 is 5-edge-colourable (Fig. 8.4 (c)), but not 4-edge-colourable as we shall see below. Therefore, $\chi'(K_5) = 5$.

Clearly, $\chi'(G) \geqslant \Delta(G)$. A surprising theorem due to Vizing states that $\chi'(G)$ is either $\Delta(G)$ or $\Delta(G)+1$.

Theorem 8.12 (Vizing, 1964) *For every graph G, $\Delta(G) \leqslant \chi'(G) \leqslant \Delta(G)+1$.*

While there are only rough estimates for $\chi(G)$, the "dual" notion of $\chi'(G)$ is strictly bounded. Therefore, determining the chromatic index of a graph reduces to deciding which of the two possible values occurs. A general strategy for finding χ' is to try to construct an edge colouring with Δ colours. From Vizing's Theorem 8.12 it follows that if such a colouring exists then $\chi' = \Delta$, otherwise $\chi' = \Delta + 1$.

Theorem 8.13 *Let $n \geqslant 3$. Then* $\chi'(K_n) = \begin{cases} n-1, & n \text{ is even} \\ n, & n \text{ is odd.} \end{cases}$

Proof. Let n be an even integer. To show that $\chi'(K_n) = n-1 = \Delta(K_n)$ it suffices to produce a proper edge colouring of K_n with $n-1$ colours. Let $\{0,1,\ldots,n-1\}$ be the set of vertices of K_n and let us consider a particular representation of K_n in Euclidean plane where vertices 0, …, $n-2$ are vertices of a regular $(n-1)$-gon and $n-1$ is its center, Fig. 8.5 (a) and (b). For each $j \in \{0,\ldots,n-2\}$ let E_j be the following set of edges of K_n:

$$E_j = \big\{\{j, n-1\}\big\} \cup \big\{\{u,v\} : u+v \equiv 2j \pmod{n-1}\big\},$$

see Fig. 8.5 (c_0), …, (c_6) where E_0, …, E_6 are depicted in case of K_8. Note that the set of edges E_j understood as a geometric configuration is a rotation of E_0 about $n-1$ through the angle $\varphi_j = j \cdot \dfrac{2\pi}{n-1}$. The edges in each E_j are independent (i.e. no two are adjacent), and $\{E_0,\ldots,E_{n-2}\}$ is a partition of $E(K_n)$ (Homework 8.8). Therefore, the edge colouring $f : E(K_n) \to \{0,\ldots,n-2\}$ given by

$$f(e) = j \quad \text{if and only if} \quad e \in E_j$$

is a proper edge colouring of K_n with $n-1$ colours.

For the second part of the proof, let n be an odd integer and let us show that $\chi'(K_n) \neq \Delta(K_n) = n-1$. Then by Vizing's Theorem 8.12 it follows that $\chi'(K_n) = \Delta(K_n)+1 = n$. Assume to the contrary that $\chi'(K_n) = \Delta(K_n) = n-1$ and let $f : E(K_n) \to \{1,\ldots,n-1\}$ be a proper edge colouring of K_n with $n-1$ colours. Let

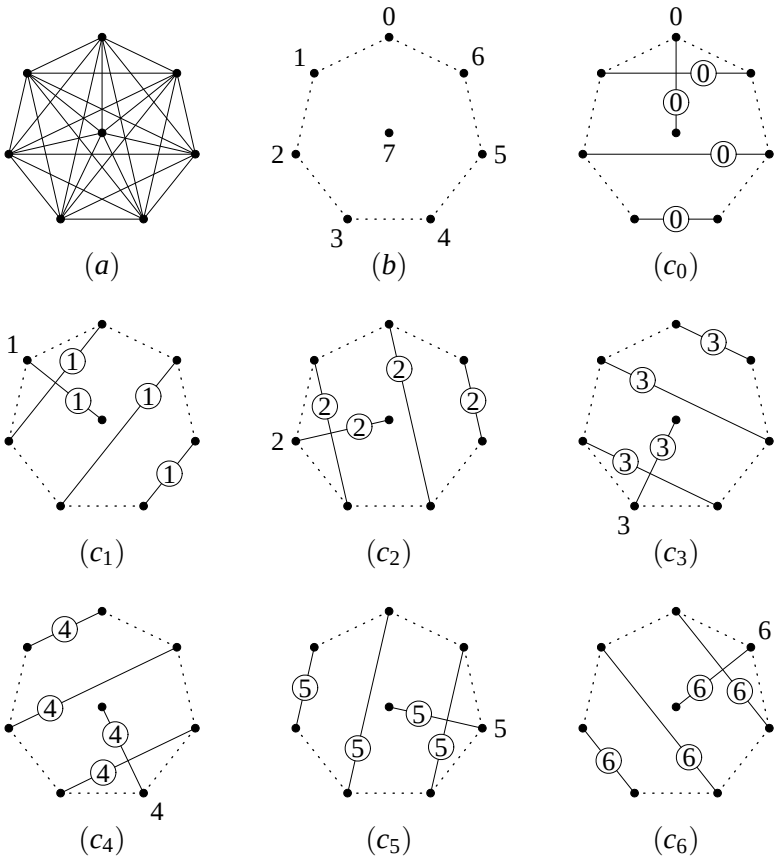


Figure 8.5: A proper edge colouring of K_8 with 7 colours

$E_1 = \{e \in E(K_n) : f(e) = 1\}$. Since every vertex of K_n is of degree $n-1$ and since edges incident to the same vertex have to be coloured by distinct colours, it follows that every vertex of K_n is incident with an edge from E_1. Therefore, $n = 2 \cdot |E_1|$ which is an even integer—contradiction. □

Theorem 8.14 (König 1916) *If G is a bipartite graph then $\chi'(G) = \Delta(G)$.*

Proof. We use induction on $m = m(G)$. If $m = 0$ the claim is obviously true. Assume that $\chi'(H) = \Delta(H)$ for all bipartite graphs H with $< m$ edges and let G be a bipartite graph with m edges. Take any $e = \{u,v\} \in E(G)$ and let $H = G - e$. Then H is a bipartite graph with $< m$ edges, so $\chi'(H) = \Delta(H)$ by the induction hypothesis. Since $\Delta(H) \leqslant \Delta(G)$ we obtain that $\chi'(H) \leqslant \Delta(G)$, i.e. H is $\Delta(G)$-edge-colourable.

Let $f : E(H) \to \{1,\dots,\Delta\}$ be a proper edge colouring of H where $\Delta = \Delta(G)$. For $x \in V(G) = V(H)$ let $B(x)$ denote the set of all colours that occur as colours of edges incident to x, that is, the set of all $c \in \{1,\dots,\Delta\}$ such that there is an edge $d \in E(H)$ incident to x and $f(d) = c$. If $c \in B(x)$ we say that the colour c is *present at x*; otherwise we say that c is *absent at x*. The colouring f takes care of all the edges of G except for the edge $e = \{u,v\}$. Our intention is to adjust f so as to turn it into the edge colouring of entire G.

Since $\delta_H(u) = \delta_G(u) - 1 < \Delta$ and $\delta_H(v) = \delta_G(v) - 1 < \Delta$, at least one colour is absent at u and at least one colour is absent at v. If there is a colour $c \in \{1,\dots,\Delta\} \setminus (B(u) \cup B(v))$ that is absent both at u and at v, we can straightforwardly extend f to $f^* : E(G) \to \{1,\dots,\Delta\}$ given by

$$f^*(d) = \begin{cases} c, & d = e \\ f(d), & \text{otherwise} \end{cases}$$

which is a proper edge colouring of G with Δ colours.

Assume now that $B(u) \cup B(v) = \{1,\dots,\Delta\}$, i.e., that every colour is present at u or at v. Let $b \in \{1,\dots,\Delta\} \setminus B(u)$ be a colour that is absent at u and $c \in \{1,\dots,\Delta\} \setminus B(v)$ a colour that is absent at v. According to the assumption, $b \neq c$, $b \in B(v)$ and $c \in B(u)$. Let $e_1 = \{v, w_1\}$ be an edge incident with v with $f(e_1) = b$ and let

$$P = v\, e_1\, w_1\, e_2\, w_2\, e_3\, w_3\, \dots\, e_k\, w_k$$

be the longest *alternating b/c-path* that starts with $v\, e_1$, that is, the longest path starting with $v\, e_1$ such that $f(e_1) = b$, $f(e_2) = c$, $f(e_3) = b$, $f(e_4) = c$ etc, Fig. 8.6.

Let us show that $u \notin \{v, w_1, \dots, w_k\}$. Suppose to the contrary that u appears as a vertex of P, say $u = w_l$, and let $P' = v\, e_1\, w_1\, e_2 \dots\, w_{l-1}\, e_l\, u$. Note first that

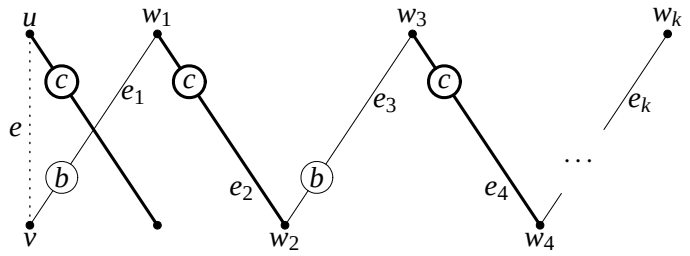


Figure 8.6: A maximal alternating b/c-path

$f(e_l) = c$ because $b \notin B(u)$. Since P, and hence P', is an alternating b/c-path that starts with an edge coloured by b, from $f(e_l) = c$ it follows that the length of P' is even. Therefore, $P' + e$ is a cycle of odd length—contradiction with the fact that G is a bipartite graph.

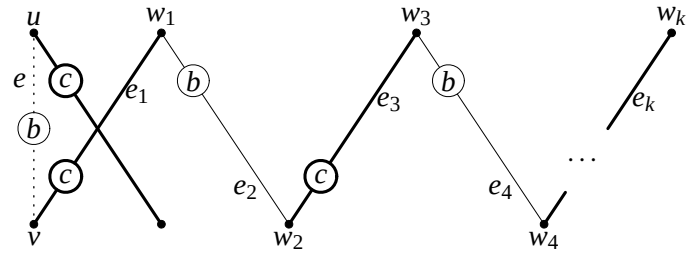


Figure 8.7: Recolouring the path

Let us recolour all the edges of P by swapping the colours b and c, Fig 8.7. By the maximality of P the adjacent edges in H are still coloured by distinct colours and thus we obtain another proper edge colouring of H (Homework 8.9). Since P does not pass through u no edge incident to u was recoloured. Therefore, in this new colouring b is absent both at u and at v, so we can use b to colour the edge e. More precisely, the required edge colouring of G is given by

$$f^*(d) = \begin{cases} b, & d = e \\ b, & d \in E(P) \text{ and } f(d) = c \\ c, & d \in E(P) \text{ and } f(d) = b \\ f(d), & \text{otherwise} \end{cases}$$

which concludes the proof. □

Homework

8.1. Show Lemma 8.2 (b) and (d).

8.2. Show that $\inf(H) \leqslant \inf(G)$ whenever $H \leqslant G$.

†**8.3.** Let G be a graph with $m(G) > 0$. Let v_1 be a vertex with the least positive degree in G and $d_1 = \delta_G(v_1)$, let v_2 be a vertex with the least positive degree in $G - v_1$ and $d_2 = \delta_{G-v_1}(v_2)$, let v_3 be a vertex with the least positive degree in $G - v_1 - v_2$ and $d_3 = \delta_{G-v_1-v_2}(v_3)$, and so on as long as there are edges in $G - v_1 - \ldots - v_j$. We thus obtain a sequence of positive integers $d_1, d_2, \ldots, d_s$. Show that $\inf(G) = \max\{d_1, d_2, \ldots, d_s\}$.

8.4. Give a direct proof of Corollary 8.4. (Hint: take a vertex $v \in V(G)$ such that $\delta(v) = \Delta(G)$ and use induction on $G - v$.)

8.5. Let $k = \chi(G)$ and let $f : V(G) \to \{1, \ldots, k\}$ be a proper vertex colouring of G. For $i \in \{1, \ldots, k\}$ let $V_i = \{v \in V(G) : f(v) = i\}$ be the set of all vertices of G coloured by the colour i. Show that $E(V_i, V_j) \neq \varnothing$ whenever $i \neq j$.

8.6. Prove Corollary 8.7. (Hint: use Theorem 8.3 and the fact that $\delta(G) \leqslant 5$ for every planar graph G.)

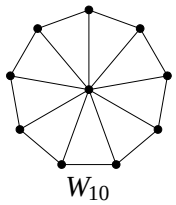
8.7. Complete the proof of Theorem 8.8 by showing that f^* is a proper vertex colouring of G.

8.8. Complete the proof of Theorem 8.13 by showing that the edges in each E_j are independent and that $\{E_0, \ldots, E_{n-2}\}$ is a partition of $E(K_n)$.

8.9. Complete the proof of Theorem 8.14 by showing that swapping the colours along a maximal alternating b/c-path in a bipartite graph produces a proper edge colouring.

Exercises

8.10. Recall that for a graph G by G^* we denote the graph obtained by adding a new vertex to G and joining the new vertex to every vertex of G. Then graph C^*_{n-1} is called the *wheel with n vertices* and denoted by W_n, see the adjacent figure. Find $\chi(W_n)$ and $\chi'(W_n)$ for $n \geqslant 4$.



W_{10}

8.11. Find χ and χ' of: (a) the Petersen graph; (b) Q_n.

8.12. Show that $\chi(G)+\chi(\overline{G}) \leqslant n+1$, where $n = n(G)$.

8.13. Show that $\chi(G)\cdot\chi(\overline{G}) \geqslant n$, where $n = n(G)$.

8.14. Show that $\chi(G)+\chi(\overline{G}) \geqslant 2\sqrt{n}$, where $n = n(G)$.

8.15. Let G be a regular graph with n vertices and let $\delta = \delta(G)$. Show that $\chi(G) \geqslant \dfrac{n}{n-\delta}$.

8.16. Let $e_1, \ldots, e_k$ be k independent edges in K_n. Find $\chi(K_n - e_1 - \ldots - e_k)$.

8.17. Let G be a graph such that $E(G) \neq \varnothing$. Show that there exists a regular graph H such that $\chi(H) = \chi(G)$, $\Delta(H) = \Delta(G)$ and G is an induced subgraph of H.

8.18. Let G be a graph with the property that every pair of odd cycles in G has a common vertex. Show that $\chi(G) \leqslant 5$.

8.19. Let $n \geqslant 4$ be an even integer and let H be a Hamiltonian cycle of K_n. Find $\chi(K_n - E(H))$.

†**8.20.** Let D be a digraph, G its base and $l(D)$ the length of the longest oriented path in D. Show that $l(D) \geqslant \chi(G) - 1$.

8.21. Find $\chi'(G)$ where G is a regular Hamiltonian graph with $\delta(G) = 3$.

8.22. Let G be a regular graph with n vertices where n is odd. Show that $\chi'(G) = \Delta(G) + 1$.

8.23. Let G be a graph such that $E(G) \neq \varnothing$. Show that there exists a regular graph H such that $\chi'(H) = \chi'(G)$, $\Delta(H) = \Delta(G)$ and G is an induced subgraph of H.

8.24. Let $\beta(G)$ denote the greatest cardinality of an independent set of edges of G. Show that $\chi'(G) = \Delta(G) + 1$ if $m(G) > \Delta(G)\cdot\beta(G)$.