3
Introduction to Graphs

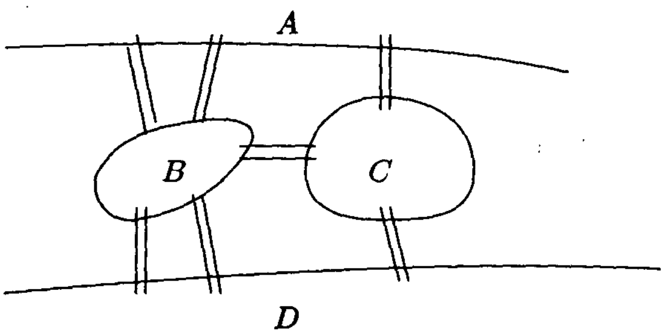
We introduce the idea of a graph via some examples, and concentrate on two types of graph, namely trees and planar graphs. Further graph-theoretic topics will be covered in the next chapter.

3.1 The Concept of a Graph

Example 3.1 (The seven bridges of Königsberg)

In the early eighteenth century there were seven bridges over the River Pregel in the Eastern Prussian town of Königsberg (now Kaliningrad). It is said that the residents tried to set out from home, cross every bridge exactly once and return home. They began to believe the task was impossible, so they asked Euler if it were possible. Euler's proof that it was impossible is often taken to be the beginning of the theory of graphs. What Euler essentially did (although his argument was in words rather than pictures) was to reduce the complexity of Figure 3.1(a) to the simple diagram of 3.1(b), where each land mass is represented by a point (vertex) and each bridge by a line (edge). If the desired walk existed, then each time a vertex was visited by using one edge, then another edge would be used up leaving the vertex; so every vertex would have to have an even number of edges incident with it. Since this is not the case, the desired walk is impossible.

The diagram of Figure 3.1(b) is an example of a graph. It has four vertices and seven edges.



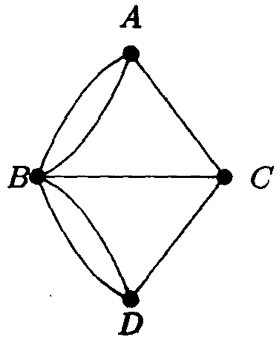


Figure 3.1

Example 3.2 (The utilities problem)

An old problem concerns three houses A, B, C which are to be joined to each of the three utilities, gas, water and electricity, without any two connections crossing each other. In other words, can the diagram of Figure 3.2 be redrawn so that no two lines cross? The diagram is another example of a graph.

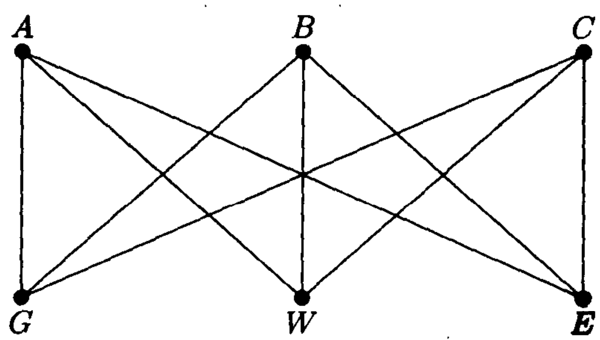


Figure 3.2 The utilities graph

Definition 3.1

A **graph** G consists of a finite set V of **vertices** and a collection E of pairs of vertices called **edges**. The vertices are represented by points, and the edges by lines (not necessarily straight) joining pairs of points. If an edge e joins vertices x and y then x and y are **adjacent** and e is **incident** with both x and y. Any edge joining a vertex x to itself is called a **loop**.

Note that we say E is a collection of pairs, not a set of pairs. This is to allow repeated edges. If two or more edges join the same two vertices, they are called **multiple edges**. For example, the graph of Figure 3.1(b) has two pairs of multiple edges. The graph of the utilities problem is **simple**, i.e. it has no loops or multiple edges.

The number of edges incident with a vertex v in a graph without loops is called the **degree** or **valency** of v and is denoted by $d(v)$. The second name recalls one of the early occurrences of graphs, as drawings of chemical molecules. For example, ethane (C_2H_6) can be represented by the graph of

Figure 3.3, where the two "inside" vertices, of valency 4, represent the two carbon atoms (carbon has valency 4), and the six other vertices, of valency 1, represent hydrogen atoms. Vertices of degree 1 are called **pendant** or **end** vertices.

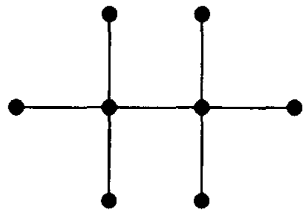


Figure 3.3 Ethane

When a graph contains a loop, the loop is considered to contribute twice to the degree of its incident vertex. This convention enables us to establish the following useful result.

Theorem 3.1

The sum of the degrees of the vertices of a graph is twice the number of edges.

Proof

Each edge contributes twice to the sum of the degrees, once at each end.

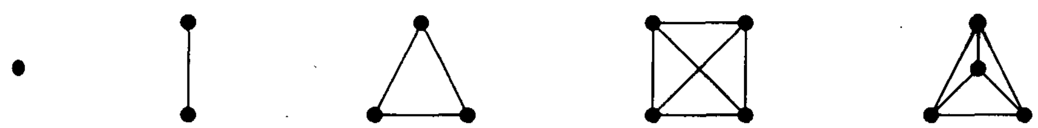
This result is sometimes called the **handshaking lemma**: at a party, the total number of hands shaken is twice the number of handshakes. It has an immediate corollary.

Corollary 3.2

In any graph, the sum of the vertex degrees is even.

Example 3.3

The **complete graph** K_n is the simple graph with n vertices, in which each pair of vertices are adjacent. Since each of the n vertices must have degree $n-1$, the number q of edges must satisfy $2q = n(n-1)$, so that $q = \frac{1}{2}n(n-1)$. This of course is as expected, since q is just the number of ways of choosing two of the n vertices, i.e. $q = \binom{n}{2} = \frac{1}{2}n(n-1)$.



K_1 K_2 K_3 K_4

Figure 3.4

The graphs $K_n, n \leq 4$, are shown in Figure 3.4. The notation K_n is in honour of the Polish mathematician K. Kuratowski (1896–1980) whose important theorem on planarity will be mentioned in Section 3.6. Note that K_4 contains K_3 within it; this idea of one graph being contained in another is formalised in the next definition.

Definition 3.2

A graph H is said to be a **subgraph** of a graph G if the vertex set of H is a subset of the vertex set of G, and the edge set of H is a subset of the edge set of G.

Thus, for example, K_m is a subgraph of K_n wherever $m < n$; simply restrict K_n to m of its vertices.

Finally in this section, we establish some standard notation. From now on, we shall use p and q to denote the numbers of vertices and edges respectively, and by a (p, q)-graph we shall mean a graph with p vertices and q edges. Thus, for example, K_4 is a $(4, 6)$-graph.

3.2 Paths in Graphs

Many important applications of graph theory involve travelling round the graph, in the sense of moving from vertex to vertex along incident edges. We make some definitions related to this idea.

Definition 3.3

A **walk** in a graph G is a sequence of edges of the form

$$v_0v_1,\ v_1v_2,\ v_2v_3, \ldots, v_{n-1}v_n.$$

This walk is sometimes, in a simple graph, represented more compactly by $v_0 \to v_1 \to v_2 \to \cdots \to v_n$. Note that there is an implied direction to the walk. v_0 is called the **initial** vertex and v_n the **final** vertex of the walk; the number (n) of edges is called the **length** of the walk.

A walk in which all the edges are distinct is called a **trail**. A trail in which all vertices $v_0, \ldots, v_n$ are distinct (except possibly $v_n = v_0$) is called a **path**; a path $v_0 \to \ldots \to v_n$ with $v_n = v_0$ is called a **cycle**.

Example 3.4

In the graph of Figure 3.5,

$z \to u \to y \to v \to u$	is a trail but not a path;
$u \to y \to w \to v$	is a path of length 3;
$u \to y \to w \to v \to u$	is a cycle of length 4.

It seems natural to consider the cycles $u \to y \to v \to u$ and $y \to v \to u \to y$

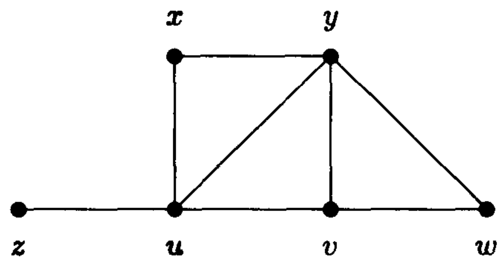


Figure 3.5

to be the same; so often we identify a cycle with the set of its edges. We use the notation (for $n > 1$)

$$C_n = \text{cycle of length } n \text{ (i.e. with } n \text{ edges and vertices);}$$

$$P_n = \text{path of length } n-1 \text{ (i.e. with } n \text{ vertices).}$$

Thus, for example, $P_2 = K_2$ and $C_3 = K_3$.

Definition 3.4

A graph is **connected** if, for each pair x, y of vertices, there is a path from x to y. A graph which is not connected is made up of a number of connected pieces, called **components**.

3.3 Trees

Definition 3.5

A tree is a connected simple graph with no cycles.
For example, the ethane graph in Figure 3.3 is a tree, as is each P_n. Note that the ethane graph has $p = 8$ and $q = 7$, while P_n has $p = n$ and $q = n - 1$; in each case, $p - q = 1$. This property in fact characterises those connected graphs which are trees. Our proof of this depends upon the following useful result.

Theorem 3.3

If T is a tree with $p \geq 2$ vertices then T contains at least two pendant vertices.

Proof

Since T has p vertices, all paths in T must have length less than p. So there must be a longest path in T, say $v_1 \to v_2 \to \ldots \to v_r$. We claim that v_1 and

v_r both have degree 1. Suppose v_1 has degree > 1; then there is another edge from v_1, say v_1v_0, where v_0 is none of $v_2, \ldots, v_r$ (otherwise there would be a cycle), so $v_0 \to v_1 \to \ldots \to v_r$ would be a longer path. So v_1 has degree 1, and a similar argument holds for v_r.

Theorem 3.4

Let T be a simple graph with p vertices. Then the following statements are equivalent:

(i) T is a tree;

(ii) T has $p - 1$ edges and no cycles;

(iii) T has $p - 1$ edges and is connected.

Proof

(i) $\Rightarrow$ (ii) We have to show that all trees with p vertices have $p - 1$ edges. This is certainly true when $p = 1$. Suppose it is true for all trees with $k \geq 1$ vertices, and let T be a tree with $k + 1$ vertices. Then, by Theorem 3.3, T has an end vertex w. Remove w and its incident edge from T to obtain a tree T' with k vertices. By the induction hypothesis, T' has $k - 1$ edges; so T has $(k - 1) + 1 = k$ edges as required.

(ii) $\Rightarrow$ (iii) Suppose T has $p - 1$ edges and no cycles, and suppose it consists of $t \geq 1$ components, $T_1, \ldots, T_t$, each of which has no cycles and hence must be a tree. Let p_i denote the number of vertices in T_i. Then $\sum_i p_i = p$, and the number of edges in T is $\sum_i (p_i - 1) = p - t$. So $p - t = p - 1$, i.e. $t = 1$, so that T is connected.

(iii) $\Rightarrow$ (i) Suppose T is connected with $p - 1$ edges, but is not a tree. Then T must have a cycle. Removing an edge from a cycle does not destroy connectedness, so we can remove edges from cycles until no cycles are left, preserving connectedness. The resulting graph must be a tree, with p vertices and $q < p-1$ edges, contradicting (ii).

This theorem can be used to establish the tree-like nature of certain chemical molecules.

Example 3.5

Show that the alkanes (paraffins) C_nH_{2n+2} have tree-like molecules.

Solution

Each molecule is represented by a graph with $n + (2n + 2) = 3n + 2$ vertices. Of these, n have degree 4 and $2n + 2$ have degree 1, so, by Theorem 3.1,

$$2q = 4n + 2n + 2 = 6n + 2$$

whence $q = 3n + 1 = p - 1$. Since molecules are connected, the graphs must be trees, by Theorem 3.4.

The first few alkanes are shown in Figure 3.6.

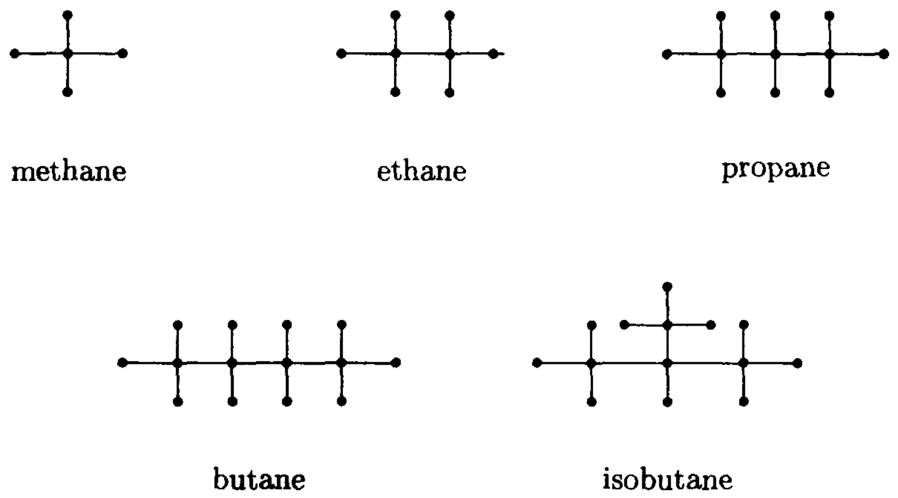


Figure 3.6 Alkanes

Note that there are two "different" trees corresponding to C_4H_{10}.

Definition 3.6

Two graphs G_1, G_2 are **isomorphic** if it is possible to label the vertices of both graphs by the same labels, so that, for each pair u, v of labels, the number of edges joining vertices u and v in G_1 is equal to the number of edges joining u and v in G_2.

Example 3.6

(i) The graphs portrayed by the last two diagrams in Figure 3.4 are isomorphic.
(ii) The butane and isobutane graphs (Figure 3.6) are not isomorphic. The second graph has one vertex of degree 4 joined to all the other vertices of degree 4, but this does not happen in the first graph.

Tree diagrams such as those in Figure 3.6 were introduced in 1864 by the chemist A. Crum Brown in his study of isomerism, the occurrence of molecules with the same chemical formula but different chemical properties. The problem of enumerating the non-isomorphic molecules C_nH_{2n+2} was eventually solved by Cayley in 1875, but his solution is beyond the scope of this book.

A related problem was: find $T(n)$, the number of non-isomorphic trees with n vertices. We have $T(1) = T(2) = T(3) = 1$, and, as the reader should check, $T(4) = 2, T(5) = 3, T(6) = 6$. No simple formula for $T(n)$ exists, although $T(n)$ is the coefficient of x^n in a known but very complicated series. However, there does exist a very nice formula for the number of trees on n given labelled vertices. For example, although $T(3) = 1$, there are three **labelled** trees with vertices labelled 1, 2, 3 as shown in Figure 3.7. It was established by Cayley in 1889 that the number of labelled trees on n vertices is n^{n-2}. A proof of this will be given in Chapter 6.

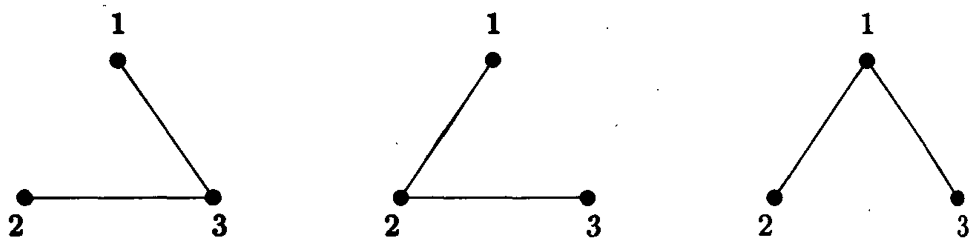


Figure 3.7 Labelled trees.

3.4 Spanning Trees

Suppose that a connected graph represents a railway system, the vertices representing the towns and the edges the railtracks. Suppose also that the government wishes to get rid of as much track as possible, nevertheless retaining a rail system which connects all the towns. What is required is a tree which is a subgraph of the given graph, containing all the vertices.

Definition 3.7

A **spanning tree** of a connected graph G is a tree which contains all the vertices of G and which is a subgraph of G.

Example 3.7

(i) K_3 has three spanning trees, as shown in Figure 3.7.

(ii) K_4 has $16 = 4^2$ spanning trees. Draw them. Do you see how this relates to Cayley's 1889 result?

(iii) In the graph of Figure 3.5, the edges zu, xu, uy, yv, yw form a spanning tree.

In the case of a **weighted** graph G, i.e. when each edge e of G has a weight $w(e)$ assigned to it, where $w(e)$ is a positive number such as the length of e, then it may be desired to find a spanning tree of smallest possible total weight. There are several different algorithms which find such a minimum weight spanning tree of G.

The greedy algorithm

This is often known as Kruskal's algorithm.

Procedure

(i) Choose an edge of smallest weight.

(ii) At each stage, choose from the edges not yet chosen the edge of smallest weight whose inclusion will not create a cycle.

iii) Continue until a spanning tree is obtained.

(If the given graph has p vertices, the algorithm will terminate after $p-1$ edges have been chosen.)

Example 3.8

Apply the greedy algorithm to the graph of Figure 3.8.

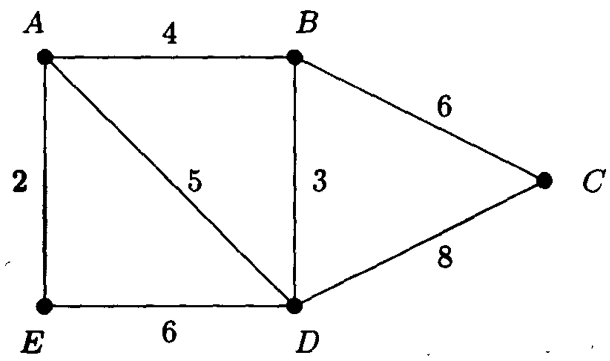


Figure 3.8

Solution

First choose AE (weight 2). Then choose $BD(3)$, then $AB(4)$. We cannot now choose $AD(5)$ since its inclusion would create a cycle $ABDA$. Similarly we cannot choose DE. So choose $BC(6)$. The edges AE, AB, BD, BC then form a minimum weight spanning tree of weight $2 + 3 + 4 + 6 = 15$.

Justification of the greedy algorithm

Suppose that the greedy algorithm produces a tree T, but that there is another spanning tree U which has smaller weight than T. Since $T \neq U$, and both have the same number of edges, there must be an edge in T not in U: let e be such an edge of minimum weight. The addition of e to U must create a cycle C, and this cycle must contain an edge e' which is not in T. Now $w(e') \geq w(e)$, since if $w(e') < w(e)$ then e' would have been chosen by the greedy algorithm rather than e. So if we remove e' from C we obtain a spanning tree V such that $w(V) \leq w(U)$, and V has one more edge in common with T than U had. By repeating this process we eventually change U into T, one edge at a time, and conclude that $w(T) \leq w(U) < w(T)$, a contradiction. So no such U can exist.

The greedy algorithm is so called because it greedily minimises the weight at each step, ignoring possible future complications; fortunately it gets away with this strategy. The disadvantage of the algorithm, however, lies in the difficulty of determining at each stage whether or not a cycle would be created by the inclusion of the smallest weighted edge available (this is particularly true when the graph is large). This problem can be overcome by using a slightly different algorithm, due to Prim (1957). In Prim's algorithm, the graph constructed is connected (and hence a tree) at each stage of the construction (unlike the greedy algorithm, which chose BD immediately after AE in the above example), and at each stage the smallest weight edge is sought which joins the existing tree to a vertex not in the tree. Clearly the inclusion of this edge cannot create a cycle.

Prim's algorithm

(i) Select any vertex, and choose the edge of smallest weight from it.

(ii) At each stage, choose the edge of smallest weight joining a vertex already included to a vertex not yet included.

(iii) Continue until all vertices are included.

Example 3.8 (revisited)

Use Prim's algorithm starting at B. Choose $BD(3)$, then $BA(4)$, then $AE(2)$, then $BC(6)$ to obtain the same spanning tree as before.

A third algorithm operates by **removing** edges from the given graph, destroying cycles, until a spanning tree is left. At each stage remove the largest-weighted edge whose removal does not disconnect the graph. In Example 3.8, we could remove DC, then DE, then AD. Clearly this approach would be quicker than the others if the graph has "few" edges.

3.5 Bipartite Graphs

Definition 3.8

A graph is **bipartite** if its vertex set V can be partitioned into two sets B, W in such a way that every edge of the graph joins a vertex in B to a vertex in W. The partition $V = B \cup W$ is called a **bipartition** of the vertex set.

Example 3.9

Labellings show that the graphs in Figure 3.9 are bipartite. In both graphs, each edge joins a B to a W.

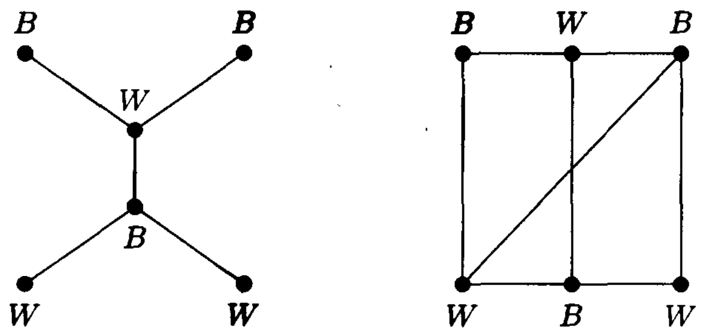


Figure 3.9 Bipartite graphs

If we interpret B and W as black and white, we see that a graph is bipartite precisely when the vertices can be coloured using two colours so that no edge joins two vertices of the same colour. For this reason, bipartite graphs are sometimes called **bichromatic**.

Example 3.10

The cycle C_n is bipartite if and only if n is even.

Theorem 3.5

A connected graph is bipartite if and only if it contains no cycle of odd length.

Proof

If a graph G contains an odd cycle (i.e. a cycle of odd length) then it cannot possibly be bipartite. So suppose now that G contains no odd cycle; we shall show how to colour its vertices B and W.

Choose any vertex v of G, and partition V as $B \cup W$ where

$$B = \{u \in V : \text{ shortest path from } v \text{ to } u \text{ has even length } \},$$

$$W = \{u \in V : \text{ shortest path from } v \text{ to } u \text{ has odd length } \}.$$

We have $u \in B$ since 0 is even; we have to check that no edge of G has both ends in B or both ends in W.

Suppose there is an edge xy with $x \in B$ and $y \in B$. Then, denoting the length of the shortest path from vertex v_1 to vertex v_2 by $d(v_1, v_2)$, we have $d(v, x) = 2m$ and $d(v, y) = 2n$ for some integers m, n. But there is a walk from v to y via x of length $2m + 1$, so $2n \leq 2m + 1$. Similarly $2m \leq 2n + 1$, so $m = n$.

Denote the shortest paths from v to x and y by $P(x)$ and $P(y)$ respectively. Then, since $m = n$, both $P(x)$ and $P(y)$ have equal lengths. Let w be the **last** vertex on $P(x)$ which is also on $P(y)$ (possibly $w = v$). Then the part of $P(x)$

from w to x and the part of $P(y)$ from w to y must be of equal length, and, since they have only w in common, they must, with edge xy, form an odd cycle. But G has no odd cycles, so the assumption of the existence of the edge xy must be false. So there is no edge with both edges in B; similarly there is no edge with both edges in W.

Corollary 3.6

All trees are bipartite.

Definition 3.9 (Complete bipartite graphs)

A simple bipartite graph with vertex set $V = B \cup W$ is **complete** if every vertex in B is joined to every vertex in W. If $|B| = m$ and $|W| = n$, the graph is denoted by $K_{m,n}$ or by $K_{n,m}$. For example, the utilities graph of Figure 3.2 is $K_{3,3}$, and the methane graph of Figure 3.6 is $K_{1,4}$.

Clearly, $K_{m,n}$ has $m+n$ vertices and mn edges; m of the vertices have degree n, and n of the vertices have degree m.

The complete graphs K_n and the complete bipartite graphs $K_{m,n}$ play important roles in graph theory, particularly in the study of planarity to which we now turn.

3.6 Planarity

A graph is **planar** if it can be drawn in the plane with no edges crossing. The concept of planarity has already appeared in the utilities problem, which can be restated as: is $K_{3,3}$ planar? If a graph is planar, then any drawing of it with no edges crossing is called a **plane** graph. For example, K_4 is planar, as was shown in Figure 3.4; the second drawing of K_4 there was a plane graph, establishing its planarity.

Planar graphs occur naturally in the **four-colour problem**. In colouring a map, it is standard procedure to give adjacent countries different colours. It became apparent that four colour always seemed to be sufficient to colour any map, and a general proof of this statement was attempted by A.B. Kempe in 1879. Ten years later, Heawood discovered that Kempe's "proof" was flawed, and instead of the four-colour theorem we had the four-colour conjecture. Eventually, in 1976, the truth of the conjecture was established by two mathematicians, K. Appel and W. Haken; as the postmark of the University of Illinois asserted, "four colours suffice".

The problem of colouring a map can be transformed into one of colouring the vertices of a planar graph. Given a map, we can represent each region by a vertex, and join two vertices by an edge precisely when the corresponding regions share a common boundary. For example, Figure 3.10 shows a map and a planar graph representing it. So the problem reduces to that of colouring the vertices

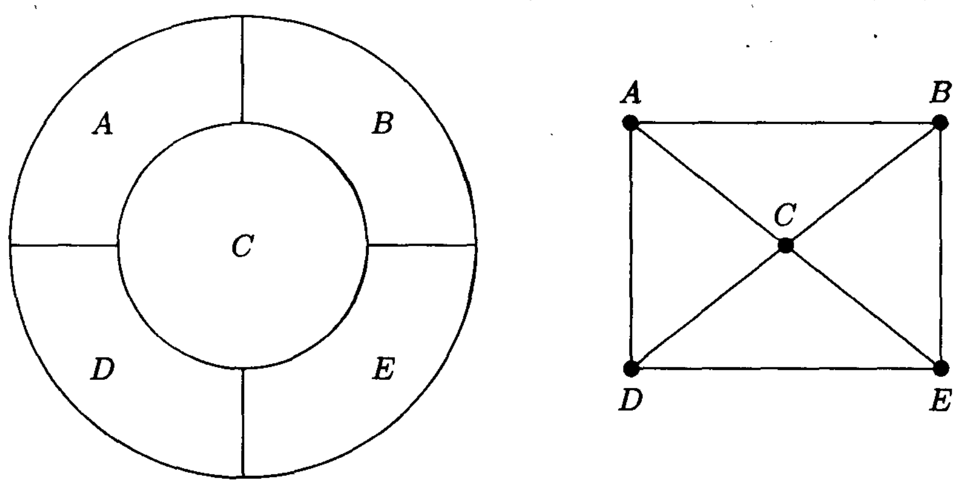


Figure 3.10

of a planar graph with four colours, so that no two adjacent vertices receive the same colour. Colourings of graphs will be discussed further in Chapter 5.

Any plane graph clearly divides the plane into disjoint regions, one of which is infinite. The basic result about plane graphs is known as Euler's formula; Euler initially studied it in the context of polyhedra, and we shall look at this in the next section.

Theorem 3.7 (Euler's formula)

Any connected plane (p, q)-graph divides the plane into r regions, where

$$p - q + r = 2.$$

Proof

If there is a cycle, remove one edge from it. The effect is to reduce q and r by 1 (since two regions are amalgamated into one), and to leave p unchanged. So the resulting graph has $p' = p, q' = q - 1, r' = r - 1$, where $p' - q' + r' = p - q + r$. Repeat this process until no cycles remain. The final graph must be a tree, with $p'' - q'' + r'' = p - (p - 1) + 1 = 2$.

Example 3.11

The plane graph in **Figure 3.10** has

$$p - q + r = 5 - 8 + 5 = 2.$$

There are four finite regions and one infinite region.

We now define the **degree** of a region of a plane graph to be the number of encounters with edges in a walk round the boundary of the region.

Example 3.12

In Figure 3.11 regions 3 and 4 have degree 3, the infinite region 1 has degree

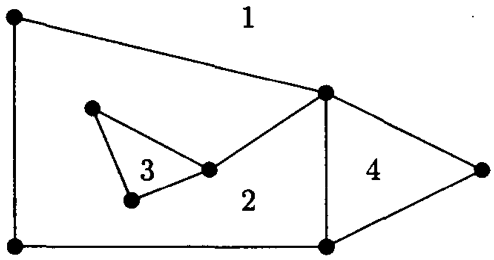


Figure 3.11

5, and region 2 has degree 9 (note that one edge is encountered twice, once on each side).

Parallel to the handshaking lemma we have:

Theorem 3.8

In a connected plane graph, $2q =$ sum of degrees of the regions.

Theorem 3.9

K_n is planar only if $n \leq 4$.

Proof

It is enough to show that K_5 is non-planar. (Why?) Now K_5 has $p = 5, q = 10$, so if a plane drawing of K_5 exists it must have $r = 2 - 5 + 10 = 7$ regions. Each of the seven regions must have degree ≥ 3, so, by Theorem 3.8, $20 = 2q \geq 7 \times 3 = 21$, a contradiction.

Theorem 3.10

$K_{3,3}$ is not planar.

Proof

$K_{3,3}$ has $p = 6$ and $q = 9$, so if a plane drawing exists it must have $r = 2 - 6 + 9 = 5$ regions. Since $K_{3,3}$ is bipartite, with no odd cycles, each region must have degree ≥ 4, so we must have $18 = 2q \geq 4 \times 5 = 20$, a contradiction.

Corollary 3.11

$K_{m,n}$ is planar $\Leftrightarrow$ $\min(m,n) \leq 2$.

The technique of counting the sum of the degrees of the regions is a useful one. We can apply it to the famous Petersen graph, shown in Figure 3.12. (See Section 4.1 and Exercise 5.17 for more about this graph.)

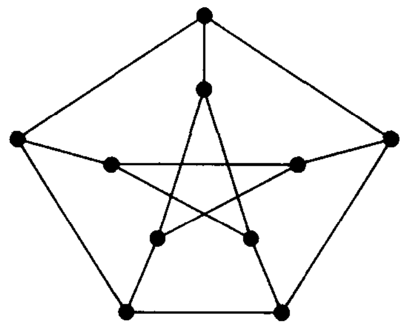


Figure 3.12 The Petersen graph

Example 3.13

The Petersen graph is not planar.

Solution

Suppose a plane drawing exists. Since $p = 10$ and $q = 15$, we would have $r = 2 - 10 + 15 = 7$. Now the shortest cycle in the graph clearly has length 5, so every region must have degree ≥ 5. So we would have a contradiction, $30 = 2q \geq 7 \times 5 = 35$.

Kuratowski's theorem

What makes a graph non-planar? Clearly, if it contains K_5 or $K_{3,3}$ as a subgraph, then it cannot possibly be planar. It was proved in 1930 by the Polish mathematician Kuratowski that, essentially, it is only the presence of a K_5 or a $K_{3,3}$ within a graph that can stop it being planar.

To clarify this statement, we first make the following observation. Since K_5 is not planar, the graph shown in Figure 3.13 cannot be planar either. For if it were, we could make a plane drawing of it, erase b from the edge ac, and obtain a plane drawing of K_5. Inserting a new vertex into an existing edge of a graph is called **subdividing** the edge, and one or more subdivision of edges creates a **subdivision** of the original graph.

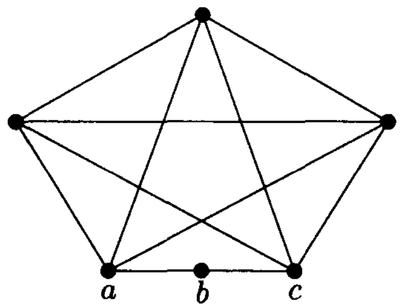


Figure 3.13

Theorem 3.12 (Kuratowski's theorem)

A graph is planar if and only if it does not contain a subdivision of K_5 or $K_{3,3}$ as a subgraph.

The proof of this deep topological result is beyond the scope of this book. But we exhibit the result's usefulness by using it to prove that the Petersen graph is non-planar.

Example 3.13 (again)

In Figure 3.14, Petersen's graph is on the left. On the right is the same graph with two edges removed. This subgraph is a subdivision of $K_{3,3}$ as shown by the labelling of the vertices.

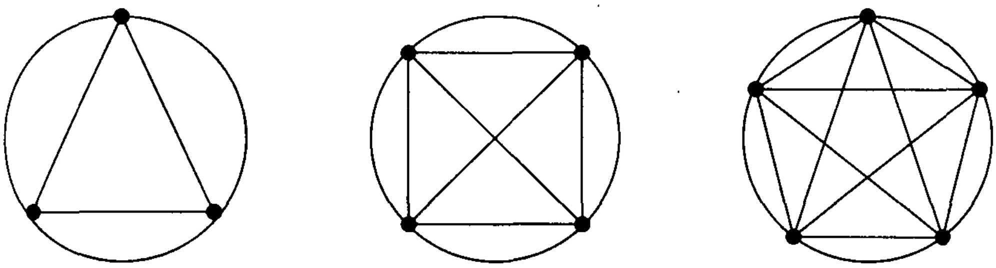
Another test for planarity will be given in Section 4.2.

Figure 3.14

Chords of a circle

We close this section on planarity with an application of Euler's formula to a well-known problem concerning chords of a circle.

Suppose we have n points spaced round a circle, and we join each pair of points by a chord, taking care to ensure that no three chords intersect at the same point. Into how many regions is the interior of the circle divided? The cases $n = 3, 4, 5$ are shown in Figure 3.15. It would appear that $n = 6$ should



$n = 3$ 4 regions $n = 4$ 8 regions $n = 5$ 16 regions

Figure 3.15

give 32 regions. But it does not! (Check!)

Suppose we have n points and have drawn the $\binom{n}{2}$ chords. There will be n regions with a circular arc as a boundary - let's lay them aside and concentrate on the remaining regions. Turn the geometrical picture into a graph by putting a vertex at each of the n given points, and at each crossing point of chords. How many crossing points are there? There is one for each pair of chords which cross. But any pair of crossing chords is obtained by choosing 4 of the given n points and drawing the "cross" chords between them; so there must be $\binom{n}{4}$ crossing points. So the resulting graph has $p = n + \binom{n}{4}$ vertices. Each of the n original vertices has degree $n - 1$, and each of the new $\binom{n}{4}$ vertices has degree 4. So by the handshaking lemma

$$2q = n(n-1) + 4\binom{n}{4}, \qquad \text{i.e. } q = \binom{n}{2} + 2\binom{n}{4}.$$

Thus

$$\begin{aligned} r &= 2 - p + q \\ &= 2 - n - \tbinom{n}{4} + \tbinom{n}{2} + 2\tbinom{n}{4} \\ &= 2 - n + \tbinom{n}{2} + \tbinom{n}{4}. \end{aligned}$$

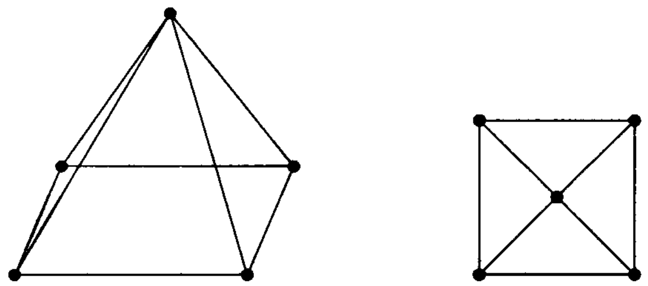
Here r includes a count of 1 for an infinite region, so there are $1 - n + \binom{n}{2} + \binom{n}{4}$ finite regions. We have to add the n boundary regions which we put aside earlier, so finally the number of regions is

$$1 + \binom{n}{2} + \binom{n}{4}.$$

Check that this gives $4, 8, 16$ for $n = 3, 4, 5$, and 31 for $n = 6$.

3.7 Polyhedra

A polyhedron is a solid bounded by a finite number of faces, each of which is polygonal. For example, the pyramid in Figure 3.16(a) is a polyhedron with five vertices, five faces (four triangular, and one square base), and eight edges.



(a) (b)

Figure 3.16 A pyramid and its plane graph

As was mentioned earlier, Euler's formula arose first in the study of polyhedra, relating the numbers of vertices, faces and edges in a convex polyhedron. (A polyhedron is convex if the straight line segment joining any two of its vertices lies entirely within it.) Such a polyhedron can be represented by a plane graph, obtained by projecting the polyhedron into a plane. The graph in Figure 3.16(b) represents the pyramid; think of the internal vertex as the top of the pyramid, and think of the base of the pyramid as being represented by the infinite region (of degree 4).

The cube is an example of a **regular** polyhedron. A polyhedron is regular if there exist integers $m \geq 3, n \geq 3$ such that each vertex has m faces (or m edges) meeting at it, and each face has n edges on its boundary. For a cube, $m = 3$ and $n = 4$. Convex regular polyhedra are known as **Platonic** solids; they were discussed at great length by the ancient Greeks who knew that there were only 5 such solids. In the next theorem we use the terminology of graphs, moving from a polyhedron to its corresponding graph.

Theorem 3.13

Suppose that a regular polyhedron has each vertex of degree m and each face of degree n. Then (m, n) is one of $(3,3), (3,4), (4,3), (3,5), (5,3)$.

Further, there exist Platonic solids corresponding to each of these pairs.

Proof

We have $p - q + r = 2$, where

$$2q = \text{ sum of vertex degrees } = mp$$

$$\text{and } 2q = \text{ sum of face degrees } = nr.$$

So $(\frac{2}{m} - 1 + \frac{2}{n})q = 2$, whence

$$(2m + 2n - mn)q = 2mn. \tag{3.1}$$

Thus, trivially, $2m+2n-mn > 0$, i.e. $(m-2)(n-2) < 4$. So $(m-2)(n-2) = 1, 2$ or 3, and the five possibilities arise.

For each possible pair (m, n), we can find q from (3.1) and then deduce the values of p and r. We tabulate these values in Table 3.1, and give the name of the corresponding Platonic solid.

Table 3.1

m	n	q	p	r	Name
3	3	6	4	4	tetrahedron
3	4	12	8	6	cube
4	3	12	6	8	octahedron
3	5	30	20	12	dodecahedron
5	3	30	12	20	icosahedron

Note that the names reflect the number r of faces. The five solids, and their plane graphs, are shown in Figure 3.17.

As well as the five regular polyhedra just discussed, there exist the **semiregular** polyhedra known as the **Archimedean solids**. Although they may well have been known to the Greeks, the first known listing of them is due to Kepler in 1619. These solids have more than one type of face, but they have the property that each vertex has the same pattern of faces around it. For example, the truncated cube, obtained by slicing off each of the eight vertices, has eight triangular faces and six octagonal faces, and, at each vertex, two octagons and one triangle meet.

Example 3.14

A polyhedron is made up of pentagons and hexagons, with three faces meeting at each vertex. Show that there must be exactly 12 pentagonal faces.

Solution

We have $p - q + r = 2$ and $2q$ = sum of vertex degrees $= 3p$. Thus $2q = 6r - 12$. Now suppose there are x pentagonal and y hexagonal faces. Then $r = x + y$ and $2q$ = sum of degrees of faces $= 5x + 6y$. Substituting into $2q = 6r - 12$ gives $5x + 6y = 6x + 6y - 12$, whence $x = 12$.

The case $x = 12, y = 0$ corresponds of course to a dodecahedron. The case $x = 12, y = 20$ corresponds to the pattern often seen on a soccer ball. The corresponding Archimedean solid is a **truncated icosahedron**; the reader

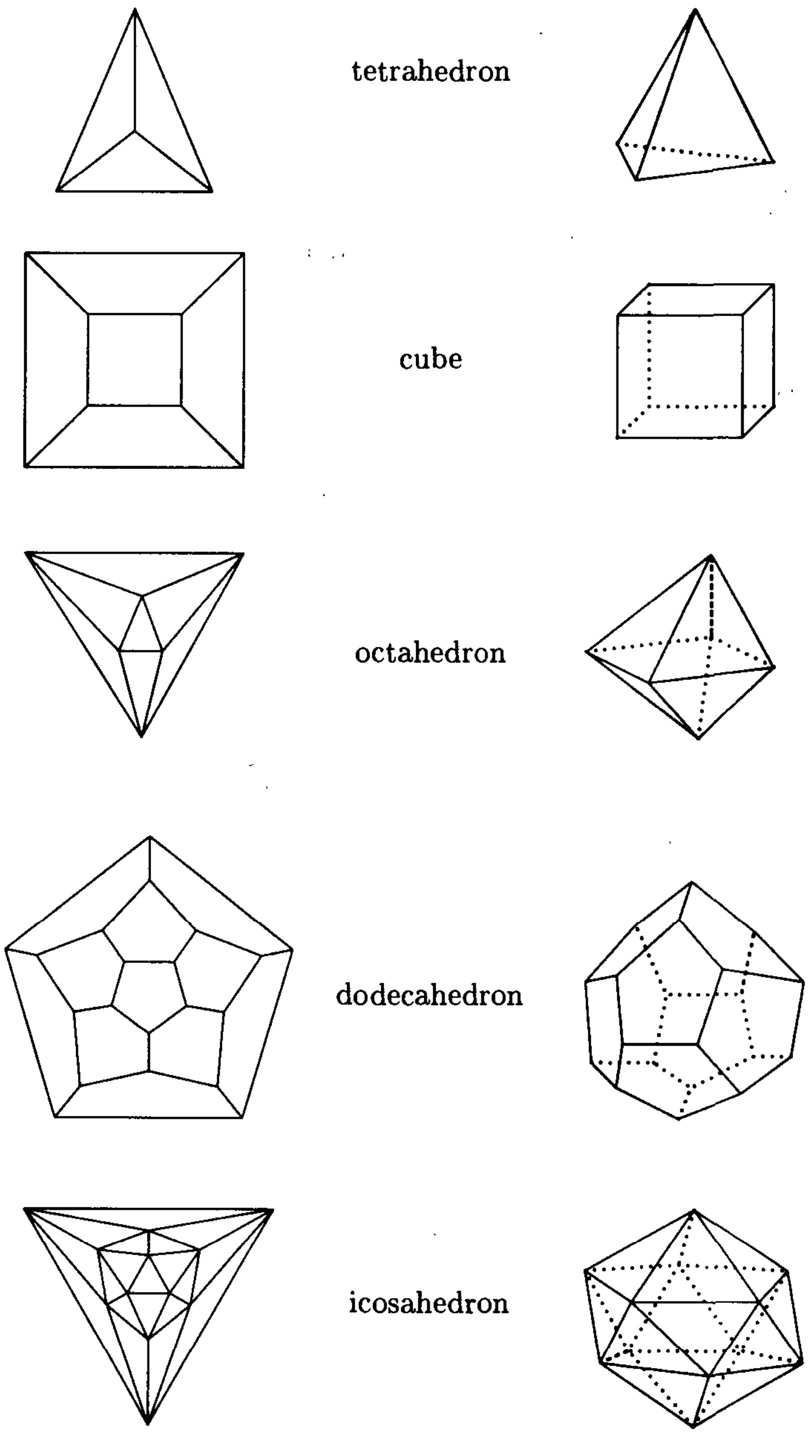


Figure 3.17

should be able to see how to obtain one by slicing vertices off an icosahedron. This solid aroused great interest in the 1990s when it was discovered that a third form of carbon existed (as well as diamond and graphite).

This form is denoted by C_{60}; the molecular structure is that of 60 carbon atoms situated at the vertices of a truncated icosahedron. The discoverers of this molecule called it **Buckminsterfullerine** (it is commonly known as a Buckyball) since they considered it as similar to a geodesic dome created by the architect R. Buckminster Fuller. But, as we have pointed out, it has been known to mathematicians for a long time.

The graphite form of carbon has the carbon atoms arranged in a flat honeycomb pattern of hexagons. Hexagons tile the plane, so need the addition of n-gons with $n < 6$ to enable a 3-dimensional form to take place. It turns out that 12 pentagons are just right to enable a complete closing up to take place. See Exercise 3.14 for the corresponding problem when pentagons are replaced by squares.

There are other fullerine molecules, such as C_{70} which has 12 pentagons and 25 hexagons; its shape is more like a rugby ball.

Exercises

Exercise 3.1

Prove that the number of vertices of odd degree in a given graph is even.

Exercise 3.2

Show that all alcohols $C_nH_{2n+1}OH$ have tree-like molecules. (The valencies of C, O, H are $4, 2, 1$ respectively.)

Exercise 3.3

Show that if G is a simple graph with p vertices, where each vertex has degree $\geq \frac{1}{2}(p-1)$, then G must be connected. (Hint: how many vertices must each component have?)

Exercise 3.4

How many spanning trees do the graphs in Figure 3.18 have?

(a) (b)

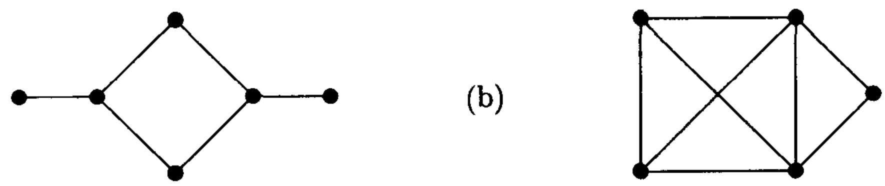


Figure 3.18

Exercise 3.5

How many edges must be removed from a connected (p, q)-graph to obtain a spanning tree?

Exercise 3.6

Let $K_{2,3}$ have bipartition $B \cup W$ where $B = \{a, b\}, W = \{x_1, x_2, x_3\}$.
(a) Explain why, in a spanning tree of $K_{2,3}$, there must be precisely one of the vertices x_i joined to both a and b.
(b) How many spanning trees does $K_{2,3}$ have?
(c) How many spanning trees does $K_{2,100}$ have?

Exercise 3.7

Use (a) the greedy algorithm, (b) Prim's algorithm to find a minimum weight spanning tree in the graph shown in Figure 3.19.

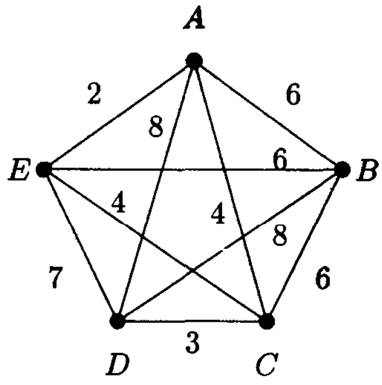


Figure 3.19

Exercise 3.8

The distances between 5 Lanarkshire towns are given in Table 3.2. Find the shortest length of a connecting road network.

Table 3.2

	G	H	A	M	EK
Glasgow	0	10	11	13	9
Hamilton	10	0	8	3	6
Airdrie	11	8	0	8	13
Motherwell	13	3	8	0	8
East Kilbride	9	6	13	8	0

Exercise 3.9

Pan Caledonian Airways (PCA) operates between 12 towns whose coordinates referred to a certain grid are $(0,2),(0,5),(1,0),(1,4),(2,3),(2,4),$ $(3,2),(3,5),(4,4),(4,5),(5,3),(6,1)$. What is the minimum number of flights necessary so that travel by PCA is possible between any two of the towns? Find the minimum total length of such a network of flights.

Exercise 3.10

Determine which of the graphs in Figure 3.20 are planar.

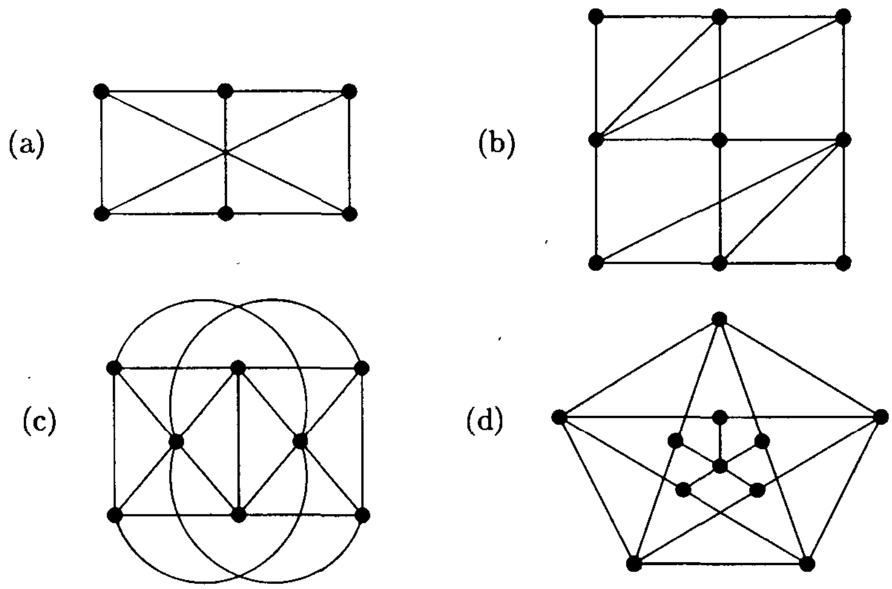


Figure 3.20

Exercise 3.11

A **complete matching** of a graph with $2n$ vertices is a subgraph consisting of n disjoint edges. How many different complete matchings are there in the graph of Figure 3.20(a)?

Exercise 3.12

The graph $G_n (n \geq 1)$ is shown in Figure 3.21.

(a) Is G_n (i) bipartite? (ii) planar?

(b) Let a_n denote the number of complete matchings of G_n. Show that

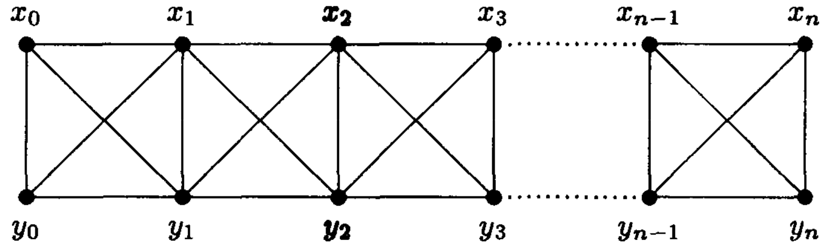


Figure 3.21

$a_1 = 3$ and $a_2 = 5$. Show that $a_n = a_{n-1} + 2a_{n-2}$ $(n \geq 3)$ and hence obtain a formula for a_n.

Exercise 3.13

(a) Show that if G is a simple planar (p, q) graph, $p \geq 3$, then $q \leq 3p-6$. Deduce that K_5 is not planar.

(b) Show that if G is a simple planar (p, q) graph, $p \geq g$, where g is the **girth** of G, i.e. the length of the shortest cycle in G, then $q \leq \frac{g}{g-2}(p-2)$.

(c) Deduce from (b) that $K_{3,3}$ and the Petersen graph are both nonplanar.

Exercise 3.14

A convex polyhedron has only square and hexagonal faces. Three faces meet at each vertex. Use Euler's formula to show that there must be exactly six square faces. The cube has no hexagonal faces: give an example with six square faces and at least one hexagonal face. (Try truncating an octahedron.)

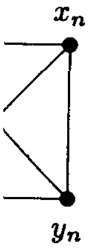
Exercise 3.15

Suppose n cuts are made across a pizza. Let p_n denote the maximum number of pieces which can result (this happens when no two cuts are parallel or meet outside the pizza, and no three are concurrent).

Prove that $p_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2}$.

Exercise 3.16

Let h_n denote the number of spanning trees in the fan graph shown in Figure 3.22. Verify that $h_1 = 1, h_2 = 3, h_3 = 8$.



) and hence

ı $q \leq 3p-6$.

g, where g
in G, then

oth nonpla-

Three faces
must be ex-
an example
truncating

: maximum
vo cuts are
:nt).

h shown in

Find a recurrence relation for h_n and hence show

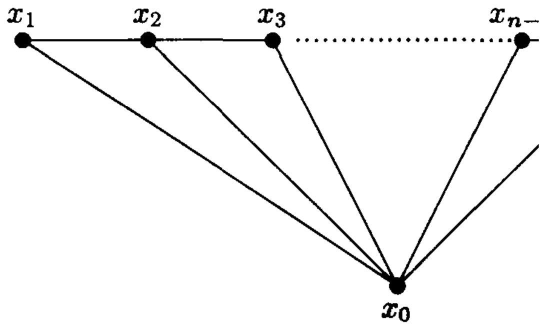


Figure 3.22

4
Travelling Round a Graph

In this chapter we consider various problems relating to the existence of certain types of walk in a graph. The reader should recall the definitions of walk, path, cycle and trail given in Section 3.2. The Königsberg bridge problem concerns the existence of a trail which is closed and contains all the edges of the graph. We study such (Eulerian) trails in more detail, but first we look at a related type of problem associated with the name of the Irish mathematician Sir William Rowan Hamilton (1805–1865).

4.1 Hamiltonian Graphs

The dodecahedron is shown at the end of Chapter 3. Hamilton posed the problem: is it possible to start at one of the 20 vertices, and, by following edges, visit every other vertex exactly once before returning to the starting point? In other words: is there a cycle through all the vertices? You should have no problem finding such a cycle (turn to Figure 4.3 if you get stuck), so it is perhaps not surprising that the commercial exploitation of this problem as a game was not a financial success.

Definition 4.1

A **hamiltonian cycle** in a graph G is a cycle containing all the vertices of G. A **hamiltonian graph** is a graph containing a hamiltonian cycle.

The name hamiltonian is, as often happens in mathematics, not entirely just, since others such as Kirkman had studied the idea before Hamilton.

Example 4.1

(a) The octahedral graph is hamiltonian: in Figure 4.1(a) take the hamiltonian cycle 1234561.

(b) The graph of Figure 4.1(b) is not hamiltonian. The easiest way to see this is to note that it has 9 vertices so that, if it is hamiltonian, it must contain a cycle of length 9. But, being a bipartite graph, it contains only cycles of even length.

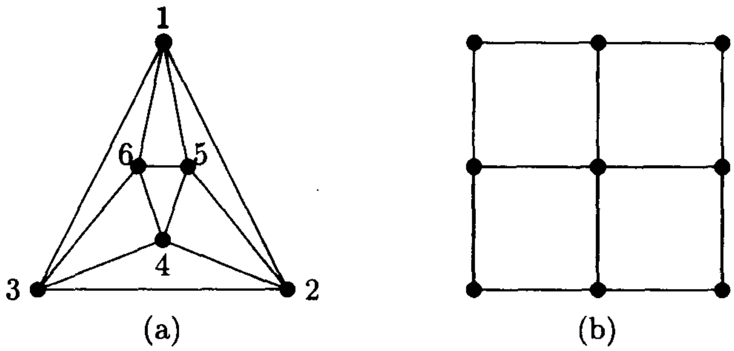


Figure 4.1

Theorem 4.1

A bipartite graph with an odd number of vertices cannot be hamiltonian.

Example 4.2

(a) K_n is hamiltonian for all $n \geq 3$.

(b) $K_{m,n}$ is hamiltonian if and only if $m = n \geq 2$.

(See Exercise 4.1.)

Example 4.3

The Petersen graph is not hamiltonian.

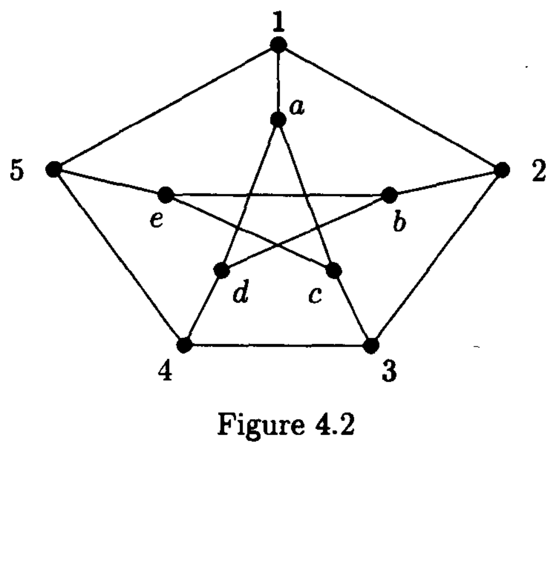


Figure 4.2

Solution

Label the vertices as shown in Figure 4.2, and suppose there is a hamiltonian cycle. Every time the cycle goes from the outside along one of the "spokes" $1a, 2b, 3c, 4d, 5e$, it has to return to the outside along another spoke. So the hamiltonian cycle must contain either 2 or 4 spokes.

(a) Suppose there are 4 spokes in a hamiltonian cycle: we can assume $5e$ is the one spoke not in it. Then 51 and 54 must be in the cycle, as must eb and ec. Since $1a$ and 15 are in the cycle, 12 is not, so 23 is. But this gives the cycle $2\,3\,c\,e\,b\,2$ as part of the hamiltonian cycle, which is clearly impossible.

(b) Suppose there are just two spokes in the hamiltonian cycle. Take $1a$ as one of them. Then ac or ad is in the cycle - say ad. Then ac is not, so $c3$ is. So spokes $b2, d4, e5$ are not in the cycle. Since $b2$ is not in the cycle, 23 must be. Similarly, since $d4$ is not in, 34 must be in the cycle. So all three edges from 3 are in the cycle, a contradiction.

There is no straightforward way of characterising hamiltonian graphs. Perhaps the best known simple sufficient condition is that given by Dirac in the following theorem, but it must be emphasised that the condition given is not at all necessary (as can be seen by considering the cycle C_n, $n \geq 5$).

Theorem 4.2 (Dirac, 1950)

If G is a simple graph with p vertices, each vertex having degree $\geq \frac{1}{2}p$, then G is hamiltonian.

Proof

Outlined in Exercise 4.6.

4.2 Planarity and Hamiltonian Graphs

There are some interesting connections between planar graphs and hamiltonian graphs. The first arose in connection with the Four Colour Conjecture (FCC), when it was realised that the presence of a hamiltonian cycle in a plane graph makes the colouring of its regions (faces) with four colours very easy. For example, consider the problem of colouring the faces of a dodecahedron using four colours. Figure 4.3 shows a hamiltonian cycle which divides the regions into an internal chain of regions, and an external chain. Colour the internal chain with colours A and B, and the external chain with colours C and D.

Early on in the history of the FCC, Tait conjectured that every polyhedral map in which every vertex has degree 3 has a hamiltonian cycle. (A map is polyhedral if any two adjacent regions meet in a single common edge or a

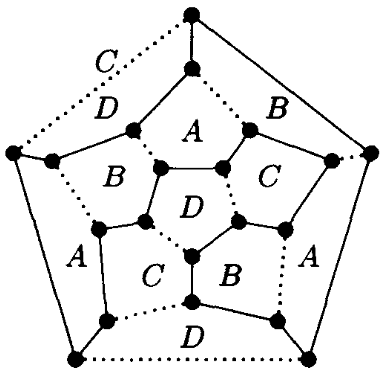


Figure 4.3 Dodecahedron

single point.) The truth of Tait's conjecture would have implied that every such map is 4-colourable; however, the conjecture was finally proved false in 1946, when Tutte constructed a counterexample.

Another connection between hamiltonicity and planarity occurs in the following algorithm which can be used to determine whether or not a given hamiltonian graph is planar. The basic idea is that if a graph G is both hamiltonian and planar, then, in a plane drawing of G, the edges of G which are not in the hamiltonian cycle H will fall into two sets, those drawn inside H and those drawn outside.

The planarity algorithm for hamiltonian graphs

1. Draw the graph G with a hamiltonian cycle H on the outside, i.e. with H as the boundary of the infinite region.
2. List the edges of G not in H: $e_1, \ldots, e_r$.
3. Form a new graph K in which the vertices are labelled $e_1, \ldots, e_r$ and where the vertices labelled e_i, e_j are joined by an edge if and only if e_i, e_j cross in the drawing of G, i.e. cannot both be drawn inside (or outside) H (such edges are said to be **incompatible**).
4. Then G is planar if and only if K is bipartite.

(If K is bipartite, with bipartition $B \cup W$, then the edges e_i coloured B can be drawn inside H, and the edges coloured W can be drawn outside.)

In practice, we introduce the edges one by one, as follows.

Example 4.4

Test the graph shown in Figure 4.4 for planarity.

Solution

1. The graph is already drawn with hamiltonian cycle *abcdefa* on the outside.

ɜd that every
roved false in

in the follow-
iven hamilto-
hamiltonian
ıre not in the
H and those

', i.e. with H

e_r and where
if e_i, e_j cross
ide) H (such

:ed B can be
)

the outside.

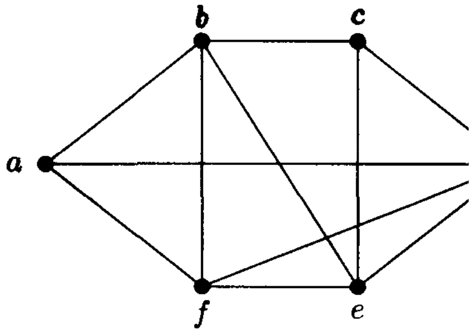
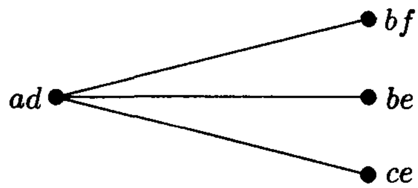
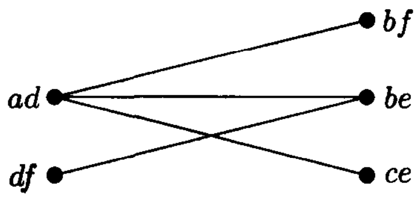


Figure 4.4

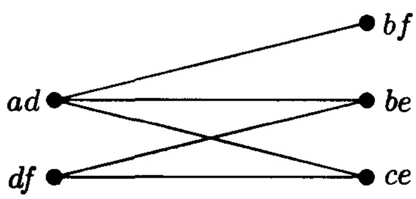
2. Edges not in the hamiltonian cycle are $ad, be, bf,$
3. Start with ad; it is incompatible with bf, be, ce:



Now consider bf. It crosses only ad. Next consid
get:



Now consider ce. It also crosses df, so we get:



4. By now we have the full graph K. (Check: the
the number of crossings of edges in G.) Since K
that G is planar, and we can draw it with ad an
outside (Figure 4.5).

Example 4.5

Show that $K_{3,3}$ is not planar.

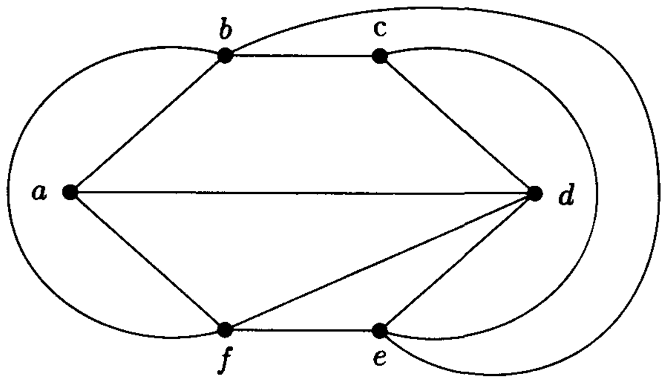


Figure 4.5

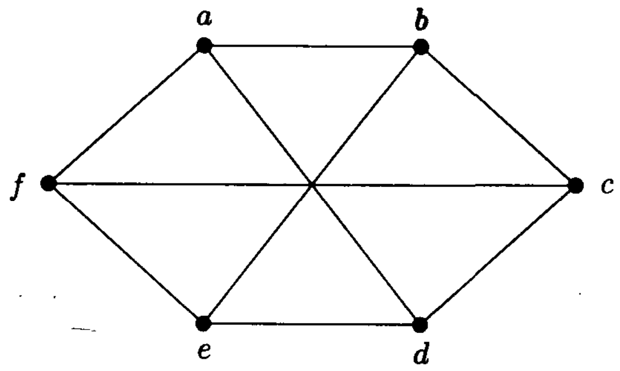
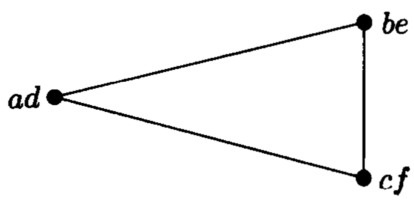


Figure 4.6

Solution

1. In Figure 4.6 we have $K_{3,3}$ drawn with hamiltonian cycle on the outside.
2. Edges not in hamiltonian cycle are ad, be, cf.
3. Obtain:



4. This is **not** bipartite, so $K_{3,3}$ is not planar.

4.3 The Travelling Salesman Problem

A sales representative of a publisher of mathematical texts has to make a round trip, starting at home, and visiting a number of university bookshops before returning home. How does the salesman choose his route to minimise the total distance travelled?

Here we consider a weighted graph, in which the vertices represent the bookshops and his home, and the edges represent the routes between them, each edge being labelled by the length of the route it represents. The salesman wishes to find a hamiltonian cycle of minimum length, i.e. of minimum total weight.

A complete graph K_n has $(n-1)!$ different hamiltonian cycles (or $\frac{1}{2}(n-1)!$ if we do not distinguish between a cycle and its "reverse"), so finding the one of minimum weight by looking at each in turn is out of the question when n is large. Even for $n = 10$, $\frac{1}{2}(n-1)! = 181\,440$. There is no really efficient algorithm yet known for solving the travelling salesman problem (TSP), so "good" rather than "best" routes are sought, as are estimates, rather than exact values, of the shortest total length.

Lower bounds

Lower bounds can be found by using spanning trees. First observe that if we take any hamiltonian cycle and remove one edge then we get a spanning tree, so

$$\text{Solution to TSP} > \text{ minimum length of a spanning tree (MST).} \qquad (4.1)$$

But we can do better. Consider any vertex v in the graph G. Any hamiltonian cycle in G has to consist of two edges from v, say vu and vw, and a path from u to w in the graph $G-(v)$ obtained from G by removing v and its incident edges. Since this path is a spanning tree of $G - \{v\}$, we have

$$\text{Solution to TSP} \geq \left\{ \begin{array}{c} \text{sum of lengths of two} \\ \text{shortest edges from } v \end{array} \right\} + \left(\begin{array}{c} \text{MST of} \\ G - \{v\} \end{array} \right). \qquad (4.2)$$

Example 4.6

Apply (4.2) to the graph of Figure 4.7.

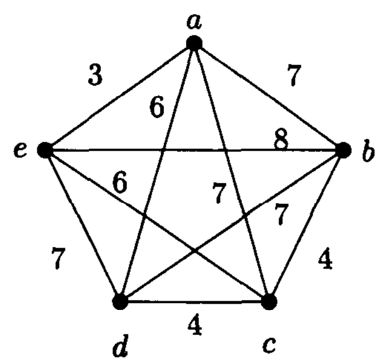


Figure 4.7

Solution

Choose vertex a. The two shortest edges from a have lengths 3 and 6. The minimum weight spanning tree of $G - \{a\}$ consists of edges bc, cd and ec, and has length 14. So, by (4.2), a lower bound for the TSP is $3 + 6 + 14 = 23$.

Instead, we could have started with b. The two shortest edges from b have lengths 4 and 7, and the minimum weight spanning tree of $G - \{b\}$ has length 13, so we obtain the lower bound $4 + 7 + 13 = 24$. This second bound gives us more information than the first.

Upper bounds

Assume that the weights are distances, satisfying the triangle inequality

$$d(x, z) \leq d(x, y) + d(y, z)$$

where $d(x, y)$ denotes the shortest distance along edges from x to y. In this case the following method gives upper bounds for the TSP in K_n.

Find a minimum spanning tree of K_n, say of weight w. We can then find a walk of length $2w$ which visits every vertex at least once, and which returns to its starting point, by going "round" the tree as shown in Figure 4.8.

Figure 4.8

We now try to reduce the length of this walk by taking shortcuts. Start at one vertex and follow the walk round. When we reach an edge which will take us to a vertex already visited, take the direct route to the next vertex not yet visited. For example, in Figure 4.8, which shows the minimum spanning tree of the graph of Example 4.6, we could start at a and obtain $aecbda$, which has length 26.

Since this method yields a hamiltonian cycle of length no greater than twice MST, we have

$$\text{MST} < \text{ solution to TSP} \leq 2\,\text{MST}, \tag{4.3}$$

and, since MST < solution to TSP, by (4.1), we have constructed a hamiltonian cycle of length at most twice the minimum possible length. In Section 4.5 we shall improve this to at most $\frac{3}{2}$ times the minimum.

4.4 Gray Codes

A Gray code of order n is a cyclic arrangement of the 2^n binary sequences of length n such that any pair of adjacent sequences differ in only one place. For example, Figure 4.9(a) shows a Gray code of order 3.

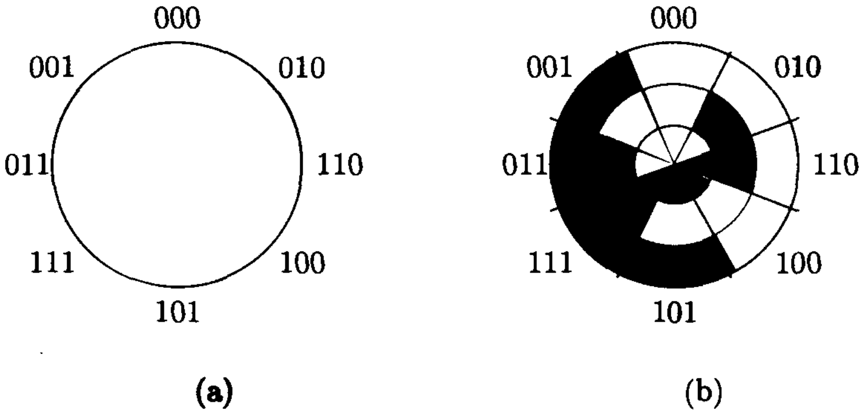


Figure 4.9

The industrial use of Gray codes is on account of their ability to describe the angular position of a rotating wheel. As in Figure 4.9(b), 0 and 1 are represented by white and black (off and on), and are read by electrical contact brushes. The fact that adjacent sequences differ in only one place reduces errors when the contact brushes are close to a boundary between segments. (Compare with 1999 changing to 2000 in a car milometer.)

Note that the code above corresponds to a hamiltonian cycle in a 3-dimensional cube (follow the arrows in Figure 4.10). Note also that the cycle involves going

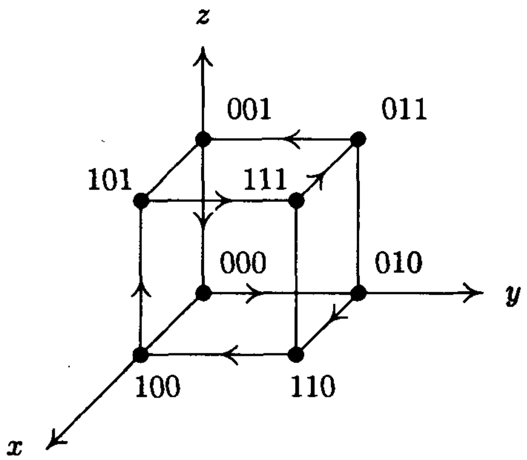


Figure 4.10

round the bottom of the cube (i.e. round a 2-dimensional cube!) with third coordinate 0, then moving up to change the third coordinate to 1, and then tracing out the 2-dimensional cube at the top, in the opposite direction. This idea generalises. So to obtain a Gray code of order 4, write down a Gray code of order 3 with 0 appended to each binary word, then follow it with the same Gray code of order 3, in reverse order, with 1 at the end of each word. This gives

$$0000 - 0100 - 1100 - 1000 - 1010 - 1110 - 0110 - 0010 - 0011-$$

$$0111 - 1111 - 1011 - 1001 - 1101 - 0101 - 0001 - 0000.$$

4.5 Eulerian Graphs

The driver of a snow plough wishes to set out from the depot, travel along each road exactly once, and return to the depot. When is this possible? Similarly, the citizens of Königsberg wished to cross every bridge exactly once and return home. Both problems ask for a closed trail of a particular type.

Definition 4.2

An **eulerian circuit** is a closed trail which contains each edge of the graph. A graph which contains an eulerian trail is called an **eulerian graph**.

It was observed in Section 3.1 that a necessary condition for the existence of an eulerian circuit is that all vertex degrees must be even. It turns out that this condition is also sufficient in connected graphs. Our proof will use the following lemma.

Lemma 4.3

Let G be a graph in which every vertex has even degree. Then the edge set of G is an edge-disjoint union of cycles.

Proof

Proceed by induction on q, the number of edges. The lemma is true for $q = 2$, so consider a graph G with k edges and suppose that the lemma is true for all graphs with $q < k$. Take any vertex v_0, and start a walk from v_0, continuing until a vertex already visited is visited for the second time. If this vertex is v_j, then the part of the walk from v_j to v_j is a cycle C. Remove C to obtain a graph H with $< k$ edges and in which every vertex has even degree. By induction, H is an edge-disjoint union of cycles, so the result follows.

Theorem 4.4

Let G be a connected graph. Then G is eulerian if and only if every vertex has even degree.

Proof

$\Rightarrow$. Already shown.

$\Leftarrow$. Suppose every vertex has even degree. Then the edges fall into disjoint cycles. Take any such cycle C_1. If C_1 does not contain all the edges of G then, since G is connected, there must be a vertex $v_1 \in C_1$ and an edge v_1v_2 not in C_1. Now v_1v_2 is in some cycle, say C_2, disjoint from C_1. Insert C_2 into C_1 at v_1 to obtain a closed trail. If this trail does not contain all edges of G, take a vertex v_3 in $C_1 \cup C_2$ and edge v_3v_4 not in $C_1 \cup C_2$. Then v_3v_4 is in some cycle C_3 which we insert into $C_1 \cup C_2$. Continue in this way until all edges are used up.

Example 4.7

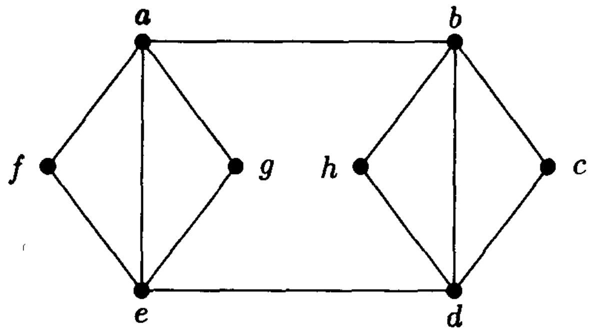


Figure 4.11

In Figure 4.11, first take cycle *abcdefa*. Then insert cycle *agea* at *a*, and finally insert cycle *bdhb* at *b* to obtain eulerian trail

$$a\,g\,e\,a\,b\,d\,h\,b\,c\,d\,e\,f\,a.$$

Definition 4.3

An **eulerian trail** is a trail which contains every edge of the graph, but is not closed. A non-eulerian graph which contains an eulerian trail is called a **semi-eulerian** graph.

The following result follows immediately from Theorem 4.4.

Theorem 4.5

A connected graph G is semi-eulerian if and only if it contains precisely two vertices of odd degree.

Example 4.8

In the Königsberg bridge problem, suppose that one further bridge is built. The resulting graph will then have two vertices of odd degree and hence will contain an eulerian trail.

An upper bound for the TSP

The following method yields a hamiltonian cycle in a complete graph whose length is at most $\frac{3}{2}$ times the length of the minimum hamiltonian cycle. This improves the bound in Section 4.3.

Given K_n, labelled by the length of the edges, first find a minimum spanning tree T. T must, by Exercise 3.1, have an **even** number $2m$ of vertices of odd degree. It is then possible to join these $2m$ vertices into m pairs by using m edges of K_n. Such a set of disjoint edges is called a **matching**. There will be many ways of choosing such a matching, so we choose a matching M of smallest total length. If we now add the edges of M to T, we obtain the new graph $M \cup T$ in which every vertex has even degree: thus $M \cup T$ possesses an eulerian circuit.

For example, with the graph of Example 4.7, T has length 17 (as in Figure 4.8) and T has four vertices of odd degree. Take $M = \{ad, bc\}$ to obtain $M \cup T$ as shown in Figure 4.12. An eulerian circuit is *aecbcda*.

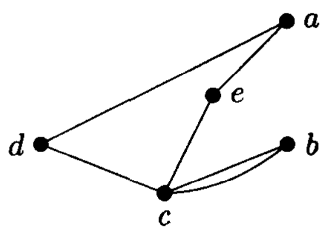


Figure 4.12

Starting at a, we can take *aecb* and, to avoid visiting c twice, go directly from b to d, and then to a, obtaining the hamiltonian cycle *aecbda* which has length 26.

We now show that the eulerian circuit obtained by this method always has length $\leq \frac{3}{2}$MST. Let TSP, EC, MST, M denote respectively the lengths of the minimum hamiltonian cycle, the eulerian circuit, the minimum spanning tree and the matching. Then

$$\text{EC} = \text{MST} + \text{M}, \quad \text{TSP} > \text{MST}.$$

The $2m$ vertices of M will occur, in some order, say $x_1, \dots, x_{2m}$, in the minimum length hamiltonian cycle. If for each $i < 2m$, we replace the part of the cycle between x_i and x_{i+1} by the edge $x_i x_{i+1}$, and we replace the part between x_{2m} and x_1 by the edge $x_{2m}x_1$, we obtain

$$\ell(x_1, x_2) + \ell(x_2, x_3) + \cdots + \ell(x_{2m}, x_1) \leq \text{TSP}$$

where $\ell(x_1, x_{i+1})$ denotes the length of the edge $x_i x_{i+1}$. Thus we have

$$(\ell(x_1, x_2) + \ell(x_3, x_4) + \cdots + \ell(x_{2m-1}, x_{2m})) + (\ell(x_2, x_3) + \cdots + \ell(x_{2m}, x_1)) \leq \text{TSP}.$$

So we obtain two matchings of $x_1, \dots, x_{2m}$ whose lengths sum to $\leq$ TSP. One of these matchings must have total length $\leq \frac{1}{2}$TSP, so that

$$\mathrm{M} \leq \frac{1}{2}\mathrm{TSP}.$$

Thus EC = MST + M $\leq$ TSP + $\frac{1}{2}$TSP = $\frac{3}{2}$TSP. Thus, on using shortcuts in the eulerian circuit to avoid repeating vertices, we obtain a hamiltonian cycle whose length is $\leq \frac{3}{2}$TSP.

4.6 Eulerian Digraphs

A **digraph** or **directed graph** is a graph in which each edge is assigned a direction, indicated by an arrow. In place of the degree of a vertex we have the **indegree**, the number of edges directed towards the vertex, and the **outdegree**, the number of edges directed away from the vertex.

Example 4.9

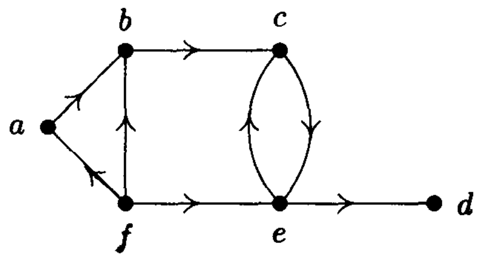


Figure 4.13

In Figure 4.13, the indegrees of $a, \dots, f$ are respectively $1, 2, 2, 1, 2, 0$, and the outdegrees are $1, 1, 1, 0, 2, 3$. It should be clear why the sum of the indegrees equals the sum of the outdegrees.

An eulerian circuit in a digraph is exactly what we would expect; it has to follow the directions of the arrows at each stage. If every vertex has its indegree equal to its outdegree then, as in Lemma 4.3, the edge set can be partitioned as an edge-disjoint union of directed cycles, and, as in Theorem 4.4, we obtain:

Theorem 4.6

A connected digraph has an eulerian circuit if and only if each vertex has its indegree and outdegree equal.

Memory wheels

It is said that the meaningless Sanskrit word *yamátárájabhánasalagám*

has been used as a memory aid by Indian drummers. It has in it every 3-tuple of accented and unaccented vowels, each 3-tuple appearing once. We can display this by replacing unaccented vowels by 0 and accented vowels by 1, to obtain

$$0\,1\,1\,1\,0\,1\,0\,0\,0\,1. \tag{4.4}$$

The 3-tuples 011, 111, 110, 101, 010, 100, 000, 001 appear in it in this order. Note that the last two digits of (4.4) are the same as the first two, so we can obtain a "memory wheel" by overlapping the ends as shown in Figure 4.14.

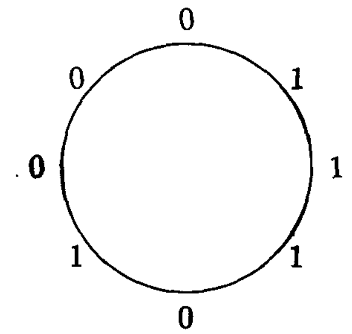


Figure 4.14

Now this arrangement achieves what a Gray code achieved, but much more efficiently. A sensor placed at the edge of the wheel can read off triples of digits and thereby determine how far the wheel has rotated. A Gray code for 8 positions would require three circles of 8 digits, i.e. 24 digits, whereas the memory wheel uses only 8.

We now try to generalise this idea: can a circular arrangement of 2^n binary digits be found which includes all 2^n n-digit binary sequences? One approach might be via hamiltonian cycles. Since, in the above example, 1<u>10</u> is followed by <u>10</u>1, and 1<u>01</u> by <u>01</u>0, we could take the triples xyz as the vertices of a graph and join xyz and yzw by an edge to obtain the directed graph of Figure 4.15.

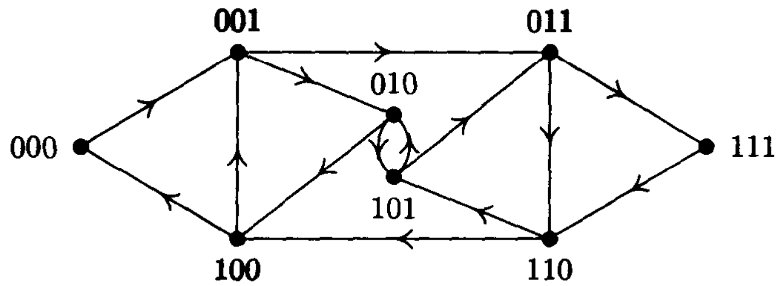


Figure 4.15

The directed hamiltonian cycle

$$000 - 001 - 011 - 111 - 110 - 101 - 010 - 100 - 000$$

yields the memory wheel of Figure 4.14. The trouble with this approach, however, is that it is not at all easy to see how to obtain a hamiltonian cycle in the corresponding digraph when $n \geq 4$.

The problem was however solved by I.J. Good, in a 1946 paper in number theory. Instead by taking the triples as the vertices, Good took the triples as the edges of a graph, in which the vertices corresponded to the overlapping 2-tuples. So, for $n = 3$, we form the digraph of Figure 4.16.

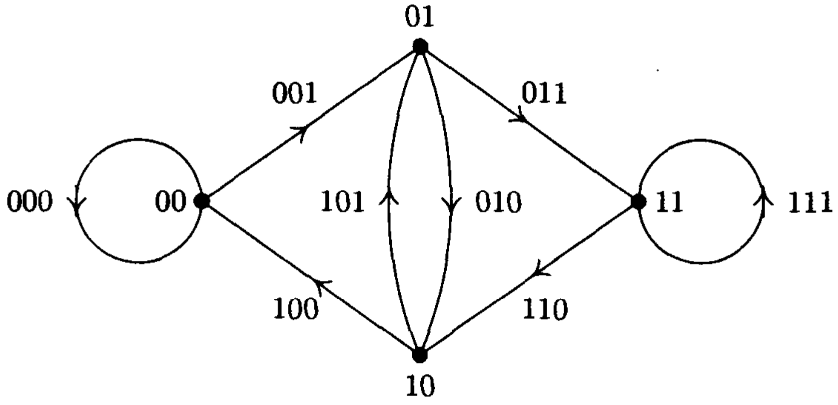


Figure 4.16

Now in this digraph all vertices have indegree and outdegree equal, so the digraph contains an eulerian circuit. Once such circuit consists of the edges

$$000 - 001 - 011 - 111 - 110 - 101 - 010 - 100 - 000,$$

and this gives the same memory wheel as before.

In general, take as vertices the $(n-1)$-digit binary sequences, and draw a directed edge from $x_1x_2 \ldots x_{n-1}$ to $x_2 \ldots x_{n-1}x_n$, labelling the edge $x_1x_2 \ldots x_n$. The resulting digraph has an eulerian circuit which yields a memory wheel.

Example 4.10

Obtain a memory wheel containing all 16 4-digit binary sequences.

Solution

Construct a digraph with 8 vertices labelled by the eight 3-digit binary sequences, and draw a directed edge from $x_1x_2x_3$ to x_2x_30 and to x_2x_31. The digraph of Figure 4.17 is obtained.

An eulerian circuit is (in terms of vertices)

$$000 - 000 - 001 - 011 - 111 - 111 - 110 - 101$$
$$-011 - 110 - 100 - 001 - 010 - 101 - 010 - 100 - 000$$

i.e. in terms of edges,

$$0000 - 0001 - 0011 - 0111 - 1111 - 1110 - 1101 - 1011$$
$$-0110 - 1100 - 1001 - 0010 - 0101 - 1010 - 0100 - 0000.$$

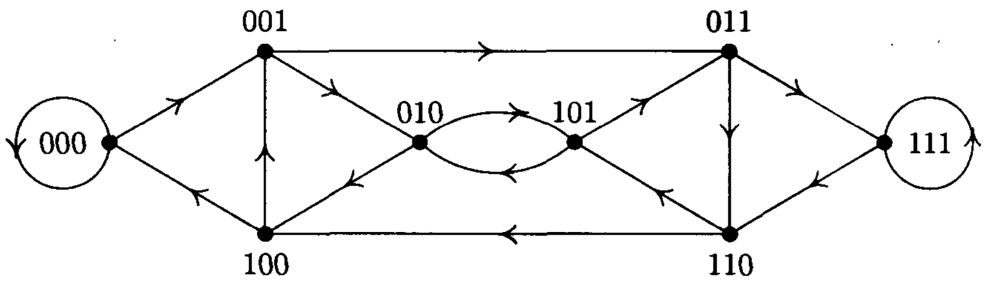


Figure 4.17

The corresponding memory wheel is as shown in Figure 4.18.

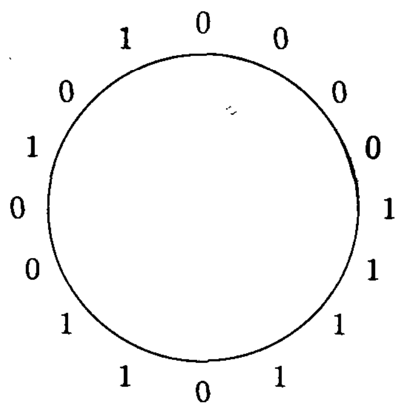


Figure 4.18

The problem of constructing memory wheels is also known as the rotating drum problem. The circular binary sequences are often called **maximum length shift register sequences**, or **de Bruijn sequences** after the Dutch mathematician N.G. de Bruijn who wrote about them in 1946 (although it turned out that they had been constructed many years before by C. Flye Sainte-Marie). They have been used worldwide in telecommunications, and there have been recent applications in biology.

Exercises

Exercise 4.1

(a) Strengthen Theorem 4.1 to: if a bipartite graph, with bipartition $V = B \cup W$, is hamiltonian, then $|B| = |W|$.

(b) Deduce that $K_{m,n}$ is hamiltonian if and only if $m = n \geq 2$.

Exercise 4.2

For each graph in Figure 4.19, determine whether (a) it is hamiltonian, (b) it is eulerian, (c) it is semi-eulerian.

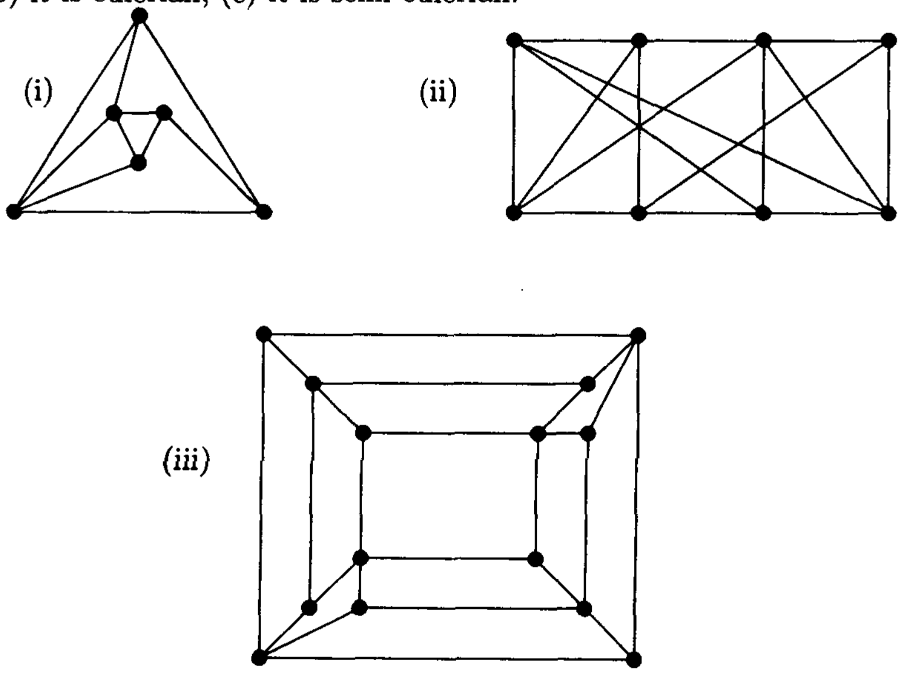


Figure 4.19

Exercise 4.3

Which of the platonic solid graphs are (a) hamiltonian, (b) eulerian?

Exercise 4.4

Use the planarity algorithm to determine whether or not the graphs in Figure 4.20 are planar.

Exercise 4.5

Construct a Gray code of order 5.

Exercise 4.6

Dirac's theorem. Prove Theorem 4.2 as follows. Suppose G is not hamiltonian. By adding edges we can assume that G is "almost" hamiltonian in the sense that the addition of any further edge will give a hamiltonian graph. So G has a path $v_1 \to v_2 \to \cdots \to v_p$ through every

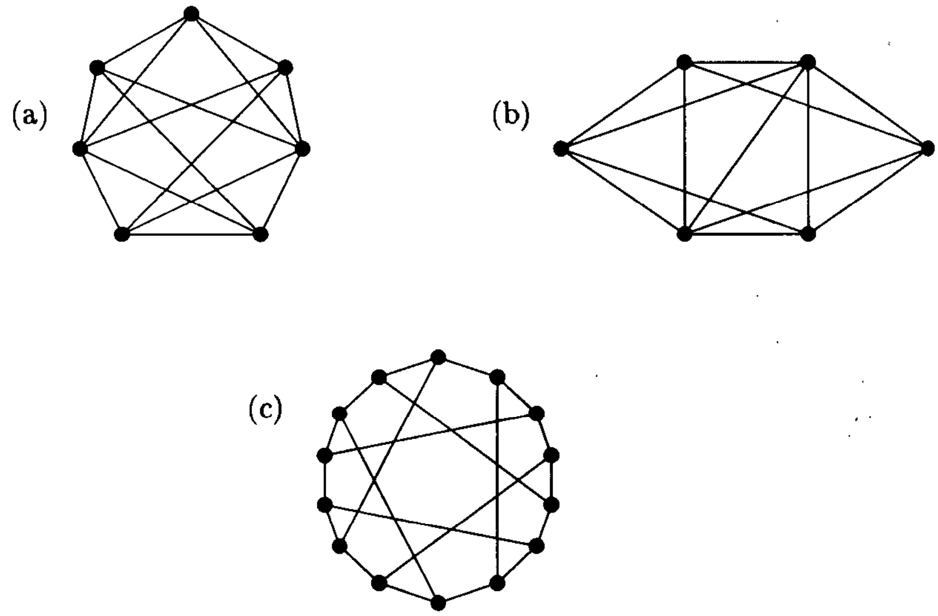


Figure 4.20

vertex, where v_1 and v_p are not adjacent. Show that there must be a vertex v_i adjacent to v_1, with v_{i-1} adjacent to v_n. This gives a hamiltonian cycle $v_1 \to \cdots \to v_{i-1} \to v_n \to \cdots \to v_{i+1} \to v_i \to v_1$.

Exercise 4.7

(a) **Ore's theorem.** Imitate the proof of Dirac's theorem to show that if G is a simple graph with $p \geq 3$ vertices, with $\deg(v) + \deg(w) \geq p$ for each pair of non-adjacent vertices v, w, then G is hamiltonian.
(b) Deduce that if G has $2 + \frac{1}{2}(p-1)(p-2)$ edges then G is hamiltonian.
(c) Find a non-hamiltonian graph with $1 + \frac{1}{2}(p-1)(p-2)$ edges.

Exercise 4.8

By removing vertex A , find a lower bound for the TSP for the graph of Exercise 3.7. Repeat, removing vertex B. Then obtain an upper bound by the method of Section 4.5.

Exercise 4.9

Find upper and lower bounds for the TSP for the situation in Exercise 3.8. How do your results compare with the exact solution?

Exercise 4.10

Construct a memory wheel containing all 32 5-digit binary sequences.

Exercise 4.11

Use digraphs to construct a memory wheel of length 9 containing all 2-digit ternary sequences (formed from the digits $0, 1, 2$). Then find one for all 3-digit ternary sequences.

Exercise 4.12

Dominoes. Can you arrange the 28 dominoes of an ordinary set in a closed loop, so that each matches with its neighbour in the usual way? Can you do so if all dominoes with a 6 on them are removed? Can you state a general theorem about dominoes with numbers $0, 1, \ldots, n$ on them? (Hint: consider each domino as an edge of a graph with vertices labelled $0, 1, \ldots, n$.)

Exercise 4.13

Figure 4.21 shows an arrangement of the numbers $1, \ldots, 5$ round a circle, so that each number is adjacent to every other number exactly once. Can you produce a similar arrangement for $1, \ldots, 7$? Use Euler's theorem to show that there is a solution for n numbers if and only if n is odd. Can you salvage a similar type of result when n is even?

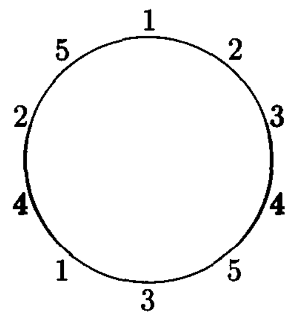


Figure 4.21

5
Partitions and Colourings

In this chapter we consider partitions of a set, introducing the Stirling numbers and the Bell numbers. We then consider vertex and edge colourings of a graph, where the vertex set and the edge set are partitioned by the colours.

5.1 Partitions of a Set

A **partition** of a set S is a collection of non-empty subsets $S_1, \ldots, S_r$ of S which are pairwise disjoint and whose union is S. The subsets S_i are called the **parts** of the partition. For example $\{1,2,4\} \cup \{3,6\} \cup \{5\}$ is a partition of $\{1, \ldots, 6\}$ into three parts. Note that it does not matter in what order the parts appear.

Example 5.1

In a game of bridge, the 52 cards of a standard pack are distributed among four people who receive 13 cards each. In how many ways can the pack of 52 cards be partitioned into four sets of size 13?

Solution

We can choose 13 cards in $\binom{52}{13}$ ways. From the remaining 39 we can choose a further 13 in $\binom{39}{13}$ ways, and then from the remaining 26 we can choose 13 in $\binom{26}{13}$ ways. This leaves a final set of 13 cards. So we have

$$\binom{52}{13}\binom{39}{13}\binom{26}{13} = \frac{52!}{13!39!} \cdot \frac{39!}{13!26!} \cdot \frac{26!}{13!13!} = \frac{52!}{(13!)^4}$$

ways of partitioning the pack. But these partitions are not all distinct, since each distinct partition arises in 4! ways, depending on which of the four sets in it is chosen first, which is chosen second, and so on. So the required number is

$$\frac{52!}{(13!)^4 4!}$$

(a vast number, greater than 10^{27}).

There is another way of approaching this counting problem. Consider a row of 52 spaces grouped into four groups of 13:

$$(\ldots\ldots)\,(\ldots\ldots)\,(\ldots\ldots)\,(\ldots\ldots).$$

The cards can be placed in the spaces in 52! ways. Within each group there are 13! ways of arranging the same 13 cards, and these different arrangements are irrelevant since they give rise to the same part of the partition, so we have to divide by $(13!)^4$, one 13! for each group. Then the four groups themselves can be arranged in 4! ways, so we have to divide by 4!, giving the same answer as before.

This argument easily generalises, to give the following result.

Theorem 5.1

A set of mn objects can be partitioned into m sets of size n in

$$\frac{(mn)!}{(n!)^m m!}$$

different ways.

Corollary 5.2

A set of $2m$ objects can be partitioned into m pairs in

$$\frac{(2m)!}{2^m m!}$$

different ways.

Example 5.2

The number of ways of pairing 16 teams in a football cup draw is

$$\frac{16!}{2^8 8!} = 2\,027\,025.$$

The same type of argument can be applied when the parts of the required partition are not all of the same size.

Example 5.3

In how many ways can a class of 25 pupils be placed into four tutorial groups of size 3, two of size 4 and one of size 5?

Solution

Consider the following grouping of 25 spaces

$$(- - -)\,(- - -)\,(- - -)\,(- - -)\,(- - - -)\,(- - - -)\,(- - - - -).$$

The 25 pupils can be placed in the spaces 25! ways. To count **distinct** partitions we have to take into account the ways of ordering the pupils within the groups - so we divide by $(3!)^4(4!)^2 5!$ - and also the ways of ordering the groups themselves - so we divide by 4! on account of the four groups of size 3 and by 2! on account of the two groups of size 4. So the required number is

$$\frac{25!}{(3!)^4(4!)^2 5!4!2!} \cong 3.6 \times 10^{15}.$$

Definition 5.1

A partition of an n-element set consisting of α_i subsets of size i, $1 \le i \le n$, where $\sum_{i=1}^{n} i\alpha_i = n$, is called a partition of **type** $1^{\alpha_1}2^{\alpha_2}\ldots n^{\alpha_n}$.

Generalising Example 5.3 gives the following result.

Theorem 5.3

The number of partitions of type $1^{\alpha_1}2^{\alpha_2}\ldots n^{\alpha_n}$ of an n-element set is

$$\frac{n!}{\prod_{i=1}^{n}(i!)^{\alpha_i}\alpha_i!}.$$

Example 5.4

The number of ways of grouping 10 people into two groups of size 3 and one group of size 4 is the number of partitions of type $3^2 4^1$ and so is

$$\frac{10!}{(3!)^2 2!4!} = 2100.$$

5.2 Stirling Numbers

In this section we think about partitioning a set into a given number of parts.

Definition 5.2

Let $S(n,k)$ denote the number of ways of partitioning an n-set into exactly k parts. Then $S(n,k)$ is called a **Stirling number of the second kind**.

These numbers are named after the Scottish mathematician James Stirling (1692–1770), who is also known for his approximation of $n!$:

$$n! \sim \sqrt{2\pi n}n^n e^{-n}.$$

Stirling also has numbers of the first kind named after him – see Exercise 5.10.

We now study $S(n,k)$. Clearly, for all $n \geq 1$,

$$S(n,1) = S(n,n) = 1. \tag{5.1}$$

Example 5.5

We show that $S(4,2) = 7$. Here are the seven ways of partitioning $\{1,2,3,4\}$ into two parts: $\{1\}\cup\{2,3,4\}, \{2\}\cup\{1,3,4\}, \{3\}\cup\{1,2,4\}, \{4\}\cup\{1,2,3\}, \{1,2\}\cup\{3,4\}, \{1,3\}\cup\{2,4\}$ and $\{1,4\}\cup\{2,3\}$.

Clearly, for large n, we need a better way of evaluating $S(n,k)$ than just writing down all possible partitions. Such a method is given by the following recurrence relation.

Theorem 5.4

$$S(n,k) = S(n-1,k-1) + k\,S(n-1,k) \tag{5.2}$$

whenever $1 < k < n$.

Proof

In any partition of $\{1,\dots,n\}$ into k parts, the element n may appear by itself as a 1-element subset or it many occur in a larger set. If it appears by itself, then the remaining $n-1$ elements have to form a partition of $\{1,\dots,n-1\}$ into $k-1$ subsets, and there are $S(n-1,k-1)$ ways in which this can be done. On the other hand, if the element n is in a set of size at least two, we can think of partitioning $\{1,\dots,n-1\}$ into k sets - this can be done in $S(n-1,k)$ ways - and then of introducing n into one of the k sets so formed - and there are k ways of doing this. So, by the addition and multiplication principles, we have $S(n,k) = S(n-1,k-1) + k\,S(n-1,k)$.

Example 5.5 (again)

$$\begin{aligned} S(4,2) &= S(3,1) + 2\,S(3,2) \\ &= \quad 1 \quad + 2\,(S(2,1) + 2\,S(2,2)) \\ &= \quad 1 \quad + 2(1+2) \quad = 7. \end{aligned}$$

Theorem 5.5

For all $n \geq 2$, $S(n,2) = 2^{n-1} - 1$.

Proof

We use induction on n. The result is true for $n = 2$, so suppose it is true for $n = k \geq 2$. Then

$$\begin{aligned} S(k+1,2) &= S(k,1) + 2\,S(k,2) \quad \text{(by 5.2)} \\ &= 1 + 2(2^{k-1} - 1) \\ &= 1 + 2^k - 2 = 2^{(k+1)-1} - 1. \end{aligned}$$

Table 5.1 gives the first few Stirling number $S(n,k)$.

Table 5.1

$n \backslash k$	1	2	3	4	5	6	7	8	$B(n)$
1	1								1
2	1	1							2
3	1	3	1						5
4	1	7	6	1					15
5	1	15	25	10	1				52
6	1	31	90	65	15	1			203
7	1	63	301	350	140	21	1		877
8	1	127	966	1701	1050	266	28	1	4140

Note the number $2^{n-1} - 1$ in the column $k = 2$. On the right of the table are the sums $B(n)$ of all the Stirling numbers in the rows. $B(n)$ is the total number of partitions of an n-set, and is called a **Bell number**, after another Scot, E.T. Bell, who emigrated to the USA. and wrote several popular books on mathematics, including *Men of Mathematics*, an idiosyncratic two-volume collection of "biographies" of famous mathematicians. We have, for $n \geq 1$,

$$B(n) = \sum_{k=1}^{n} S(n,k). \tag{5.3}$$

If we define $B(0) = 1 = S(0,0)$ (accept this as a useful convention, like $\binom{0}{0} = 1$), we can obtain a recurrence relation for the Bell numbers.

Theorem 5.6

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

The nth element of the set being partitioned will appear in one of the sets of the partition along with $j \geq 0$ other elements. There are $\binom{n-1}{j}$ ways of choosing these j elements. The remaining $n-1-j$ elements can then be partitioned in $B(n-1-j)$ ways. So

$$\begin{aligned} B(n) &= \sum_{j=0}^{n-1} \binom{n-1}{j} B(n-1-j) \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} B(k) \quad \text{(on putting } n-1-j=k\text{).} \end{aligned}$$

Example 5.6

$$B(9) = \sum_{k=0}^{8} \binom{8}{k} B(k)$$

$$= 1 + 8.1 + 28.2 + 56.5 + 70.15 + 56.52 + 28.203 + 8.877 + 1.4140$$

$$= 21\,147.$$

For an interesting (but useless!) formula for $B(n)$, see Exercise 5.9.

5.3 Counting Functions

The Stirling numbers arise naturally in the enumeration of all functions $f : X \to Y$ which can be defined from an m-set X to an n-set Y. There are n^m such functions since, for each $x \in X$, there are n possible values for $f(x)$.

Recall that the **image** of $f : X \to Y$ is the set of elements of Y which actually arise as a value $f(x)$ for some $x \in X$:

$$\text{im} f = \{y \in Y : y = f(x) \text{ for some } x \in X\}.$$

Each function $f : X \to Y$ has as its image a subset of Y. How many such functions have an image of size k? If f takes precisely k values then X can be partitioned into k parts, the ith of which will consist of those elements of X which are mapped onto the ith member of im(f). So a function $f : X \to Y$ with image of size k can be constructed as follows:

(i) partition X into k parts $X_1, \dots, X_k$ (this can be done in $S(m,k)$ ways);

(ii) choose the image set of size k in Y (this can be done in $\binom{n}{k}$ ways);

(iii) pair off each X_i with one of the members of the image set (this can be done in $k!$ ways).

So the number of functions $f : X \to Y$ with image of size k is $S(m,k)\binom{n}{k}k!$. Thus, since k can take any value from 1 to n, and since there are n^m functions $f : X \to Y$ altogether, we obtain:

Theorem 5.7

Let $|X| = m$ and $|Y| = n$ where $m, n \geq 1$.

(a) The number of functions $f : X \to Y$ with image of size k is $S(m,k)\binom{n}{k}k!$.

(b)

$$n^m = \sum_{k=1}^{n} S(m,k)\binom{n}{k}k!. \tag{5.4}$$

Note as a special case that the number of **surjections** from X to Y, i.e. functions whose image set is the whole of Y, is $n!S(m,n)$.

Example 5.7

We check (5.4) in the case $n = 4, m = 5$.

$$\sum_{k=1}^{4} S(5,k)\binom{4}{k}k! = 4S(5,1) + 12S(5,2) + 24S(5,3) + 24S(5,4)$$

$$= 4 + 180 + 600 + 240 = 1024 = 4^5.$$

Note that if we define $S(m,0) = 0$ for all $m \geq 1$, and $S(0,0) = 1$, then we can rewrite (5.4) as

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

This identity can be inverted.

Theorem 5.8

For all $m \geq 1, n \geq 0, m \geq n$,

$$n!S(m,n) = \sum_{k=0}^{n} (-1)^{n-k}\binom{n}{k}k^m. \tag{5.5}$$

Proof We can use Corollary 1.15, putting $a_k = k^m$ and $b_k = S(m,k)k!$. Alternatively we shall be able to use the inclusion-exclusion principle in the next chapter: see Section 6.2.

Example 5.8

$$S(5,3) = \frac{1}{3!}\left(\sum_{k=0}^{3}(-1)^{3-k}\binom{3}{k}k^5\right) = \frac{1}{6}(-0+3-3.2^5+3^5) = 25.$$

5.4 Vertex Colourings of Graphs

To colour the vertices of a graph G is to assign a colour to each vertex in such a way that no two adjacent vertices receive the same colour. If we define an **independent** set of vertices of G to be a set of vertices no two of which are adjacent, then a vertex colouring can be thought of as a partition of the set V of vertices into independent subsets. Often we are concerned with the **smallest** number of colours required, i.e. the smallest number of independent sets which partition V; we call this number the chromatic number of G.

Definition 5.3

The **chromatic number** $\chi(G)$ of a graph G is the smallest value of k for which the vertex set of G can be partitioned into k independent subsets.

We have met the idea of colouring vertices already; in Section 3.5 we noted that bipartite graphs are bichromatic; so if G is bipartite with at least one edge then $\chi(G) = 2$. Also, the four colour theorem asserts that $\chi(G) \leq 4$ for all planar graphs G.

Theorem 5.9

(i) $\chi(K_n) = n$.

(ii) $\chi(C_n) = 2$ if n is even; $\chi(C_n) = 3$ if n is odd.

Proof

(i) No two vertices can receive the same colour since they are adjacent.

(ii) If n is even, we can alternate colours round the cycle; if n is odd we need a third colour for the "last" vertex coloured.

Example 5.9

The graph of Figure 5.1(a) has chromatic number 3; it needs at least three colours since it contains C_3, and three colours are sufficient, as shown in Figure 5.1(b).

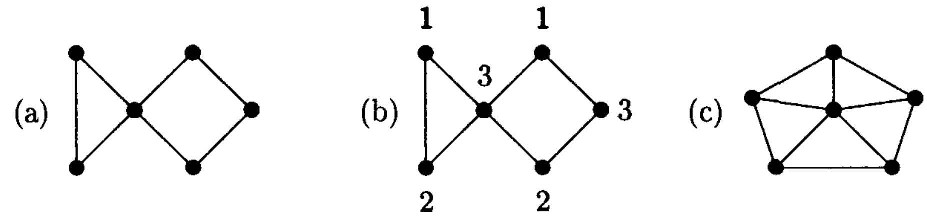


Figure 5.1

Note that the case of C_n, n odd, contradicts the belief of some amateur four-colour-theorem-provers, that a graph needs m colours only if it contains K_m as a subgraph. Another counterexample to this belief is the graph of Figure 5.1(c) which needs four colours (why?) although it does not contain K_4.

There is no easy way of finding $\chi(G)$ for a given graph G. The greedy algorithm, which we now describe, will give an upper bound for $\chi(G)$ related to the maximum vertex degree. In our description of the algorithm we denote colours by $C_1, C_2, C_3, \ldots$ and call C_i the ith colour.

The greedy algorithm for vertex colouring

1. List the vertices in some order: $v_1, \ldots, v_p$.
2. Assign colour C_1 to v_1.
3. At stage $i+1$, when v_i has just been assigned a colour, assign to v_{i+1} the colour C_j with j as small as possible which has not yet been used to colour a vertex adjacent to v_{i+1}.

Example 5.10

We use the greedy algorithm to colour the graph of Figure 5.2 for each of the two vertex orderings shown.

With vertices listed as in (a), we assign colours as follows:

v:	1	2	3	4	5	6	7
C:	1	2	1	3	4	1	2

This colouring uses four colours. However, with the vertices labelled as in (b), we get:

v:	1	2	3	4	5	6	7
C:	1	2	3	1	3	1	2

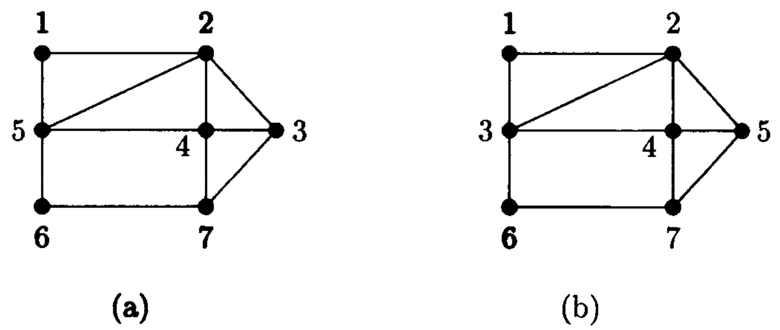


Figure 5.2

This second colouring shows that $\chi(G) \leq 3$; in fact $\chi(G) = 3$ since G is not bipartite.

Clearly, the bound for $\chi(G)$ obtained by the algorithm depends on the particular order in which the vertices are considered. But note that, if a vertex v has degree d then, when it is the turn of v to be assigned a colour, at most d of the colours are ineligible, so it must be given some colour C_i where $i \leq d+1$. Thus we have the following bound.

Theorem 5.10

If G has maximum vertex degree Δ, then the greedy algorithm will colour the vertices of G using at most $\Delta + 1$ colours, so that $\chi(G) \leq \Delta + 1$.

Example 5.11 (A timetabling problem)

The University of Central Caledonia has nine vice-principals, Professors $A, B, \dots, I$, who serve on eight committees. The memberships of the committees are as follows.

Committee 1 :	A,	B,	C,	D
2 :	A,	C,	D,	E
3 :	B,	D,	F,	G
4 :	C,	F,	G,	H
5 :	A,	H,	J	
6 :	H,	I,	J	
7 :	G,	H,	J	
8 :	E,	I.		

Each committee is to meet for a day; no two committees with a member in common can meet on the same day. Find the smallest number of days in which the meetings can take place.

Solution

Represent each committee by a vertex, and join two vertices by an edge precisely when the corresponding committees have overlapping membership. Then the

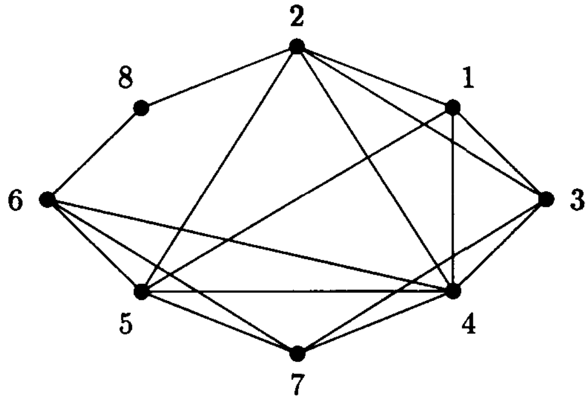


Figure 5.3

minimum number of days required is the chromatic number of the graph G, shown in Figure 5.3. Note that vertices $1, 2, 3, 4$ form a K_4, so at least four colours (days) are needed. But four colours are sufficient: e.g.

$$\{1, 7, 8\} \cup \{3, 5\} \cup \{2, 6\} \cup \{4\}$$

is a partition of $\{1, \ldots, 8\}$ into independent sets. So $\chi(G) = 4$, and four days are enough.

5.5 Edge Colourings of Graphs

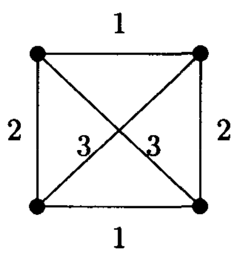
An **edge colouring** of a graph G is an assignment of colours to the edges of G so that no two edges with a common vertex receive the same colour. The minimum number of colours required in an edge colouring of G is called the **chromatic index** of G and is denoted by $\chi'(G)$.

Thus to edge colour a graph is to partition the edge set into subsets such that no two edges in the same subset have a vertex in common, i.e. so that all edges in any part of the partition are **disjoint**. A set of disjoint edges in a graph is often called a **matching**. Clearly, in an edge colouring, all edges at a vertex v must receive different colours, so $\chi'(K_n) \geq n - 1$ for each n.

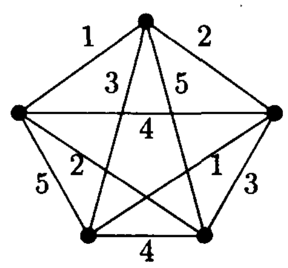
Example 5.12

(a) $\chi'(K_4) = 3$, since $\chi'(K_4) \geq 3$ and three colours suffice, as shown in Figure 5.4(a).

(b) $\chi'(K_5) = 5$. Here, $\Delta = 4$ colours are not enough. For there are 10 edges and no more than two edges in any matching. However, 5 colours are enough, as shown in Figure 5.4(b).



(a)



(b)

Figure 5.4

Theorem 5.11

(i) $\chi'(K_n) = n$ if n is odd.

(ii) $\chi'(K_n) = n - 1$ if n is even.

Proof

(i) If n is odd, any matching in K_n can have at most $\frac{1}{2}(n-1)$ edges. So at most $\frac{1}{2}(n-1)$ edges can be given any one colour. But there are $\frac{1}{2}n(n-1)$ edges in K_n, so at least n colours are needed. Now we can colour the edges using n colours in the following way. Represent K_n as a regular n-gon, with all diagonals drawn. Colour the boundary edges by $1, \ldots, n$; then colour each diagonal by the colour of the boundary edge parallel with it. This gives an edge colouring using n colours. The case $n = 5$ is as in Figure 5.4(b).

(ii) Now suppose n is even. Certainly $\chi'(K_n) \geq n-1$; we show how to use only $n-1$ colours. Since $n-1$ is odd, we can colour K_{n-1} using $n-1$ colours, as described above. Now take another vertex v and join each vertex of K_{n-1} to v, thus obtaining K_n. At each vertex of K_{n-1}, one colour has not been used. The colours missing at each vertex of K_{n-1} are all different, so we can use these $n-1$ colours to colour the added edges at v. This gives an edge colouring of K_n using $n-1$ colours.

The appearance of $\Delta(= n-1)$ and $\Delta + 1(= n)$ as the chromatic indices of K_n, according as n is even or odd, is in accordance with the following result.

Theorem 5.12 (Vizing, 1964)

If G is a simple graph with maximum vertex degree Δ, then $\chi'(G) = \Delta$ or $\Delta + 1$.

We omit the proof of this result; a proof can be found in [**9**]. But we include the statement of the result because it has led to a great deal of work on determining which graphs are **class one** graphs, i.e. satisfy $\chi'(G) = \Delta$, and which are **class two**, i.e. satisfy $\chi'(G) = \Delta + 1$.

Example 5.13

The Petersen graph is class 2. Here $\Delta = 3$, so we have to show that $\chi'(G) \neq 3$. So suppose an edge colouring using only three colours exists. Then the outer 5-cycle uses three colours, and, without loss of generality, we can assume that it is coloured as in Figure 5.5(a). The spokes are then uniquely coloured, as in

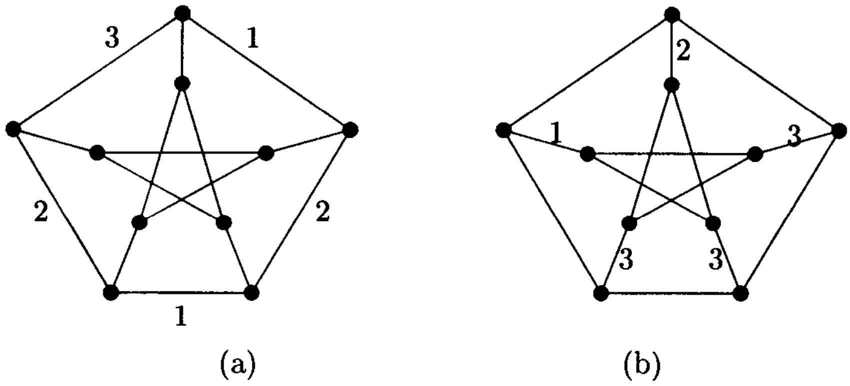


Figure 5.5

Figure 5.5(b). But this leaves two adjacent inside edges which have to be given colour 2. So there is no edge colouring with three colours.

We close this section by establishing that all bipartite graphs are class 1. This result is due to König, the Hungarian author of the first major book on graph theory [**14**].

Theorem 5.13 (König)

$\chi'(G) = \Delta$ for all bipartite graphs G.

Proof

Proceed by induction on q, the number of edges. The theorem is clearly true for graphs with $q = 1$; so suppose it is true for all bipartite graphs with k edges, and consider a bipartite graph G with maximum vertex degree Δ and with $k + 1$ edges. Choose any edge vw of G, and remove it, thereby forming a new bipartite graph H. H has k edges and maximum vertex degree $\leq \Delta$, so, by the induction hypothesis, H can be edge coloured using at most Δ colours.

Now, in H, v and w both have degree $\leq \Delta - 1$, so there is at least one colour missing from the edges from v, and at least one missing from the edges of w. If there is a colour missing at both vertices then it can be used to colour edge vw. If there is no colour missing from both, then let C_1 be a colour missing at v, and C_2 a colour missing at w. Now there is some edge, say vu, coloured C_2; if there is an edge coloured C_1 from u, go along it, and continue along edges coloured C_1 and C_2 alternately as far as possible. The path so constructed will never reach w since if it did it would have to reach w along an edge coloured C_1 and so would be a path of even length, giving, with edge vw, an odd cycle in a bipartite graph. So the connected subgraph K, consisting of vertex v and all vertices and edges of H which can be reached by a path of edges coloured C_1 and C_2, does not contain w. So we can interchange the colours C_1 and C_2 in K without interfering with the colours in the rest of H. This gives a new edge colouring of H in which v and w have no edge coloured C_2, and we can use C_2 to colour vw.

This idea of swapping colours along a path was used by Kempe in his unsuccessful 1879 attempt to prove the four colour theorem. Despite the fact that it did not work there as Kempe had hoped, it nevertheless has proved to be a very useful technique in graph theory.

Example 5.14

Eight students require to consult certain library books. Each is to borrow each required book for a week. The books B_j required by each student S_i are as follows:

$$S_1 : B_1, B_2, B_3 \qquad S_2 : B_2, B_4, B_5, B_6 \qquad S_3 : B_2, B_3, B_5, B_7$$

$$S_4 : B_3, B_5 \qquad S_5 : B_1, B_6, B_7 \qquad S_6 : B_2, B_4, B_6$$

$$S_7 : B_4, B_5, B_7 \qquad S_8 : B_3, B_6.$$

What is the minimum number of weeks required so that each student can borrow all books required?

Solution

Draw a bipartite graph G with vertices labelled $B_1, \ldots, B_7, S_1, \ldots, S_8$, and with S_i joined by an edge to B_j precisely when student S_i has to consult book B_j. Then G has maximum vertex degree $\Delta = 4$, so, by König's theorem, $\chi'(G) = 4$. Thus four colours (weeks) are required. You should be able to partition the set of edges into four disjoint matchings.

Exercises

Exercise 5.1

How many ways are there of arranging 16 football teams into four groups of four?

Exercise 5.2

A class contains 30 pupils. For a chemistry project, the class is to be put into four groups, two of size 7 and two of size 8. In how many ways can this be done?

Exercise 5.3

In the early versions of the Enigma machine, used in Germany in the 1930s, the plugboard swapped six pairs of distinct letters of the alphabet. In how many ways can this be done (assuming 26 letters)?

Exercise 5.4

Any permutation is a product of cycles. For example, the permutation 351642 ($3 \to 1, 5 \to 2, 1 \to 3, 6 \to 4, 4 \to 5, 2 \to 6$) can be written as (31)(2645). How many permutations of $1, \ldots, 8$ are a product of a 1-cycle, two 2-cycles and a 3-cycle?

Exercise 5.5

Prove that (a) $S(n, n-1) = \binom{n}{2}$, (b) $S(n, n-2) = \binom{n}{3} + 3\binom{n}{4}$.

Exercise 5.6

Prove by induction that $S(n, 3) > 3^{n-2}$ for all $n \geq 6$.

Exercise 5.7

Show that $S(n, k) = \sum_{m=k-1}^{n-1} \binom{n-1}{m} S(m, k-1)$ and hence given another proof of Theorem 5.6.

Exercise 5.8

Find $B(10)$.

Exercise 5.9

Use Theorem 5.6 and induction to prove that $B(n) = \frac{1}{e} \sum_{j=0}^{\infty} \frac{j^n}{j!}$.

Exercise 5.10

The (signless) Stirling numbers $s(n,k)$ of the first kind are defined by: $s(n,k)$ is the number of permutations of $1,\ldots,n$ consisting of exactly k cycles. Verify that $s(2,1)=1, s(3,1)=2, s(3,2)=3, s(4,2)=11$ and that $s(n,1)=(n-1)!$. Prove that $s(n,k)=(n-1)s(n-1,k)+s(n-1,k-1)$, and deduce the value of $s(6,2)$.

Exercise 5.11

Find $\chi(G)$ and $\chi'(G)$ for each of the graphs of Exercise 4.2.

Exercise 5.12

Let G be a graph with p vertices and let $\alpha(G)$ denote the size of the largest independent set of vertices of G. Show that $\chi(G)\alpha(G) \geq p$.

Exercise 5.13

Apply the greedy vertex colouring algorithm to the graph of Figure 5.3, taking the vertices (a) in the order $1,\ldots,8$, (b) in order $8,\ldots,1$. Do you get a colouring using four colours?

Exercise 5.14

As Exercise 5.13, but this time choose vertices in (a) increasing, (b) decreasing order of vertex degrees. Which approach would you expect to require fewer colours in general?

Exercise 5.15

Explain why there is always an ordering of the vertices for which the greedy algorithm will lead to a colouring with $\chi(G)$ colours.

Exercise 5.16

Find the chromatic index of each of the five Platonic solid graphs.

Exercise 5.17

A graph in which every vertex degree is 3 is called a **cubic** graph. Prove that all hamiltonian cubic graphs have chromatic index 3. (Note however that not all cubic graphs have chromatic index 3, e.g. the Petersen graph.)

Exercise 5.18

Let G be a graph with an odd number $p = 2k + 1$ of vertices, each of which has the same degree r.

(a) Show that G has $(k + \frac{1}{2})r$ edges.

(b) Explain why no more than k edges can have the same colour in any edge colouring, and hence show that $\chi'(G) = r + 1$. Thus every regular graph with an odd number of vertices is class 2. (This includes K_n, n odd, as shown in Theorem 5.11.)

Exercise 5.19

Let $f_\lambda(G)$ denote the number of ways of colouring the vertices of G using λ given colours.

(a) Show that $f_\lambda(K_n) = \lambda(\lambda - 1)(\lambda - 2) \dots (\lambda - n + 1)$.

(b) Show that $f_\lambda(T) = \lambda(\lambda - 1)^{n-1}$ for all trees T with n vertices.

(c) Let xy be any edge of G. Let G' be the graph obtained from G by removing the edge xy, and let G'' be the graph obtained by identifying vertices x and y. Then $f_\lambda(G) = f_\lambda(G') - f_\lambda(G'')$. Deduce that $f_\lambda(G)$ is a polynomial in λ: it is called the **chromatic polynomial** of G.

(d) Note that the solution $a_n = 2^n + (-1)^n 2$ of Example 2.4 can be interpreted as: $f_3(C_n) = 2^n + (-1)^n 2$. By replacing 3 colours by λ colours, show similarly that $f_\lambda(C_n) = (\lambda - 1)^n + (-1)^n(\lambda - 1)$. Note that this gives $f_2(C_n) = 0$ whenever n is odd, as expected!