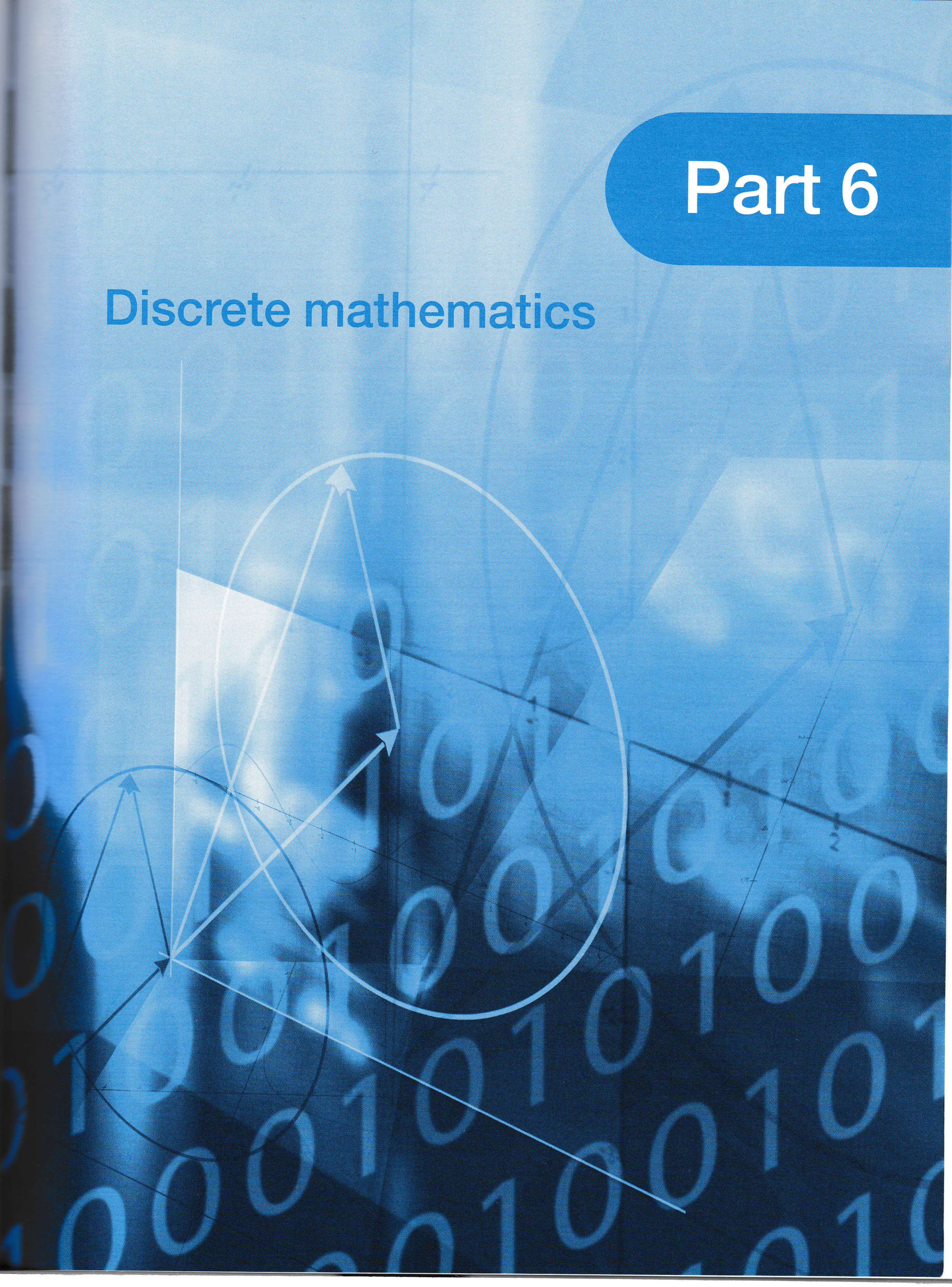
Part 6

Discrete mathematics



Sets

35

CONTENTS

We are often interested in grouping together objects that have common characteristics or features. We might be interested in the integers 1, 2, 3, 4, or in all the integers. Such a group is called a set. The set of all points in a plane would consist of pairs of numbers of the form (x, y), where x and y are coordinates which can take any real values. These examples all involve numbers, but the elements of sets can be other objects such as functions, or matrices, or Fourier series, or Laplace transforms, etc.

35.1 Notation

A set is a collection of objects or elements. The elements in the set can be defined by a rule or in any descriptive manner. Sets are usually denoted by capital letters such as S, A, B, X, etc., and their elements by lowercase letters such as s, a, b, x, etc. The elements in a set are listed between braces $\{\dots\}$. If the set A consists of just two numbers 0 and 1, then we write

$$A = \{0, 1\}, \quad \text{or} \quad A = \{1, 0\}, \tag{35.1}$$

the order being a matter of indifference. We say that 0 and 1 are the elements or members of the set A, or belong to A. We write

$$0 \in A, \qquad 1 \in A,$$

read as '0 belongs to the set A', etc. The number 2 does not belong to A, and we write

$$2 \notin A,$$

that is '2 does not belong to the set A'.

The set defined by (35.1) is the binary set, which could represent the *on* and *off* states of a system. This could be the state of a light switch, for example.

Sets can be either finite, having a finite number of elements, or infinite, in which case the set contains an infinite number of elements. Thus the set given by (35.1) defines a finite set A, while

$$B = \{1, 2, 3, \ldots\},$$

the list of all positive integers, defines an infinite set.

Some of the more common sets have their own special symbols:

Notation for sets of numbers

$\mathbb{R}$, the set of all real numbers
$\mathbb{C}$, the set of all complex numbers
$\mathbb{R}^+$, the set of all positive real numbers (excludes zero)
$\mathbb{Z}$, the set of all integers (positive, negative, and zero)
$\mathbb{N}^+$, the set of all positive integers
$\mathbb{N}^-$, the set of all negative integers
$\mathbb{Q}$, the set of all rational numbers (i.e. numbers of the form p/q where $q \neq 0$ and p are integers) (35.2)

Often the elements are defined by a rule rather than by a list or formula. We write the set as

$$S = \{x \mid x \text{ satisfies specified rules}\},$$

which can be translated as 'S is the set of values of x which satisfy the stated rules'. The rules occur after the vertical |. Thus

$$S = \{x \mid x \in \mathbb{N}^+ \text{ and } 2 \leqslant x \leqslant 8\}$$

is an alternative way of writing $S = \{2, 3, 4, 5, 6, 7, 8\}$. As another example,

$$S = \{x \mid x \in \mathbb{R} \text{ and } 0 \leqslant x \leqslant 1\}$$

is the closed interval $[0, 1]$, that is, *all* real numbers between 0 and 1 including 0 and 1.

Self-test 35.1

List in full the elements in the following sets:

(a) $S_1 = \{x \mid x \in \mathbb{N}^+ \text{ and } -2 \leqslant x \leqslant 8\}$,
(b) $S_2 = \{p/q \mid p \in \mathbb{N}^+, q \in \mathbb{N}^+, 1 \leqslant p \leqslant 3 \text{ and } 2 \leqslant q \leqslant 4\}$.

35.2 Equality, union, and intersection

Two sets A and B are said to be equal if they contain exactly the same elements. If this is the case, we write

$$A = B.$$

For example,

$$A = \{1, 2, 3\}, \qquad B = \{3, 2, 1\}, \qquad C = \{3, 1, 2, 1\}$$

are all equal, that is $A = B = C$. The order of the elements is immaterial, and repeated elements are discounted.

In a given context, the set of all elements of interest is known as the universal set, usually denoted by U. The definition of U depends on the context. For example, for the set A above, the universal set might be $\mathbb{N}$ (the set of all positive integers), or $\mathbb{R}^+$ (the set of all positive numbers), or some other set which includes $\{1, 2, 3\}$, depending on the particular application.

We now define how sets can be combined to create new sets. The union of two sets A and B is the set of all elements that belong to A, *or to* B, *or to both*. It is written as

$$A \cup B = \{x \mid x \in A \text{ or } x \in B \text{ or both}\},$$

and read as 'A union B'.

Example 35.1 *Find the union of*

$$A = \{x \mid x \in \mathbb{R} \text{ and } 0 \leqslant x \leqslant 2\} \quad \textit{and} \quad B = \{x \mid x \in \mathbb{R} \text{ and } 1 \leqslant x \leqslant 3\}.$$

The elements in the union have to belong to one or other of the intervals $0 \leqslant x \leqslant 2$, or $1 \leqslant x \leqslant 3$, or to both. The condition is satisfied by all numbers in the interval $0 \leqslant x \leqslant 3$, and by no others. Hence

$$A \cup B = \{ x \mid \mathbb{R} \text{ and } 0 \leqslant x \leqslant 3\}.$$

The intersection of two sets A and B is the set $A \cap B$ that contains *all elements common to both A and B*. It is written and defined by

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

Example 35.2 *Find the intersection of the sets A and B in Example* 35.1.

The elements in the intersection have to belong simultaneously to both intervals, that is to the overlapping part of the intervals $[0, 2]$ and $[1, 3]$, which is $[1, 2]$. Thus

$$A \cap B = \{x \mid x \in \mathbb{R} \text{ and } 1 \leqslant x \leqslant 2\}.$$

In the definitions of $A \cup B$ and $A \cap B$ above, we can see that the logical operation 'or' is associated with union, while 'and' is associated with intersection.

If A and B have no elements in common, then A and B are said to be disjoint. The set with no elements is called the empty set and denoted by $\varnothing$. Thus, if A and B are disjoint, then $A \cap B = \varnothing$. Thus if $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$ then $A \cap B = \varnothing$.

The complement of a set A is the set of all those elements which belong to the universal set U but do not belong to A. We denote this set by $\bar{A}$ (the notations A^c and A' are also frequently used): it will depend on the definition of U. Hence, the complement of A is, assuming that $x \in U$,

$$\bar{A} = \{x \mid x \notin A\}.$$

We say that A is a subset of B, expressed as $A \subseteq B$, if every element of A also belongs to the set B. It follows that $A \subseteq U$ if $B \subseteq U$. If there are elements of B which are not in A, then A is called a proper subset of B and written $A \subset B$. The statement $A \subseteq B$ includes the possibility that $A = B$, while $A \subset B$ does not. If $A \subseteq B$ and $B \subseteq A$, then all elements in A are contained in B, and vice versa; in other words, $A = B$.

The sets of integers $\mathbb{Z}$ and rational numbers $\mathbb{Q}$ are proper subsets of the real numbers $\mathbb{R}$, that is

$$\mathbb{Z} \subset \mathbb{R}, \qquad \text{and} \qquad \mathbb{Q} \subset \mathbb{R}.$$

We can summarize the results as follows.

Set operations

(a) **Union:** $A \cup B = \{x \mid x \in A \text{ or } x \in B \text{ or both}\}$.
(b) **Intersection:** $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$.
(c) **Complement:** $\bar{A} = \{x \mid x \notin A\}$.
(d) **Empty set:** $\varnothing$, the set with no elements.
(e) **Subset:** $A \subseteq B$ means that A is a **subset** of B.
(f) **Proper subset:** $A \subset B$ means that $A \subseteq B$ but $A \neq B$. (35.3)

Self-test 35.2

Find the union and intersection of $A = \{x \mid x \in \mathbb{R} \text{ and } -1 \leqslant x \leqslant 23\}$, $B = \{x \mid x \in \mathbb{N}^{+} \text{ and } 1 \leqslant x \leqslant 4\}$.

35.3 Venn diagrams

Useful graphical views and interpretations of sets and operations on them can be provided by Venn diagrams. We represent sets by regions in the plane, with the interpretation that the region stands for those elements belonging to the given set. The diagrams are symbolic: the set $A = \{1, 2\}$, for example, could be represented by the circle as shown in Fig. 35.1. Usually, sets are represented by the interiors of circles, but any closed curves can be used. In a given context, all the sets are subsets of a certain universal set U, whose nature will differ according to the context.

If the universal set is represented by a rectangle, then a subset A of U is represented by the interior of a circle within the rectangle shown in Fig. 35.2. This is a Venn diagram for U and A. Remember that A could represent an infinite number of elements, or one element, or be the empty set $\varnothing$. Two regions in U may have elements in common. For example, $A = \{1, 2, 3\}$ and $B = \{3, 4, 5\}$ have the common element 3. In a Venn diagram this is represented by intersecting regions A and B as in Fig. 35.3a.

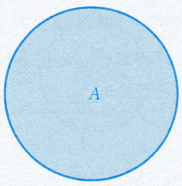


Fig. 35.1

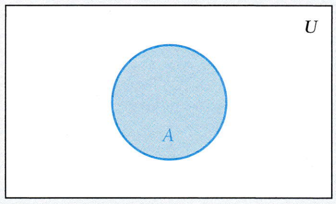


Fig. 35.2 Venn diagram for the universal set U and a set A.

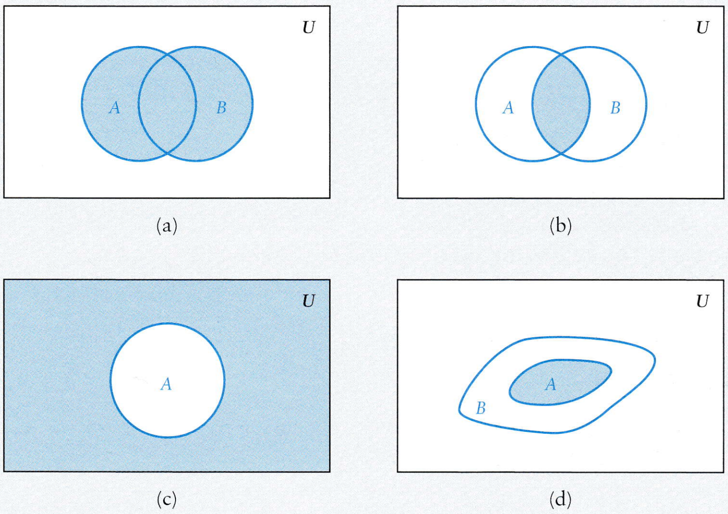


Fig. 35.3 (a) Union $A \cup B$. (b) Intersection $A \cap B$. (c) Complement $\bar{A}$. (d) Proper subset $A \subset B$.

The union, intersection, complement, and proper subset can be represented by the Venn diagrams shown in Fig. 35.3. The shaded regions indicate the elements defined by the operations.

From the definitions of union, intersection, and complement, or from Venn diagrams, the following laws of the algebra of sets can be deduced:

(a) Algebra of sets

$$A \cup A = A, \qquad A \cap A = A.$$

(b) Commutative laws:

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A.$$

(c) Associative laws (see Fig. 35.4):

$$(A \cup B) \cup C = A \cup (B \cup C),$$
$$(A \cap B) \cap C = A \cap (B \cap C).$$

↗

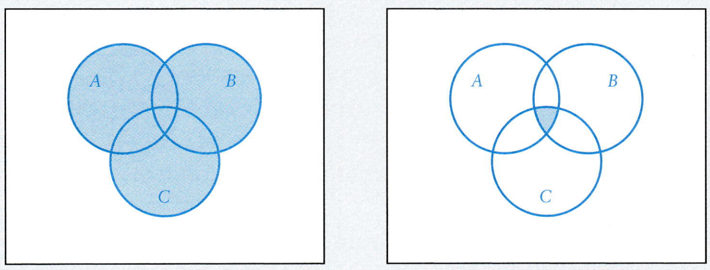


Fig. 35.4 (a) $(A \cup B) \cup C$ or $A \cup (B \cup C)$. (b) $(A \cap B) \cap C$ or $A \cap (B \cap C)$.

(d) **Distributive laws:**

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \qquad (35.4)$$

Sets also satisfy the following identity and complementary laws:

Identity laws: $A \cup \varnothing = A, \qquad A \cap \boldsymbol{U} = A.$

Complementary laws:

$$A \cup \bar{A} = \boldsymbol{U}, \qquad A \cap \bar{A} = \varnothing, \qquad \bar{\bar{A}} = A. \qquad (35.5)$$

For example, $\bar{A}$ consists of all elements that do not belong to A, and none that do; so there are no elements common to A and $\bar{A}$. Therefore $A \cap \bar{A} = \varnothing$.

The difference of the sets A and B, written as $A \backslash B$, consists of the set of those elements that belong to A but do not belong to B. Thus

$$A \backslash B = \{x \mid x \in A \text{ and } x \notin B\} \quad \text{or} \quad \overline{A \cap B} \cap A.$$

(The notation $A - B$ is also used for $A \backslash B$.) Figure 35.5b shows a Venn diagram for $A \backslash B$.

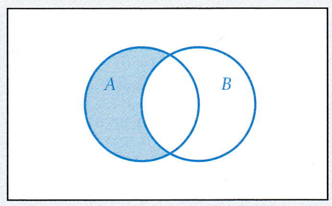


Fig. 35.5 Venn diagram for the difference $A \backslash B$ (shaded).

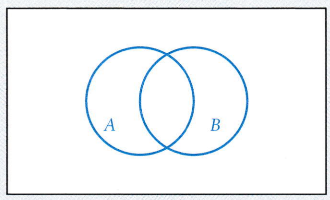


Fig. 35.6

Example 35.3 *Using Fig. 35.6 as the Venn diagram of two sets A and B, mark by shading the following sets:*

(a) $A \cup \bar{B}$, (b) $A \cap \bar{B}$, (c) $\bar{A} \cap \bar{B}$, (d) $\bar{A} \cup \bar{B}$, (e) $\overline{A \cup B}$, (f) $\overline{A \cap B}$.

Venn diagrams of the sets are shown in Fig. 35.7.

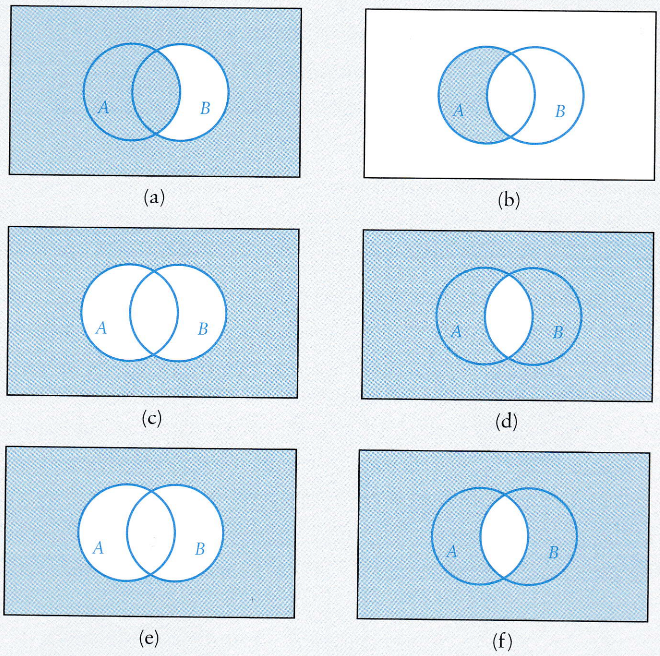


Fig. 35.7

Examples 35.3c, e and 35.3d confirm de Morgan's laws, which are

De Morgan's laws

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \qquad \overline{A \cap B} = \bar{A} \cup \bar{B}. \qquad (35.6)$$

Example 35.4 *Using Fig. 35.8 as the Venn diagram of three sets A, B, and C, shade the following sets:*

(a) $(A \cap B) \cup C$, (b) $(A \cap B) \cap C$, (c) $(A \cap B) \cap (A \cap C)$, (d) $(A \cup B) \cup (A \cap C)$.

The required sets are shown in Fig. 35.9. (Figs. 35.9b, c confirm that $(A \cap B) \cap C = (A \cap B) \cap (A \cap C)$.)

Example 35.4 continued

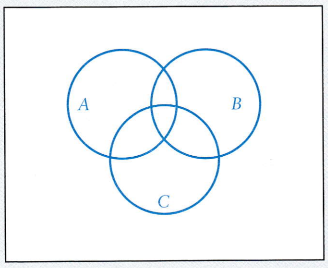


Fig. 35.8

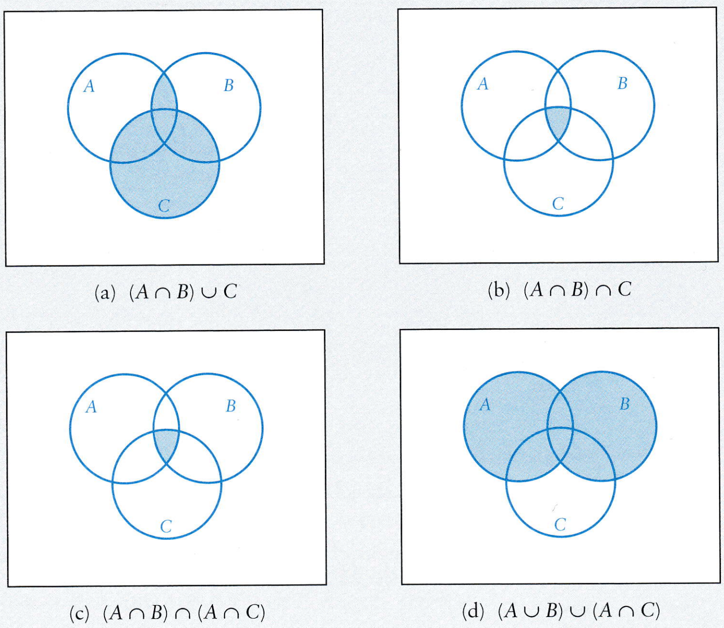


Fig. 35.9

Example 35.5 *Show that* $(A \cap B) \cup (A \cap \bar{B}) = A$.

By the distributive law (35.4),

$$
\begin{aligned}
(A \cap B) \cup (A \cap \bar{B}) &= A \cap (B \cup \bar{B}) \\
&= A \cap U \quad \text{(by the complementary law)} \\
&= A \quad \text{(by the identity law).}
\end{aligned}
$$

This can be confirmed graphically by drawing a Venn diagram.

Example 35.6 *Show that* $(A \cup B) \cup (A \backslash B) = A \cup B$.

From Fig. 35.5, we can observe that $A \backslash B = A \cap \bar{B}$. Hence

$$
\begin{aligned}
(A \cup B) \cup (A \backslash B) &= (A \cup B) \cup (A \cap \bar{B}) \\
&= A \cup (B \cup (A \cap \bar{B})) \quad \text{(associative law)} \\
&= A \cup ((B \cup A) \cap (B \cup \bar{B})) \quad \text{(distributive law)} \\
&= A \cup ((B \cup A) \cap U) \\
&= A \cup (B \cup A) \quad \text{(identity law)} \\
&= (B \cup A) \cup A \quad \text{(commutative law)} \\
&= B \cup (A \cup A) = B \cup A \\
&= A \cup B.
\end{aligned}
$$

Alternatively, and more intuitively, we may notice that, since $A \backslash B$ is a subset of A, it is therefore also a subset of $A \cup B$, and so adds nothing to $A \cup B$ when united with it.

Example 35.7 *In a manufacturing process, a product passes through three production stages and is given a quality check at all three stages, which it either passes or fails. Let P_i represent the set of products passing the quality check at stage i. Draw a Venn diagram of the process. Interpret the quality failures of the products in the sets given by* $\bar{P}_1$, $P_2 \backslash (P_1 \cup P_3)$, and $(P_1 \cup P_2) \cap P_3$. *What set represents the completely satisfactory products?*

A production run of 1000 *occurs, of which* 8 *fail all stages,* 20 *pass only stage* P_1, 31 *only stage* P_2, *and* 17 *only stage* P_3; 814 *pass stages* P_1 and P_2, 902 *stages* P_2 *and* P_3, *and* 800 *stages* P_3 *and* P_1. *Determine the final number which pass all quality checks.*

$\bar{P}_1$ represents all products which fail the P_1 quality check.
$P_2 \backslash (P_1 \cup P_3)$ represents those products which pass only P_2 stage.
$(P_1 \cup P_2) \cap P_3$ represents those products which are satisfactory at stages P_3 and P_1 or P_2.
The set $P_1 \cap P_2 \cap P_3$ represents those products which are satisfactory at all stages.

The numbers associated with each subset of the universal set U are shown in Fig. 35.10. Since 8 fail all quality checks, then the number of elements in $P_1 \cup (P_2 \cup P_3)$ is 992. In the figure, k represents the number of products which pass all the quality checks. Hence $800 - k$, for example, represents those products which are satisfactory in stages P_1 and P_2, but fail in P_3. Thus $P_1 \cup P_2 \cup P_3$ contains

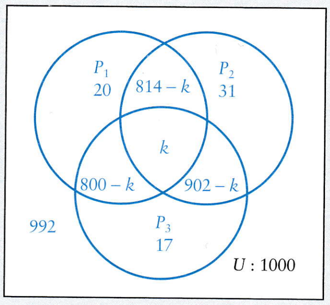


Fig. 35.10

Example 35.7 continued

$$992 = 20 + 31 + 17 + (814 - k) + (902 - k) + (800 - k) + k$$

products. Hence $992 = 2584 - 2k$, and so

$$k = 796.$$

Of the 1000 products manufactured, 796 passed all the quality checks.

In the previous Example, we are really interested in the *numbers* of elements in each of the sets. For example, the number of elements in U is 1000 and the number of elements in $P_2 \backslash (P_1 \cup P_3)$, those products which pass only stage 2, is 31. We write

$$\mathrm{n}(U) = 1000, \qquad \mathrm{n}[P_2 \backslash (P_1 \cup P_3)] = 31.$$

The number of elements in the set S is $\mathrm{n}(S)$: this number is known as the cardinality of S. Many sets can have infinite cardinality. For example, $\mathrm{n}(\mathbb{Q})$, where $\mathbb{Q}$ is the set of rational numbers, is an infinite number. We write $\mathrm{n}(\mathbb{Q}) = \infty$. The empty set $\varnothing$ has no elements: hence $\mathrm{n}(\varnothing) = 0$.

The following results apply to *finite sets*. If two finite sets A and B are disjoint, then they have no elements in common. It follows that

$$\mathrm{n}(A \cup B) = \mathrm{n}(A) + \mathrm{n}(B).$$

This result applies to any number of disjoint sets. It is clear that they must be disjoint, since otherwise elements would be counted more than once.

This last result is also a useful method of counting elements when combined with a Venn diagram. Consider just two sets A and B as shown in Fig. 35.11. The sets representing each of the subsets in the Venn diagram $A \backslash B$, $A \cap B$, and $B \backslash A$ are shown in Fig. 35.11. Since these sets are disjoint, then we can obtain a formula for the number of elements in the union of A and B, namely

$$\mathrm{n}(A \cup B) = \mathrm{n}(A \backslash B) + \mathrm{n}(A \cap B) + \mathrm{n}(B \backslash A). \tag{35.7}$$

For sets A and B separately,

$$\mathrm{n}(A) = \mathrm{n}(A \backslash B) + \mathrm{n}(A \cap B), \qquad \mathrm{n}(B) = \mathrm{n}(B \backslash A) + \mathrm{n}(A \cap B). \tag{35.8}$$

Elimination of $\mathrm{n}(A \backslash B)$ and $\mathrm{n}(B \backslash A)$ between (35.7) and (35.8) leads to the alternative result

$$\mathrm{n}(A \cup B) = \mathrm{n}(A) + \mathrm{n}(B) - \mathrm{n}(A \cap B).$$

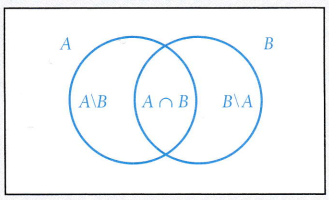


Fig. 35.11 Counting elements in the union of two sets.

For three finite sets A, B, and C the corresponding result is

$$\mathrm{n}(A\cup B\cup C)=\mathrm{n}(A)+\mathrm{n}(B)+\mathrm{n}(C)+\mathrm{n}(A\cap B\cap C)-\mathrm{n}(B\cap C)$$
$$-\,\mathrm{n}(C\cap A)-\mathrm{n}(A\cap B).$$

This result can be constructed from the Venn diagram.

Further discussion of sets and their algebra can be found in Garnier and Taylor (1991).

Self-test 35.3

In Fig. 35.8, shade
(a) $A\cap(B\cap C)$; (b) $A\cup B$; (c) $(A\cup C)\cap(B\cup C)$.

Problems

35.1 (Section 35.1). List the elements in the following sets:
(a) $S=\{x\,|\,x\in\mathbb{N}^+ \text{ and } 3\leqslant x\leqslant 10\}$;
(b) $S=\{x\,|\,x\in\mathbb{N}^+ \text{ and } -2\leqslant x\leqslant 4\}$;
(c) $S=\{x\,|\,x\in\mathbb{Z} \text{ and } -2\leqslant x\leqslant 4\}$;
(d) $S=\{x\,|\,x\in\mathbb{N}^+ \text{ or } \mathbb{N}^-, \text{ and } -2\leqslant x\leqslant 4\}$;
(e) $S=\{1/x\,|\,x\in\mathbb{N}^+ \text{ and } 3\leqslant x\leqslant 8\}$;
(f) $S=\{x^2\,|\,x\in\mathbb{N}^+ \text{ and } |x|\leqslant 3\}$;
(g) $S=\{x+\mathrm{i}y\,|\,x\in\mathbb{N}^+, y\in\mathbb{N}^+, 1\leqslant x\leqslant 4, 2\leqslant y\leqslant 5\}$.

35.2 (Section 35.3). Show on Venn diagrams the following sets:
(a) $A\cup\bar{B}$; (b) $\bar{A}\cap\bar{B}$;
(c) $A\cap(B\cup C)$; (d) $(A\cap B)\cup(B\cap C)$;
(e) $\overline{A\cap B}$;
(f) $(A\backslash B)\cap C$;
(g) $A\backslash(B\cap C)$;
(h) $\overline{(A\backslash B)\cup(B\backslash C)}$.

35.3 (Section 35.2). Determine the union $A\cup B$ of each of the following pairs of sets A and B:
(a) $A=\{x\,|\,x\in\mathbb{R} \text{ and } -1\leqslant x\leqslant 2\}$, $B=\{x\,|\,x\in\mathbb{R} \text{ and } -1\leqslant x\leqslant 4\}$;
(b) $A=\{x\,|\,x\in\mathbb{R} \text{ and } -1\leqslant x< 0\}$, $B=\{x\,|\,x\in\mathbb{R} \text{ and } 0< x< 1\}$;
(c) $A=\{1,2,3,4\}$, $B=\{-4,-3,-2,-1\}$;
(d) $A=\{y\,|\,y=\cos x, x\in\mathbb{R}, \text{ and } 0\leqslant x\leqslant \tfrac{1}{2}\pi\}$, $B=\{y\,|\,y=\sin x, x\in\mathbb{R}, \text{ and } -\tfrac{1}{2}\pi\leqslant x\leqslant \tfrac{1}{2}\pi\}$.

35.4 (Section 35.2). Determine the intersections $A\cap B$ of the following sets:
(a) $A=\{x\,|\,x\in\mathbb{R}, \text{ and } -2\leqslant x\leqslant 1\}$, $B=\{x\,|\,x\in\mathbb{R}, \text{ and } -1\leqslant x\leqslant 2\}$;
(b) $A=\{x\,|\,x\in\mathbb{N}^+ \text{ and } -5\leqslant x\leqslant 2\}$, $B=\{x\,|\,x\in\mathbb{R}, \text{ and } -5\leqslant x\leqslant 2\}$;
(c) $A=\{n\,|\,n=1/m \text{ and } m\in\mathbb{N}^+\}$, $B=\{n\,|\,n=1/m^2 \text{ and } m\in\mathbb{N}^+\}$;
(d) $A=\{x\,|\,x\in\mathbb{R} \text{ and } x^2-3x+2=0\}$, $B=\{x\,|\,x\in\mathbb{R} \text{ and } 2x^2+x-3=0\}$;
(e) $A=\{x\,|\,x\in\mathbb{R} \text{ and } |x|\leqslant 2\}$, $B=\{x\,|\,x\in\mathbb{R} \text{ and } |x-1|\leqslant 1\}$.

35.5 (Section 35.3). Construct a set formula for the shaded sets of Fig. 35.12:

(a)
(b)

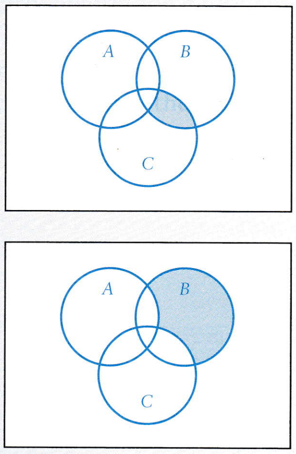
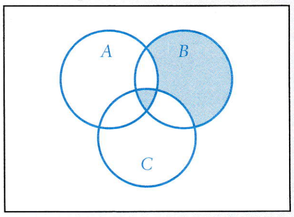


Fig. 35.12

(c)



(d)

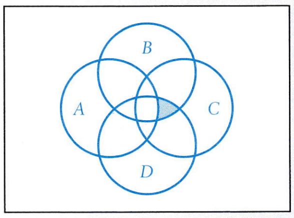


Fig. 35.12 (*continued*)

35.6 The set S consists of products, each of which is given n pass/fail tests, numbered 1 to n. The set S_r, consists of those products that pass test r. What is the set of products that
(a) fails all tests, (b) fails only test 1,
(c) fails some tests?

35.7 At Keele University, all first-year students must take three subjects of which at least one must be a science subject, and at least one must be a humanities or social science subject. Let A be the set of all first-year students in a given year, A_1 the set of students who take exactly one science subject, B_1 the set of students who take just one humanities subject, and B_2 the set of those who take two social science subjects. Draw a Venn diagram to represent the different sets of students classified by groups of subjects. Give set formulae for students who take
(a) just one social science subject,
(b) no humanities subject,
(c) one subject from each group.

35.8 (Section 35.3). The rules listed in (35.4) illustrate the duality principle which states that every statement involving sets which is true *for all sets* has a dual in which $\cup$ and $\cap$ are interchanged, and $\varnothing$ and U are interchanged everywhere.

Use Venn diagrams to establish the following:
(a) $(A\backslash B) \cap C = (A \cap C)\backslash B$;
(b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
What are their dual identities?

35.9 Three sets A, B, and C satisfy

$$A \cap B \cap C = (A \cap C) \cup (B \cap C).$$

Explain why the duality principle of Problem 34.8 does not apply. What condition of the duality principle is violated?

35.10 The cartesian product of two sets A and B is the set of all ordered pairs $\{(a, b)\}$, where $a \in A$ and $b \in B$. It is written as

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

If $A = B$, then we write $A \times A = A^2$. Let $A = \{1, 2\}$ and $B = \{1, 2, 3\}$; write down all the elements in the sets $A \times B$, $B \times A$, A^2, and B^2.

35.11 The cartesian product extends to the products of three or more sets. Thus

$$A \times B \times C = \{(a, b, c) \mid a \in A \text{ and } b \in B \text{ and } c \in C\}.$$

Let $A = \{1, 2, 3\}$, $B = \{0, 1\}$, and $C = \{1, 2\}$. Write down all the elements in

$$A \times B \times C,\ A^2 \times C,\ (A \cup B) \times C,\ (A \cap B) \times C.$$

35.12 At the end of a production process, 500 electrical components pass through three quality checks P, Q, and R. It is found that 38 components fail check P, 29 fail Q, 30 fail R, 7 fail P and Q, 5 fail Q and R, 8 fail R and P, and 3 fail all checks. Determine how many components:
(a) pass all checks,
(b) fail just one check,
(c) fail just two checks.

35.13 (Section 35.4). For three finite sets A, B, and C, show that the number of elements in the union of the sets is given by

$$\begin{aligned} n(A \cup B \cup C) &= n(A) + n(B) + n(C) \\ &\quad + n(A \cap B \cap C) - n(B \cap C) \\ &\quad - n(C \cap A) - n(A \cap B). \end{aligned}$$

35.14 If A and B are two finite sets, explain why, for the cartesian product (defined in Problem 35.10 above),

$$n(A \times B) = n(A)n(B).$$

35.15 The menu in a restaurant contains three courses: 4 starters (set A), 5 main courses (set B), and 3 sweets (set C). Customers can choose either the full menu or, alternatively, a main course and a sweet. In terms of cartesian products what is the set of all possible meals (the answer is really a *set of pairs and triples*). For how many different orders can customers ask?

35.16 Given $A = \{1, 2, 3\}$, $B = \{3, 4\}$, and $C = \{2, 3, 4, 5\}$, find the elements in the sets $B \cup C$, $B \cap C$, and the cartesian products $A \times B$ and $A \times C$. Verify that

$$A \times (B \cup C) = (A \times B) \cup (A \times C),$$

$$A \times (B \cap C) = (A \times B) \cap (A \times C).$$

(This example suggests general results which are true for all sets.)

Boolean algebra: logic gates and switching functions

36

CONTENTS

We are now going to present some new operations between special entities. They have analogies with ordinary addition and multiplication, and the symbols for them will be similar; but not the same, since we need to emphasize that these are Boolean operations. The algebra involved is named after George Boole (1815–64) who first developed the modern ideas of symbolic logic. Boolean algebra has applications in logic and switching circuits.

36.1 Laws of Boolean algebra

Consider a set B which consists of just two elements 0 and 1, that is $B = \{0, 1\}$. We shall denote the sum of two elements a and b of B by $a \oplus b$ (the notations $\vee$, $\cup$, and $+$, and the alternative term join are also used); we denote the product of the two elements by $a * b$ (the notations $\wedge$, $\cap$, $\times$, and $\cdot$, or simply ab, and the alternative term meet are also in use) and the complement of a by $\bar{a}$ ($\sim a$ and $\neg a$ are used in logic). These binary operations applied to the members of B are defined to give the elements shown in Table 36.1.

Table 36.1 *Binary operations*

Sum

a	b	$a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	1

Product

a	b	$a * b$
0	0	0
0	1	0
1	0	0
1	1	1

Complement

a	$\bar{a}$
0	1
1	0

Thus, for example

$$0 \oplus 1 = 1, \quad 1 \oplus 1 = 1, \qquad 0 * 0 = 0, \quad 1 * 1 = 1, \qquad \bar{0} = 1, \quad \bar{1} = 0.$$

The elements of B are known as Boolean variables. We have restricted our set B to one with just two elements or binary digits, because this is the main application in circuits and computer design, but definitions can be interpreted for more general sets. A Boolean algebra is a set with the operations $\oplus$, $*$, and $\bar{\ }$ defined on it, together with the following laws on any elements a, b, c which belong to B:

Commutative laws:

$$a \oplus b = b \oplus a, \qquad a * b = b * a;$$

Associative laws:

$$a \oplus (b \oplus c) = (a \oplus b) \oplus c, \qquad a * (b * c) = (a * b) * c;$$

Distributive laws:

$$a * (b \oplus c) = (a * b) \oplus (a * c),$$
$$a \oplus (b * c) = (a \oplus b) * (a \oplus c). \qquad (36.1)$$

In addition, the set must contain distinct identity elements 0 and 1 for the operations $\oplus$ and $*$ respectively. For these elements we must have the identity and complement laws:

Identity laws:

$$a \oplus 0 = a, \qquad a * 1 = a$$

Complement laws:

$$a \oplus \bar{a} = 1, \qquad a * \bar{a} = 0 \qquad (36.2)$$

To summarize, we can say that a Boolean algebra consists of the collection

$$(B, \oplus, *, \bar{\ }, 0, 1),$$

in other words, a set B, the binary operations $\oplus$ and $*$, the complement $\bar{\ }$, and the identity elements 0 and 1.

In our case $B = \{0, 1\}$, the binary set, which consists simply of identity elements. We can check that the definitions in Table 36.1 satisfy the laws in (36.1). They are essentially the laws of set operations with sum $\oplus$ and product $*$ replacing union $\cup$ and intersection $\cap$, and with 1 replacing the universal set U and 0 the empty set $\varnothing$.

Just as with sets, we can deduce further laws, some of which are included in (36.3):

Absorption laws:

$$a \oplus (a * b) = a, \qquad a * (a \oplus b) = a;$$

de Morgan's laws: $\overline{a \oplus b} = \bar{a} * \bar{b}, \qquad \overline{a * b} = \bar{a} \oplus \bar{b};$

Identity laws:

$$1 \oplus a = a \oplus 1 = 1, \qquad 0 * a = a * 0 = 0;$$

Reflexive law: $\bar{\bar{a}} = a. \qquad (36.3)$

Note that $*$ *takes precedence over* $\oplus$ in the absence of brackets. Thus, in the first absorption law, $a \oplus a * b$ means $a \oplus (a * b)$; in the second absorption law, the brackets are essential.

We will prove one of the absorption laws to illustrate how proofs are approached in Boolean algebra.

Example 36.1 *Prove that* $a \oplus a * b = a$.

For all $a,b \in B$

$$\begin{aligned} a \oplus a * b &= a * 1 \oplus a * b \quad \text{(identity law)} \\ &= a * (1 \oplus b) \quad \text{(distributive law).} \end{aligned}$$

Now

$$\begin{aligned} 1 \oplus b &= (1 \oplus b) * 1 \quad \text{(identity law)} \\ &= 1 * (b \oplus 1) \quad \text{(associative law)} \\ &= (b \oplus \bar{b}) * (b \oplus 1) \quad \text{(complement law)} \\ &= b \oplus \bar{b} * 1 \quad \text{(distributive law)} \\ &= b \oplus \bar{b} \quad \text{(identity law)} \\ &= 1 \quad \text{(complement law).} \end{aligned}$$

Finally

$$a \oplus a * b = a * 1 = a.$$

36.2 Logic gates and truth tables

Any expression made up from the elements of B and the operations $\oplus$, $*$, and $^{-}$ is known as a Boolean expression. For example,

$$a \oplus b, \quad a \oplus \bar{b}, \quad a \oplus \bar{a} * b,$$

are Boolean expressions. For the binary set, the elements 1 and 0 can represent 'on' or 'off' states in digital circuits. The basic components in a computer are logic gates which can produce an output from inputs. All the outputs and inputs can be in one of two states, usually either low voltage (0) or high voltage (1).

The fundamental Boolean operations of $\oplus$, $*$, and $^{-}$ correspond to devices known respectively as the OR gate, AND gate, and NOT gate. As with circuit components such as resistance and inductance, each has its own symbol.

The OR gate has two inputs and a single output represented by the symbol in Fig. 36.1. The output is $f = a \oplus b$. The inputs a and b can each take either of the values 0 or 1. Hence there are four possible inputs into the device as listed in Table 36.2. The final column f can be completed using the sum rule in Table 36.1. Then, if a is 'on' (1) and b is 'off' (0), the output f is 'on' (1). Table 36.2 is known as the truth table of the OR gate.

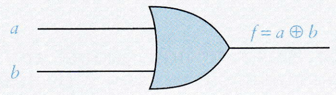


Fig. 36.1 The OR gate.

Table 36.2 *Truth table for the* OR *gate*

a	b	$f = a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	1

The symbol and truth table for the AND gate are shown in Fig. 36.2 and Table 36.3. Again the device has two inputs and the single output $f = a * b$, the product of a and b.

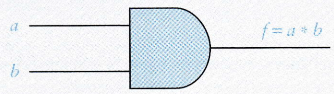


Fig. 36.2 The AND gate.

Table 36.3 *Truth table for the* AND *gate*

a	b	$f = a * b$
0	0	0
0	1	0
1	0	0
1	1	1

Finally the NOT gate is shown in Fig. 36.3 with its truth table given as Table 36.4. The NOT gate has a single input and a single output which is the complement of its input.

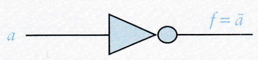


Fig. 36.3 The NOT gate.

Table 36.4 *Truth table for the* NOT *gate*

a	$f = \bar{a}$
0	1
1	0

There is further jargon associated with these gates. The output $a \oplus b$ is known as the disjunction of a and b, while $a * b$ is known as the conjunction of a and b, and $\bar{a}$ is called the negation of a.

These devices can be connected in series and parallel to create new logic devices, each of which will have its own truth table.

A series connection between a NOT gate and an AND gate is shown in Fig. 36.4a. The output $a * b$ of the AND gate becomes the input of the NOT gate which results in the output $\overline{a * b}$. This combined device is known as the NAND gate, and it has its own symbolic representation shown in Fig. 36.4b. Its truth table is given in Table 36.5.

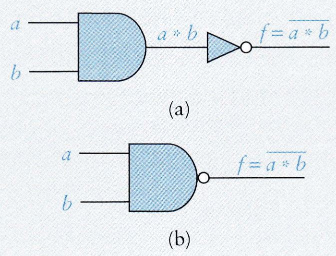


Fig. 36.4 The NAND gate.

Table 36.5 *Truth table for the* NAND *gate*

a	b	$f=\overline{a * b}$
0	0	1
0	1	1
1	0	1
1	1	0

A series connection between a NOT gate and an OR gate produces the NOR gate as shown in Fig. 36.5a. The output f is the complement of the sum of a and b. The NOR gate also has its own symbol contraction shown in Fig. 36.5b. It has the truth table shown in Table 36.6.

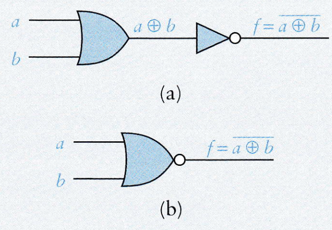


Fig. 36.5 The NOR gate.

Table 36.6 *Truth table for the* NOR *gate*

a	b	$f=\overline{a \oplus b}$
0	0	1
0	1	0
1	0	0
1	1	0

Self-test 36.1

The output of an AND gate (Fig. 36.2) is attached to a NOT gate (Fig. 36.3). Construct the truth table for the system.

36.3 Logic networks

The five gates introduced in the previous section can be linked in series and parallel combinations to create further logic networks. Some examples are presented here.

Example 36.2 *Construct the Boolean expression for the output f of the device shown in Fig. 36.6.*

Starting from the left in Fig. 36.6, the upper AND gate produces an output $a * b$ and the lower OR gate has an output $c \oplus d$. These become the inputs into the OR gate on the right. Hence the final output is

Example 36.2 continued

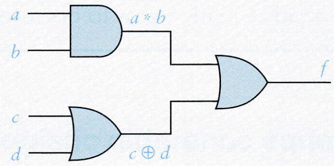


Fig. 36.6

$$f = (a * b) \oplus c \oplus d.$$

Since there are four inputs, the output f can be determined for each of the $2^4 = 16$ possible inputs. Hence if, for example, $a = 1$, $b = 0$, $c = 0$, $d = 1$, then the output $f = 1$.

Example 36.3 *Figure 36.7 shows a logical network with three inputs a, b, c, and four devices. Find a Boolean expression for the output f. Write down the truth table for the system.*

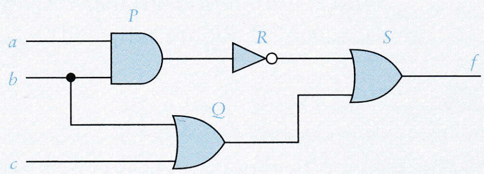


Fig. 36.7

Note that the input b is the same in both devices P and Q. The output from the AND gate P is $a * b$, and the output from R is $\overline{a * b}$. The output from Q is $b \oplus c$. Hence the inputs $\overline{a * b}$ and $b \oplus c$ into S produce an output

$$f = \overline{a * b} \oplus b \oplus c.$$

The truth table for this network is given in Table 36.7. Whatever the inputs, the device is always 'on'.

Table 36.7

a	b	c	$\overline{a * b}$	$b \oplus c$	$\overline{a * b} \oplus (b \oplus c)$
0	0	0	0	1	1
0	0	1	0	1	1
0	1	0	0	1	1
0	1	1	0	1	1
1	0	0	0	1	1
1	0	1	0	1	1
1	1	0	1	0	1
1	1	1	1	0	1

Example 36.4 *Show that, using just the* NOR *gate, it is possible to build a logic network to model any Boolean expression.*

Given inputs a and b, we have to show that devices can be constructed using just NOR gates with outputs of $a \oplus b$, $a * b$, and $\bar{a}$. For inputs of a and b, the single NOR gate generates an output of $\overline{a \oplus b}$. Figure 36.8 shows three devices which simulate the required outputs.

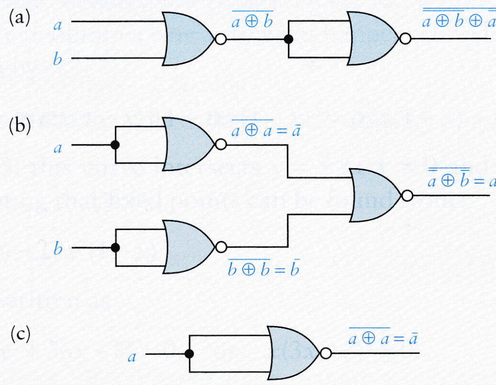


Fig. 36.8 The simulations are: (a) OR gate; (b) AND gate (c) NOT gate.

Example 36.5 *Design a logic network using* OR, AND, *and* NOT *gates to reproduce the Boolean expression* $f = a * \bar{b} \oplus a$ *for inputs* a *and* b.

From input b we obtain $\bar{b}$ by a NOT gate. The inputs a and $\bar{b}$ are then fed into an AND gate to produce $a * \bar{b}$. Finally a spur from the a input and the $a * \bar{b}$ output are fed into an OR gate as shown in Fig. 36.9.

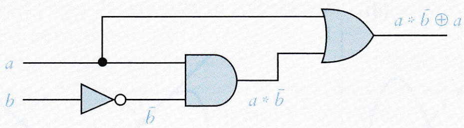


Fig. 36.9

Self-test 36.2

An AND gate with inputs a and b, and a NOT gate with input c are connected to a NOR gate. Find a Boolean expression for the output f, and construct a truth table for the system.

36.4 The inverse truth-table problem

In this problem we attempt the inverse problem; of creating a Boolean expression for a given truth table. For example, Table 36.8 is a truth table for two inputs a and b. We illustrate a method for the construction of a Boolean expression which will generate this truth table. Pick out cases for which $f=1$. For the case $a=0, b=1$, write down $\bar{a} * b$, and for $a=1$, $b=0$ write down $a * \bar{b}$, using in the products, the *complement of any zero element*. Thus, for example, if $a=0$ and $b=1$, then $\bar{a}=1$ and $\bar{a} * b=1$. Similarly $a * \bar{b}=1$. Hence by Table 36.3

$$\bar{a} * b \oplus a * \bar{b} = 1$$

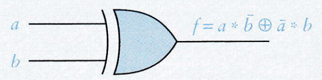


Fig. 36.10 The exclusive-OR gate.

Table 36.8 *Truth table for EXOR gate*

a	b	f
0	0	0
0	1	1
1	0	1
1	1	0

for these cases. f remains zero for the remaining outputs. We obtain

$$f=\bar{a} * b \oplus a * \bar{b}, \tag{36.4}$$

and the final output f can be checked.

This particular gate is known as the exclusive-OR gate, or EXOR gate, and has its own symbol shown in Fig. 36.10. This form of f obtained by the construction just described is known as the disjunctive normal form. By the definitions in Table 36.1, the construction *guarantees* a Boolean expression for any truth table.

Applied to the truth table for the OR gate (Table 36.2), the disjunctive form gives

$$f=(\bar{a} \oplus b) \oplus (a \oplus \bar{b}) \oplus (a \oplus b),$$

which is evidently a more complicated version of $a \oplus b$.

The method can be applied to more complex truth tables. Table 36.9 shows an output for three inputs. The output 1 appears in rows 2, 4, 5, 7, 8. In row 2, $a=0$,

Table 36.9

a	b	c	f
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

$b = 0$ and $c = 1$. Hence we introduce $\bar{a} * \bar{b} * c$ which equals 1. Apply the same procedure to rows 4, 5, 7, 8 introducing the complement for zero Boolean variables. The disjunctive normal form for a corresponding Boolean expression is, following the rules for products of elements and their complements,

$$f = \bar{a} * \bar{b} * c \oplus \bar{a} * b * c \oplus a * \bar{b} * \bar{c} \oplus a * b * \bar{c} \oplus a * b * c.$$

Check that f does give the required output. The disjunctive normal form always guarantees an answer, but it is not necessarily the simplest or most efficient in circuit architecture.

Self-test 36.3

Construct a Boolean expression for the truth table

a	b	$\overline{a * b}$
0	0	1
0	1	1
1	0	1
1	1	0

using the disjunctive normal form. Compare the answer with the answer of Self-test 36.2.

36.5 Switching circuits

A circuit of on–off switches can also be represented by Boolean expressions. For example, Fig. 36.11 shows a simple on–off switch in part of a circuit. Current flows if the switch S is in the *on* or closed position ($a = 1$), and does not flow if the switch is in the *off* or open position ($a = 0$). The variable a represents the state of the switch.

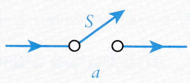


Fig. 36.11 On–off switch.

Consider two switches S_1 and S_2 in series (Fig. 36.12). Current only flows if both switches are closed, that is when $a_1 = 1$ and $a_2 = 1$, where a_1 and a_2 represent the states of the switches. Hence the truth table for the series switches is as shown in Table 36.10. Thus the state of current flow is given by $f = a * b$, the product of a and b.

Similarly two switches in parallel (Fig. 36.13) correspond to the sum of a and b. The truth table is given in Table 36.11. The final column indicates that $f = a \oplus b$.

The complement of a, the state of switch S_1, is another switch S_2 in the circuit which is always in the complementary state to S_1, off when S_1 is on and vice versa.

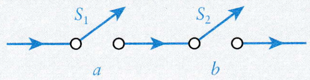


Fig. 36.12 Two switches in series.

Table 36.10 *Truth table for two switches in series*

a	b	f
0	0	0
0	1	0
1	0	0
1	1	1

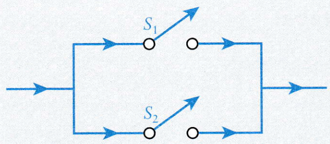


Fig. 36.13 Two switches in parallel.

Table 36.11 *Truth table for two switches in parallel*

a	b	f
0	0	0
0	1	1
1	0	1
1	1	1

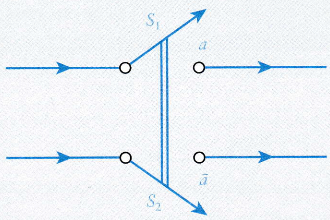


Fig. 36.14 Complement of a switch using a rigid tie.

It can be represented symbolically by Fig. 36.14, in which the switches S_1 and S_2 are joined by a rigid tie.

These devices are analogous to the gates of Section 36.3. For switching circuits, the Boolean expressions are often referred to as switching functions.

Example 36.6 *Find a switching function f for the system shown in Fig. 36.15.*

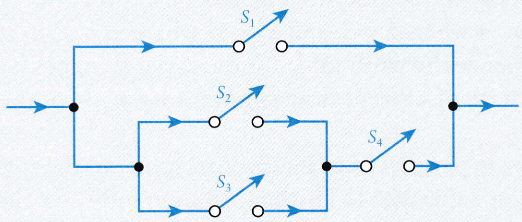


Fig. 36.15

↗

Example 36.6 continued

Let a_1, a_2, a_3, a_4 represent respectively the states of each switch S_1, S_2, S_3, S_4. Since S_2 and S_3 are in parallel, their output will be $a_2 \oplus a_3$. This combined in series with a_4 will give an output of $(a_2 \oplus a_3) * a_4$. In turn, this is in parallel with S_1. Hence, the final output is

$$(a_2 \oplus a_3) * a_4 \oplus a_1.$$

Example 36.7 *A light on a staircase is controlled by two switches S_1 and S_2, one at the bottom of the stairs and one at the top. Switches can be separately 'up' or 'down'. If both switches are up, the light is off. Either switch changed to down switches the light on, and any subsequent change to a switch alters the state of the light. Design a truth table for the circuit.*

The truth table is shown in Table 36.12, where the state of S_i $(i = 1, 2)$ is $a_i = 0$ when the switch is up (off) and $a_i = 1$ when the switch is down (on). The light on is $f = 1$, and the light off is $f = 0$. This truth table is the same as that for the exclusive-OR gate in Section 36.4. Hence, from (36.4), the circuit can be represented by the switching function

$$f = a_1 * \bar{a}_2 \oplus \bar{a}_1 * a_2.$$

The actual circuit is shown in Fig. 36.16, where S_1 and S_2 are one-pole two-way switches. At S_1, the state a_1 represents the switch 'up' and its complement $\bar{a}_1$ is the switch down. A similar state operates at S_2.

Table 36.12

Switch S_1	Switch S_2	Light	a_1	a_2	f
up	up	off	0	0	0
down	up	on	1	0	1
down	down	off	1	1	0
up	down	on	0	1	1

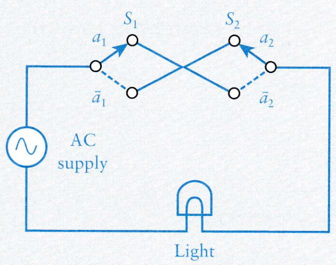


Fig. 36.16 Two-switch light control.

Further explanation of Boolean algebra with many applications to switching circuits can be found in Garnier and Taylor (1991).

Problems

36.1 Read through Example 36.1. Now prove the other absorption law:

$$a * a \oplus b = a.$$

(Example 36.1 and this result illustrate the duality principle, which states that any theorem which can be proved in Boolean algebra implies another theorem with $*$ and $\oplus$ interchanged for the same elements.)

36.2 (Section 36.1). Prove the de Morgan result

$$\overline{a \oplus b} = \bar{a} * \bar{b},$$

by showing that $(a \oplus b) \oplus (\bar{a} * \bar{b}) = 1$. Explain how the duality result (Problem 36.1) gives the other de Morgan theorem.

36.3 (Section 36.1). Let B be the Boolean algebra with the two elements 0 and 1. For arbitrary $a,b \in B$, prove the following:
(a) $a * (\bar{a} \oplus b) = a * b$;
(b) $(a \oplus b) * (a \oplus \bar{b}) = a$;
(c) $(a \oplus b) * \bar{a} * \bar{b} = 0$.

36.4 (Section 36.1). Using the laws of Boolean algebra for the set with two elements 0 and 1, show that:
(a) $a * b \oplus a * \bar{b} = a$;
(b) $a \oplus \bar{a} * \bar{b} * c = a \oplus \bar{b} * c$.
Use the result to obtain the truth tables in each case.

36.5 (Section 36.4). In Problem 36.4b, it is shown that

$$a \oplus \bar{a} * \bar{b} * c = a \oplus \bar{b} * c.$$

Design two sequences of gates which give the same output for the inputs a, b, and c. The resultant gates are said to be logically equivalent.

36.6 (Section 36.4). Design a circuit of gates to produce the output

$$(a \oplus \bar{b}) * (a \oplus \bar{c}).$$

Construct the truth table for this Boolean expression.

36.7 (Section 36.1). Show that the Boolean expressions $(a \oplus b) * (\bar{a} \oplus b) \oplus a$ and $a \oplus b$ are equivalent.

36.8 (Section 36.1). Show that the following Boolean expressions are equivalent:
(a) $a \oplus b$; (b) $a \oplus b * b$.

36.9 (Section 36.3). Find a Boolean expression f which corresponds to the truth table shown in Table 36.13.

Table 36.13

a	b	c	f
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0

36.10 (Section 36.3). Construct Boolean expressions for the output f in the devices shown in Figs 36.17a–d. Construct the truth tables in each case.

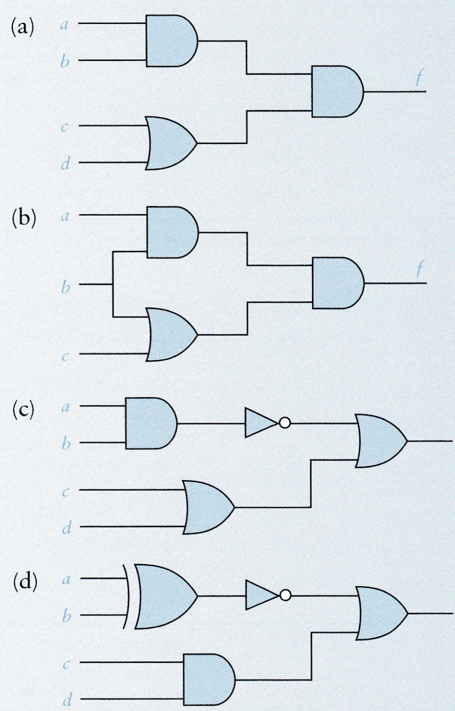


Fig. 36.17

36.11 Find the outputs f and g in the logic circuits shown in Fig. 36.18. This device can represent binary addition in which g is the 'carry' in the binary table shown in Table 36.14. The output g gives the '1' in the '10' in the binary sum $1 + 1 = 10$.

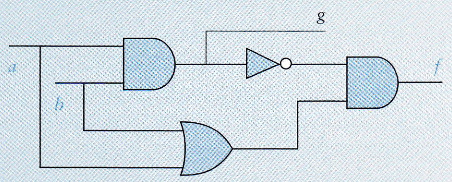


Fig. 36.18

Table 36.14

x	y	$x + y$
0	0	0
0	1	1
1	0	1
1	1	10

36.12 (Section 36.3). Reproduce the logic gate in Fig. 36.6 using just the NOR gate.

36.13 (Section 36.4). Using the disjunctive normal form, construct a Boolean expression f for the truth tables given in Tables 36.15 and 36.16.

36.14 (Section 36.3). Show that any Boolean expression can be modelled using just a NAND gate. (Hint: use a method similar to that explained in Example 36.4.)

Table 36.15

a	b	f
0	0	0
0	1	1
1	0	1
1	1	1

Table 36.16

a	b	c	f
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	0

36.15 (Section 36.4). Find switching functions for the switching circuits shown in Figs 36.19a,b.

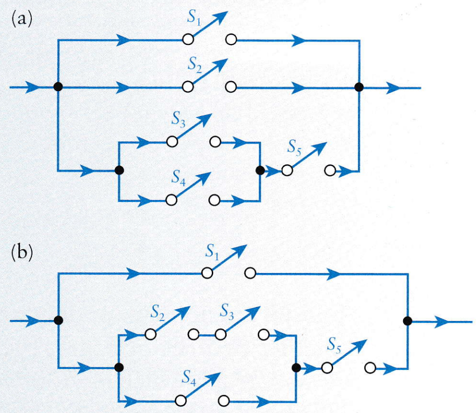


Fig. 36.19

36.16 A lecture theatre has three entrances and the lighting can be controlled from each entrance; that is, it can be switched on or off independently. The light is 'on' if the output f equals 1 and 'off' if $f = 0$. Let $a_i = 1$ $(i = 1, 2, 3)$ when switch i is up, and let $a_i = 0$ $(i = 1, 2, 3)$ when it is down. Construct a truth table for the state of the lighting for all states of the switches. Also specify a Boolean expression which will control the lighting.

37 Graph theory and its applications

CONTENTS

A graph is a network or diagram composed of points, or nodes or vertices, joined together by lines or edges, each of which has a vertex at each end. Figure 37.1 shows a graph which has four vertices {a, b, c, d} and six edges {ab, ab, ad, bd, bc, cd}. Two vertices are not joined in this graph, namely a and c, while a and b are joined by two edges. Generally, it is not the shape of the graph which is important; it is usually the number and connection of the edges which is significant. The terminology is unfortunate. Graphs in this context should not be confused with curves generated by functions as in Chapter 1. 'Networks' might be a more appropriate term but historical precedent is difficult to overturn. However the context usually fixes the meaning.

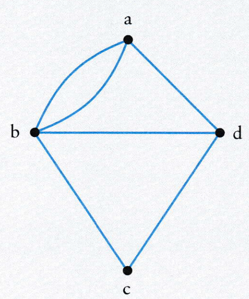


Fig. 37.1

37.1 Examples of graphs

Here are some practical examples of situations and objects which can be usefully represented by graphs.

(i) *Electrical circuits*. Figure 37.2a shows an electrical circuit with three resistors R_1, R_2, and R_3, an inductor L, and a voltage source V_1. Each edge has just one component, and the joins between components are the vertices (the term node is frequently used in circuit theory) in the graph. Care has to be taken with the definition of nodes (see Section 37.6): they are not necessarily where three or more wires meet. This circuit has four vertices a, b, c, d, and it can be represented by the graph in Fig. 37.2b if we are only interested in the links, not what they contain. The presence of a line or edge between two nodes in the graph indicates that there is a component between the nodes.

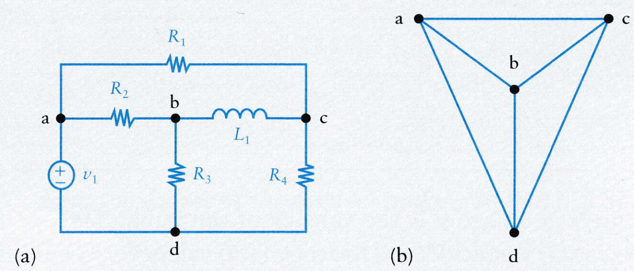


Fig. 37.2

Figure 37.3 shows another circuit with six vertices in which the boxes indicate electrical components. The wires joining c to f and b to e cross over each other. In the design of printed circuits, it is useful to know whether the circuit can be redrawn so that no wires cross. Such a graph, with no edges crossing, is known as a planar graph. The graph in Fig. 37.2 is planar, but the graph of the circuit in Fig. 37.3 has no planar drawing: at least two edges will cross in any plane diagram of it. We shall discuss this notion in Section 37.8.

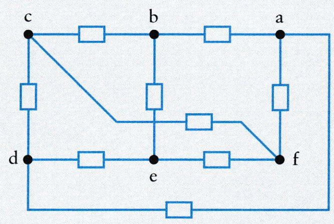


Fig. 37.3

(ii) *Chemical molecules*. Molecular diagrams look like candidates for graphs. The molecule of ethanol can be represented by Fig. 37.4a. In its graph representation in Fig. 37.4b, the vertices represent atoms and the edges bonds. The number of

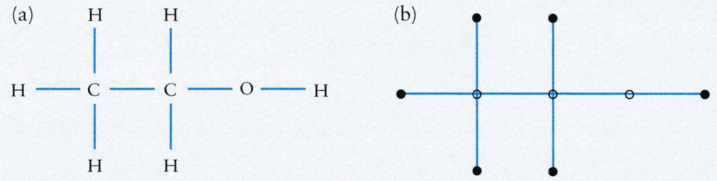


Fig. 37.4 Ethanol molecule.

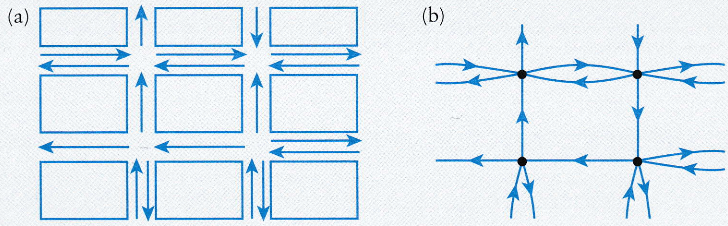


Fig. 37.5 (a) Traffic flow in a road grid, (b) Digraph representation of the roads in (a).

bonds which meet at an atom is the valency of the atom. Thus carbon (C) has valency 4, oxygen (O) valency 2, and hydrogen (H) valency 1. Generally in graphs, the number of edges that meet at a vertex is known as the degree of the vertex.

(iii) *Road maps*. Road maps and street plans are graphs with roads as edges and junctions as vertices. However, most road networks include one-way streets. Hence graphs need to be modified to indicate directions in which movement or flow is permitted. Figure 37.5a shows a typical section of a street plan with some one-way streets. We have to associate directions with the edges as shown in the graph of the plan in Fig. 37.5b. Note that two-way streets now have *two directed edges* associated with them. This is an example of a directed graph, which is also known by the shortened term digraph.

(iv) *Shortest paths*. Figure 37.6 shows a digraph with weights associated with each edge. The graph could represent routes between towns S and F which pass

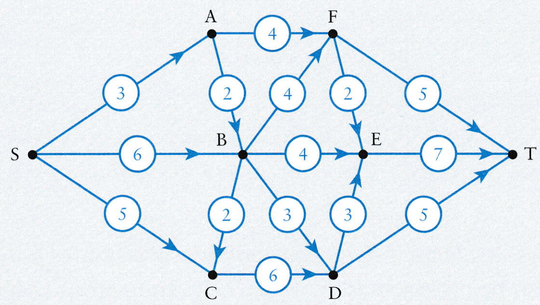


Fig. 37.6

through intermediate towns A, B, ... , the weights associated with each directed edge could stand for distances or times. This graph is shown as a digraph, but weights could be present without directions in some cases. We might be interested in this example in the shortest distance between the start (S) and the finish (F).

37.2 Definitions and properties of graphs

As we have seen, a graph is an object composed of vertices and edges with one vertex at each end of every edge. An edge which joins a vertex to itself is known as a loop. If two or more edges join the same two vertices then they are known as multiple edges. A graph with no loops or multiple edges is known as a simple graph. A graph with loops and/or multiple edges is known as a multigraph.

A graph in which every vertex can be reached from every other vertex along a succession of edges is said to be connected. Otherwise the graph is said to be disconnected. A connected graph is in one piece; a disconnected graph is in two or more pieces.

The degree of a vertex x is the number of edges that meet there, denoted by deg(x). If, in a graph G, all the vertices have the same degree r, then G is said to be regular of degree r.

Example 37.1 *Find the degree of the vertices in the graph in Fig.* 37.1.

Three edges meet at the vertex a. Hence deg(a) = 3. Four edges meet at b. Hence deg(b) = 4. Similarly, deg(c) = 2 and deg(d) = 3.

A simple graph in which every vertex is joined to every other vertex by just one edge is called a complete graph (see also Section 37.8).

Figure 37.7 shows some examples of the various graphs described above.

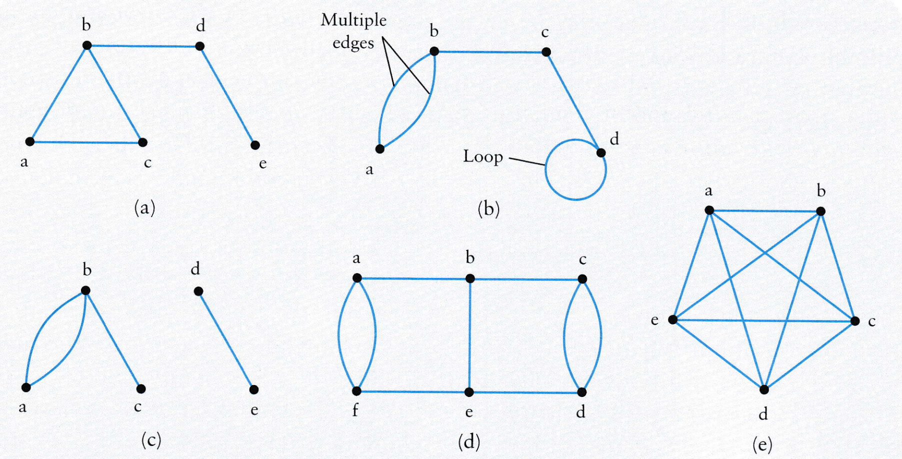


Fig. 37.7 (a) Connected simple graph. (b) Connected multigraph. (c) Disconnected multigraph. (d) Regular graph of degree 3. (e) Complete graph with five vertices: deg(a) = 4.

Since every edge has a vertex at each end, it follows that the sum of all the vertex degrees equals twice the number of edges. This is known as the handshaking lemma. For example, from Example 37.1,

$$\deg(a) + \deg(b) + \deg(c) + \deg(d) = 3 + 4 + 2 + 3 = 12,$$

which is twice the number of edges in the graph shown in Fig. 37.1.

There are two immediate consequences of the handshaking lemma:

(i) the sum of all the vertex degrees in a graph is an even number;

(ii) the number of vertices of odd degree is even.

Self-test 37.1

List the degrees of each vertex, as an increasing sequence, for each graph in Fig. 37.7.

37.3 How many simple graphs are there?

Graphs can be described as labelled, in which case the vertices are distinguishable as in Fig. 37.8a or unlabelled as in Fig. 37.8b. If we look at graphs with just three vertices, there are eight labelled simple graphs as shown in Fig. 37.9, but there are just four distinct unlabelled graphs as shown in Fig. 37.10. In Fig. 37.9, the three labelled graphs with one edge will correspond to the one unlabelled graph in Fig. 37.10.

The number of labelled simple graphs with n vertices is fairly easy to calculate. Between any two vertices, there is the possibility of an edge. Any vertex can be joined to $n-1$ other vertices. Since this will duplicate edges, there will be $\frac{1}{2}n(n-1)$ possible edges. Each edge may be either present or not. Hence the number of possible combinations of present and absent edges will be $2^{\frac{1}{2}n(n-1)}$, which is the number of labelled graphs (this number increases extremely rapidly, with n). Thus there must be $2^{\frac{1}{2}4(4-1)} = 2^6 = 64$ labelled graphs with four vertices; of these,

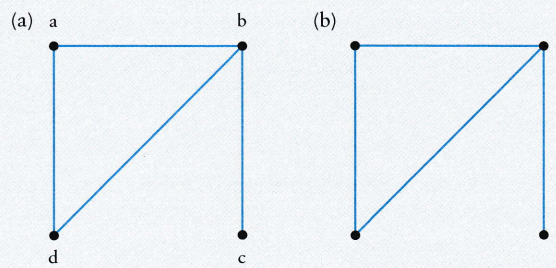


Fig. 37.8

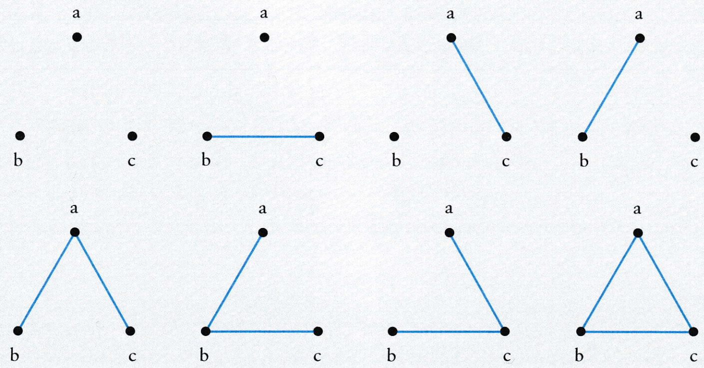


Fig. 37.9 Labelled graphs with three vertices.

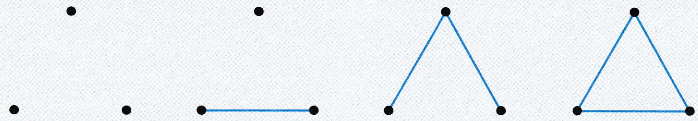


Fig. 37.10 Unlabelled graphs with three vertices.

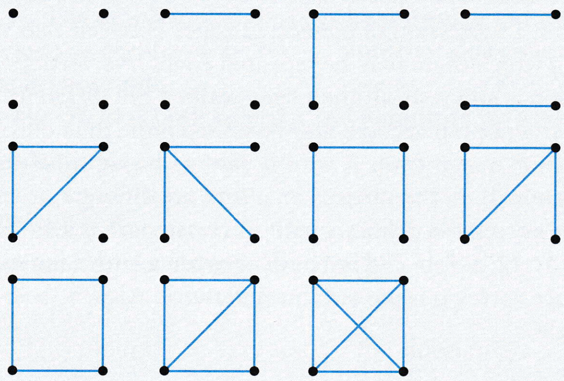


Fig. 37.11 All unlabelled graphs with four vertices.

11 can be identified as unlabelled graphs. The latter graphs are shown in Fig. 37.11. Of the 11 unlabelled graphs it can be seen that six are connected and four are regular.

For applications involving electrical circuits, the main interest is in connected graphs. The numbers of the various categories of graphs up to $n = 7$ vertices are

Table 37.1

n	1	2	3	4	5	6	7
Labelled graphs	1	2	8	64	1024	32 768	2 097 152
Unlabelled graphs	1	2	4	11	34	156	1 044
Connected graphs	1	1	2	6	21	112	853
Regular graphs	1	2	2	4	3	8	6

given in Table 37.1. It can be seen from the table that the number of unlabelled graphs is a considerable reduction on the labelled set, and that regular graphs are comparatively rare. The counting of unlabelled graphs does not follow from a simple formula.

Self-test 37.2

List all unlabelled simple graphs with five vertices. Indicate which graphs are connected, and which are regular. What are the degrees of the regular graphs?

37.4 Paths and cycles

Suppose we follow a succession of connected edges between two vertices a and z in a graph, along which there may be repeated edges and vertices. This is known as a walk between a and z. If all the edges walked are different (i.e. no edge is covered more than once but vertices may be visited more than once), then the walk defines what is known as a trail. A trail is said to be closed if the first and last vertices are the same. If all the vertices on a trail are different, except possibly the end pair, then the succession defines a path. A closed path is known as a cycle. For example, in Fig. 37.12, a–f–b–c–d is a path between a and d, but a–b–f–e–b–c–d is only a trail since vertex b is passed through twice. Also, a–b–c–d–e–f–a is an example of a cycle.

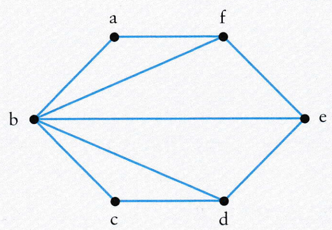


Fig. 37.12

Example 37.2 *Electrical circuits are usually such that every edge of their representative graph is part of a cycle. List all the distinct cycles in the graph in Fig. 37.2a.*

The graph of the circuit is repeated in Fig. 37.13. The complete list of cycles is:

3-edge cycles: a–b–c–a, a–b–d–a, a–d–c–a, b–d–c–b;
4-edge cycles: a–b–d–c–a, a–d–b–c–a, a–b–c–d–a.

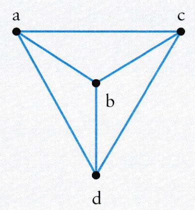


Fig. 37.13

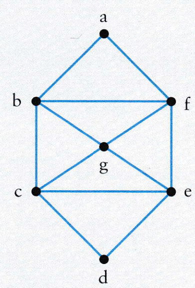


Fig. 37.14

Some graphs have special closed-path and cycle properties. A connected graph G is said to be **eulerian** if there exists a closed trail that includes every edge in G. A connected graph G is said to be **hamiltonian** if there exists a cycle that includes every vertex in G. The graph in Fig. 37.13 is hamiltonian but not eulerian. One hamiltonian cycle in its graph is a–b–d–c–a. Note that this cycle does not have to cover *every* edge in the graph.

The graph in Fig. 37.14 is both eulerian and hamiltonian. An eulerian trail is

a–b–c–d–e–f–g–e–c–g–b–f–a,

and a hamiltonian cycle is

a–b–c–d–e–g–f–a.

It can be shown that a connected graph is eulerian if and only if every vertex has even degree. This provides an easy test for the eulerian property of a graph.

37.5 Trees

A connected graph which has no cycles is known as a **tree**. An example of a tree is shown in Fig. 37.15. The edges in a tree are called **branches**.

Suppose that a graph G consists of the set $V(G)$ of vertices and the set $E(G)$ of edges. Then any graph whose vertices and edges are subsets of $V(G)$ and $E(G)$ respectively is called a **subgraph**. It is important to note that the subgraph must be a graph whose vertices and edges come from G; and only edges that join two vertices of the subgraph are permitted in the subset of $E(G)$.

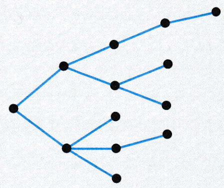


Fig. 37.15 An example of a tree.

Suppose that G is a connected graph.

A spanning tree of G is a subgraph of G which is a tree and includes all vertices of G.

Figure 37.16a shows a connected graph G and Fig. 37.16b shows a spanning tree of G. Graphs can have many different spanning trees. The set of edges that are not part of the spanning tree (the broken edges in Fig. 37.16b) is known as the cotree and its edges are called links.

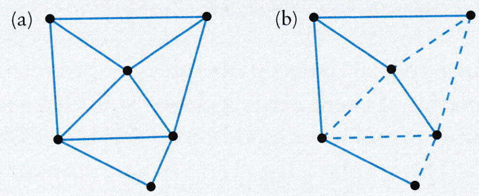


Fig. 37.16 (a) Connected graph. (b) The same graph with a spanning tree.

Construct a tree from a vertex by adding edges. Each edge added must introduce a new vertex, since otherwise a cycle would be created and the graph would no longer be a tree. A tree with two vertices has one edge, a tree with three vertices has two edges and so on. Hence a tree with n vertices must have just $n - 1$ branches. It follows that a graph with n vertices must have a cotree with $e - n + 1$ links, where e is the number of edges of the graph.

We now introduce the cutset, by which we can disconnect a graph into two subgraphs which together contain all the original vertices, by removing a minimum set of edges in the graph.

Cutset

In a connected graph, a cutset is a set of edges (a) whose removal disconnects the graph into two subgraphs and (b) no proper subset of the cutset disconnects the graph.

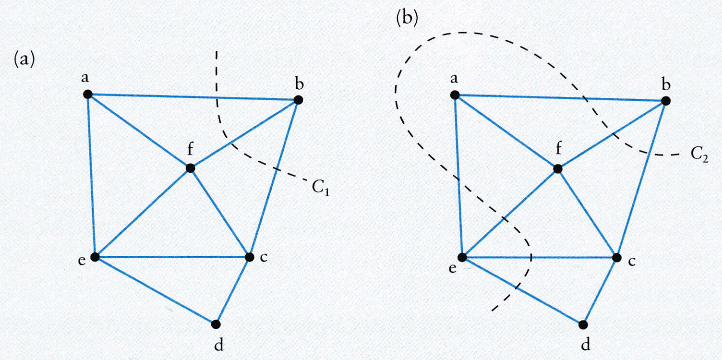


Fig. 37.17 A cutset of a graph.

A proper subset of the cutset is one which does not include the cutset. There must be no redundancy in the cutset. Thus, for example in Fig. 37.17a, the broken line C_1, which removes the edges ba, bf, and bc, defines a cutset {ba, bf, bc}, since $\{b\}$ and $\{a, c, d, e, f\}$ are disconnected subgraphs. C_2 in Fig. 37.17b does not define a cutset, since the *subset* {ab, bf, bc} of edges disconnects the graph.

Self-test 37.3

(a) What are the degrees of the vertices in the spanning tree in Fig. 37.16(b)? Design a spanning tree with vertex degrees $\{1, 1, 2, 2, 2\}$. (b) Indicate spanning trees of the graph in Fig. 37.14 with vertex degrees (i) $\{1, 1, 2, 2, 2, 2, 2\}$, (ii) $\{1, 1, 1, 1, 2, 2, 4\}$.

37.6 Electrical circuits: the cutset method

In this section we give a brief description of the representation of circuits by graphs, and show how Kirchhoff's laws can be applied to cutsets of the resulting graphs. Figure 37.18a shows a plan of a circuit with seven resistors, a voltage supply, and two capacitors. This particular circuit has 10 components and 10 edges.

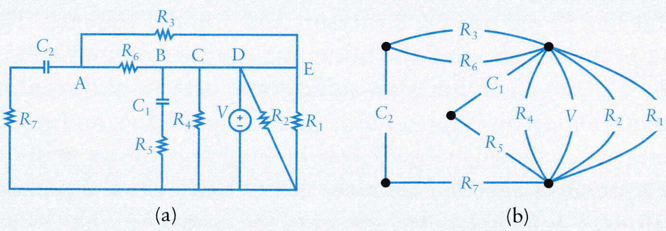


Fig. 37.18 A circuit and its graph.

Note that A will be a vertex or node (a preferred term in circuits) but that the joins B, C, and D are not separate nodes but can be coalesced into a single node. The equivalent graph is shown in Fig. 37.18b: it has five nodes and 10 edges. Note that it is a multigraph with two nodes joined by two edges and two nodes joined by four edges.

A circuit loop in the circuit is a cycle in the graph.

Kirchhoff's laws have already been stated in eqn (21.8), but for convenience they are given again here in graph terms. They state (i) that *the algebraic sum of the voltages around any loop is zero*, and (ii) that *the algebraic sum of the currents entering any node is zero*.

In addition, for resistors we also have Ohm's law which states that *the voltage across a resistor is directly proportional to the current flowing through it*, that is

$$v \propto i \qquad \text{or} \qquad v = Ri,$$

where the constant R is measured in units called *ohms* (Ω). Figure 37.19 shows a circuit with two independent maintained current sources i_X and i_Y: the symbol of the circle enclosing an arrow represents a maintained current in the direction of the arrow.

The corresponding six-node digraph with currents $i_1, i_2, \ldots, i_8$ in the directions indicated is shown in Fig. 37.20. If any current turns out to be negative then its direction will be opposite to that shown.

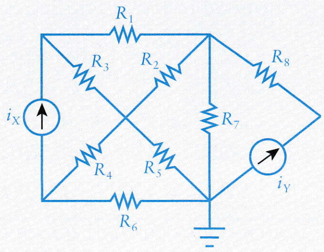


Fig. 37.19

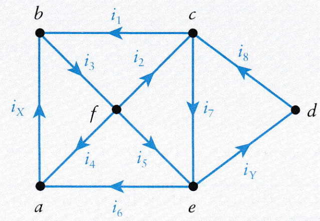


Fig. 37.20

Now introduce nodal voltages $v_a, v_b, \ldots, v_f$ as shown in Fig. 37.21. The use of nodal voltages means that effectively Kirchhoff's first law is automatically satisfied. The earthing at e makes $v_e = 0$ and other voltages can be measured relative to this zero ground potential.

This circuit has 13 unknowns: 8 currents and 5 nodal voltages. The problem with circuits is the selection of the minimum number of consistent equations from Kirchhoff's laws and Ohm's law sufficient to determine the unknowns.

The graph of this circuit is the same as that in Fig. 37.16a, and we shall use the same spanning tree as shown in Fig. 37.16b. In this graph, the number of nodes n is 6, the number of edges e is 10. Hence the cotree has, from the previous section, $e - n + 1 = 10 - 6 + 1 = 5$ links. Any cutset of the original graphs which contains one and only one branch of the spanning tree (the rest of the cutset consisting of links) is known as a fundamental cutset of the circuit. Hence we can associate five

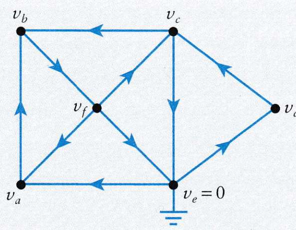


Fig. 37.21

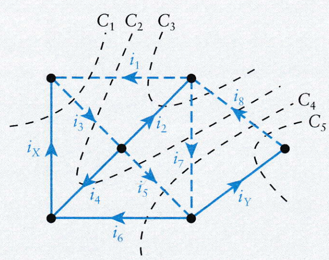


Fig. 37.22

fundamental cutsets with the spanning tree in Fig. 37.16b. Five possible cutsets $C_1, C_2, \ldots, C_5$ are shown in Fig. 37.22.

By repeated use of Kirchhoff's second law to the nodes on one side of a cutset, it follows that the algebraic sum of the currents crossing the cutset must be zero. Hence the five cutset equations are:

$$C_1\colon i_1 - i_3 + i_X = 0, \tag{37.1}$$

$$C_2\colon i_1 - i_3 + i_4 + i_5 + i_7 - i_8 = 0, \tag{37.2}$$

$$C_3\colon i_i - i_2 + i_7 - i_8 = 0, \tag{37.3}$$

$$C_4\colon i_6 - i_5 - i_7 + i_8 = 0, \tag{37.4}$$

$$C_5\colon i_Y - i_8 = 0. \tag{37.5}$$

These equations must be independent since each one contains a current from a branch of the spanning tree which does not appear in any other equation. Further any non-fundamental cutset equation will be a linear combination of the five fundamental cutset equations. The number of branches in the spanning tree defines the number of independent equations.

We can also apply Ohm's law to each resistor (note that current flows from high to low potential). Thus the voltage difference across R_1 is $v_c - v_b$, so that

$$i_1 = (v_c - v_b)/R_1. \tag{37.6}$$

Similarly

$$i_2 = (v_f - v_c)/R_2, \tag{37.7}$$

$$i_3 = (v_b - v_f)/R_3, \tag{37.8}$$

$$i_4 = (v_f - v_a)/R_4, \tag{37.9}$$

$$i_5 = v_f/R_5, \tag{37.10}$$

$$i_6 = (-v_a)/R_6, \tag{37.11}$$

$$i_7 = v_c/R_7, \tag{37.12}$$

$$i_8 = (v_d - v_c)/R_8. \tag{37.13}$$

We can now substitute for the currents from (37.6) to (37.13) into (37.1) to (37.5) resulting in five linear equations to determine the nodal voltages v_a, v_b, v_c, v_d, v_f in terms of the known currents i_X and i_Y. The remaining currents can then be calculated from (37.6) to (37.13).

Example 37.3 *Using the cutset method, find all currents and nodal voltages in the circuit shown in Fig. 37.23.*

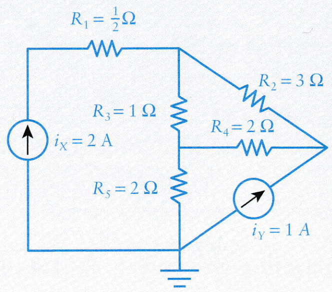


Fig. 37.23

The circuit can be represented by a graph with five nodes (Fig. 37.24) with the currents i_1, i_2, i_3, i_4, i_5 in the directions shown.

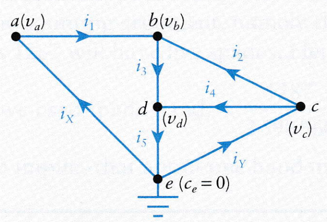


Fig. 37.24

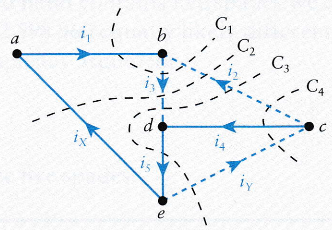


Fig. 37.25 Fundamental cutsets.

A spanning tree with three links is shown in Fig. 37.25 together with cutsets C_1, C_2, C_3, C_4. Hence Kirchhoff's second law implies:

$$C_1\colon i_1 - i_3 + i_2 = 0, \tag{37.14}$$
$$C_2\colon i_X - i_3 + i_2 = 0, \tag{37.15}$$
$$C_3\colon -i_Y + i_5 - i_3 + i_2 = 0, \tag{37.16}$$
$$C_4\colon -i_Y + i_4 + i_2 = 0. \tag{37.17}$$

With $v_e = 0$, the currents in terms of the nodal voltages v_a, v_b, v_c, v_d are, by Ohm's law:

$$i_1 = (v_a - v_b)/R_1 = 2(v_a - v_b), \tag{37.18}$$
$$i_2 = (v_c - v_b)/R_2 = \tfrac{1}{3}(v_c - v_b), \tag{37.19}$$
$$i_3 = (v_b - v_d)/R_3 = v_b - v_d, \tag{37.20}$$
$$i_4 = (v_c - v_d)/R_4 = \tfrac{1}{2}(v_c - v_d), \tag{37.21}$$
$$i_5 = v_d/R_5 = \tfrac{1}{2}v_d. \tag{37.22}$$

Eliminate the currents in (37.14) to (37.17) using (37.18) to (37.22):

$$2v_a - \tfrac{10}{3}v_b + \tfrac{1}{3}v_c + v_d = 0, \tag{37.23}$$
$$\tfrac{4}{3}v_b - \tfrac{1}{3}v_c - v_d = 2, \tag{37.24}$$
$$-\tfrac{4}{3}v_b + \tfrac{1}{3}v_c - \tfrac{3}{2}v_d = 2, \tag{37.25}$$
$$-\tfrac{1}{3}v_b + \tfrac{5}{6}v_c - \tfrac{1}{2}v_d = 1. \tag{37.26}$$

↗

Example 37.3 continued

These are linear equations which can be solved using the methods of Chapter 12. Computer algebra is also very useful in solving sets of equations of this type (see the computer algebra applications for Chapter 12 in Chapter 42). The answers are

$$v_a = 5\text{ V}, \qquad v_b = 4\text{ V}, \qquad v_c = 4\text{ V}, \qquad v_d = 2\text{ V}.$$

Since $v_c = v_b$, no current flows through the resistor on bc.

We can summarize the result for an earthed circuit which contains only resistors and current sources. Suppose that the representative graph of the circuit contains n nodes and e edges of which f contain known current sources. The curcuit will have $e - f$ unknown currents and $n - 1$ unknown nodal voltages giving $e - f + n - 1$ unknowns in total. Its spanning tree will have $n - 1$ edges which will lead to $n - 1$ fundamental cutset equations, and Ohm's law will apply to $e - f$ resistors. Hence we shall always have a consistent set of $e - f + n - 1$ equations to find the unknowns.

This result can be extended to circuits with current sources, voltage sources (batteries), and resistors. If the representative graph has n nodes and e edges of which f contain current sources and s maintained voltage sources, then the number of unknown currents will be $e - f$ and the number of unknown nodal voltages will be $n - 1 - s$ since the nodal voltage difference across a battery will be known. Hence the number of unknowns is $e - f + n - 1 - s$ which will satisfy $n - 1$ cutset equations and $e - f - s$ Ohm's laws.

37.7 Signal-flow graphs

Figure 37.26 shows a block diagram of a negative-feedback control system. The input into the system is $P(s)$ and the output $Q(s)$. All operations are defined by their transfer functions (see Section 25.4). The boxes represent devices or controllers. The circle represents a sum operator, and the return sign on $F(s)$ indicates positive or negative feedback. The output signal $Q(s)$ is fed back into the input through $H(s)$, and it is a negative feedback which will reduce the output. In a later problem, we shall consider a device with a positive feedback. Thus the input into $G(s)$ is

$$A(s) = P(s) - F(s). \tag{37.27}$$

The boxes each produce outputs given by the transfer functions

$$Q(s) = G(s)A(s), \tag{37.28}$$

$$F(s) = H(s)Q(s). \tag{37.29}$$

We wish to find $Q(s)$ in terms of $P(s)$, $G(s)$, and $H(s)$, from the equations (37.27) to (37.29). Thus, from (37.28)

$$Q(s) = G(s)A(s) = G(s)[P(s) - F(s)], = G(s)[P(s) - H(s)Q(s)].$$

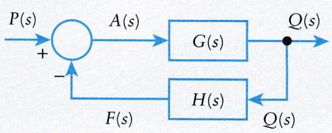


Fig. 37.26 Negative-feedback control system.

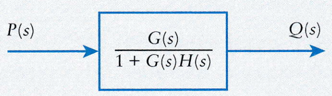


Fig. 37.27 Block-reduced diagram for Fig. 37.26.

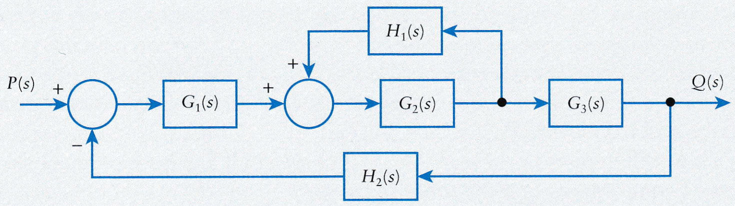


Fig. 37.28 A multiple-feedback control system.

Hence the output transfer function is

$$Q(s) = \frac{G(s)}{1 + G(s)H(s)} P(s).$$

This is the closed-loop transfer function. The actual signal can be obtained by finding the inverse Laplace transform for $Q(s)$. Hence the system is equivalent to that shown in Fig. 37.27.

If the feedback reinforces the input signal it is called positive feedback. Figure 37.28 shows a multiple-feedback control system with a positive and a negative feedback. The output signal is given by

$$Q(s) = \frac{G_1(s)G_2(s)G_3(s)}{1 - G_2(s)H_1(s) + G_1(s)G_2(s)G_3(s)H_2(s)} P(s), \tag{37.30}$$

which can be obtained by the method of block-diagram reduction. For example, the feedback through H_1 makes the system equivalent to that shown in Fig. 37.29. We can now combine the series devices which reduce the system to the negative-feedback control system considered at the beginning of this section. The details are omitted here.

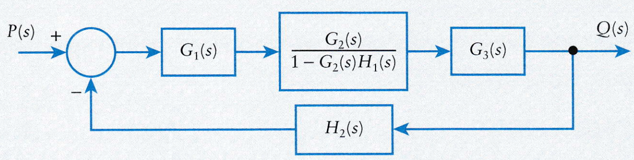


Fig. 37.29 First stage in the block reduction of the multiple-feedback control system.

This block-reduction method can get quite complicated for a complex feedback system. Instead of using block reduction in this way, represent the system by a weighted digraph as shown in Fig. 37.30, where the weights are the transfer functions – except that the edges representing the input and output are assigned

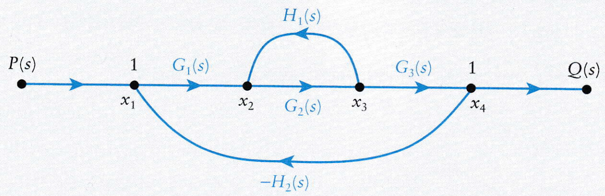


Fig. 37.30 Signal-flow graph for the multiple-feedback control system shown in Fig. 37.28.

weight 1 since they carry no devices. Also the negative feedback is replaced by $-H_2(s)$, to make sure that it reduces the input into $G_1(s)$. This is the signal-flow graph of the system. Let the inputs into the nodes be x_1, x_2, x_3, and x_4 as shown; then, for the positive-feedback cycle,

$$x_3 = G_2x_2, \qquad x_2 = G_1x_1 + H_1x_3.$$

(The argument (s) has now been dropped from the working.) Hence

$$x_3 = \frac{G_1G_2x_1}{1 - G_2H_1}.$$

In other words, we can replace (a) by (b) in Fig. 37.31.

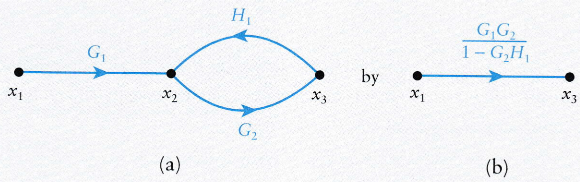


Fig. 37.31

There are other rules, and a complete list now follows for the replacements for subgraphs in the graph.

(a) *Multiple edges*. See Fig. 37.32. This follows since

$$x_2 = Gx_1 + Hx_1 = (G + H)x_1.$$

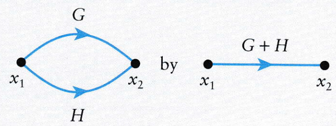


Fig. 37.32 Multiple edges.

(b) *Edges in series*. See Fig. 37.33. This follows since

$$x_3 = Hx_2 = H(Gx_1) = HGx_1.$$

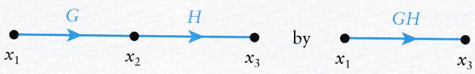


Fig. 37.33 Edges in series.

(c) *Cycles*. See Fig. 37.34. This follows since

$$x_3 = Hx_2 \quad \text{and} \quad x_2 = Gx_1 + Jx_3.$$

Assume that $HJ \neq 1$; otherwise there is infinite gain.

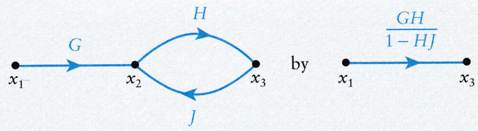


Fig. 37.34 Cycle.

(d) *Loops*. See Fig. 37.35. This follows since

$$x_2 = Gx_1 + Hx_2$$

with $H \neq 1$.

(e) *Stems*. See Fig. 37.36. This follows since

$$x_2 = Gx_1, \qquad x_3 = Hx_2 = HGx_1, \qquad x_4 = Jx_2 = JGx_1.$$

Apply these rules to the successive reduction of the feedback system in Fig. 37.30. The sequence of steps in the reduction of the signal-control graph to a single-edge graph is shown in Fig. 37.37. The weight of the final edge agrees with the output in eqn (37.38).

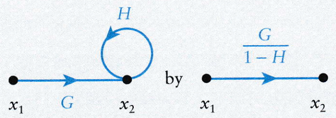


Fig. 37.35 Loop.

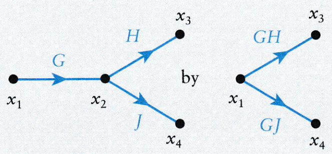


Fig. 37.36 Stem.

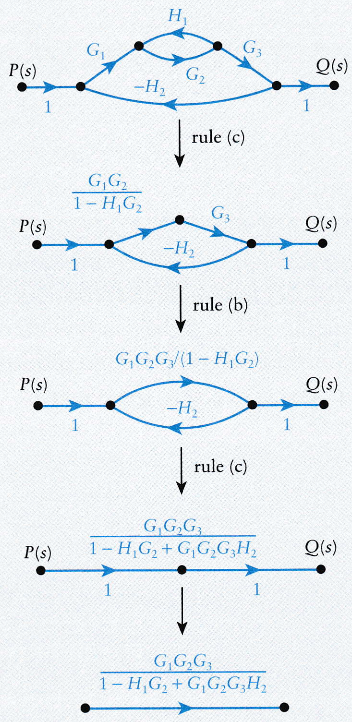


Fig. 37.37 Successive steps in the reduction of the signal-flow graph of the control system shown in Fig. 37.28.

Essentially the operations in a signal-flow graph are those applied to a *weighted digraph* as illustrated in the following example.

Example 37.4 *Find the output–input relation in the signal-flow graph shown in Fig. 37.38.*

Applying rule (a) to the multiple edge, and rule (c) to the cycle, the graph is reduced to Fig. 37.39. Apply the series rule to the divided edges to give Fig. 37.40. Finally the multiple-edge and series rules give Fig. 37.41. Thus the output is given by

$$q = \frac{abd}{1 - bc} + he(g + f).$$

In the actual control system $a, b, c, \ldots$ will be transfer functions.

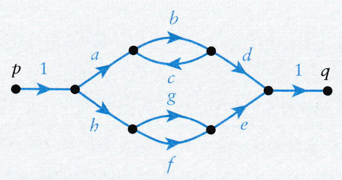


Fig. 37.38

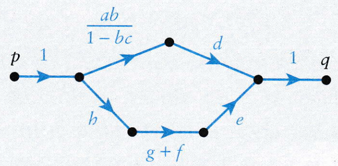


Fig. 37.39

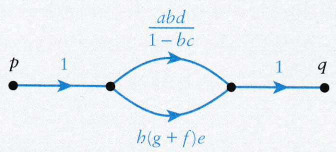


Fig. 37.40

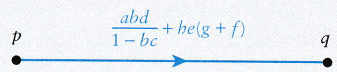


Fig. 37.41

Self-test 37.4

Suppose that c is in the opposite direction in the signal-flow graph Fig. 37.38. Find the new output–input relation.

37.8 Planar graphs

As we remarked in Section 37.1, planar graphs are important in circuit design since planar circuits can be manufactured as a single board. A planar graph is a graph that can be drawn with no edges crossing or meeting except at vertices. The standard example of a simple application which cannot be represented by a planar graph is the delivery of three services, water (W), gas (G), and electricity (E), to three houses A, B, C (Fig. 37.42). This graph has no plane drawing. The reorganization of the graph in Fig. 37.43 shows the impossibility of this; if W and C are connected last then this edge must cross either AE or BG.

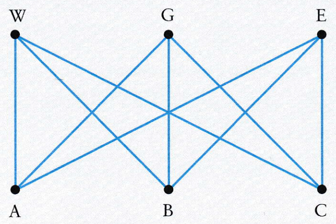


Fig. 37.42 Bipartite graph $K_{3,3}$.

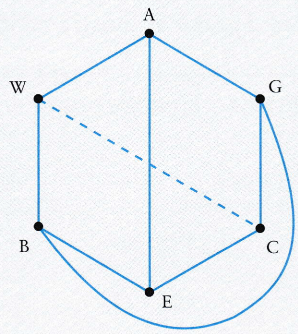


Fig. 37.43

The graph in Fig. 37.42 is an example of a bipartite graph in which one set of vertices may be connected to another set of vertices, but not to vertices in the same set. If every vertex in one set is connected by one edge to every vertex in the other set then it is called a complete bipartite graph. If the sets have m and n vertices respectively, then the notation $K_{m,n}$ denotes the complete bipartite graph. Figure 37.42 shows the graph $K_{3,3}$ and this graph is not planar. Check that the graphs $K_{2,2}$ and $K_{2,3}$ are planar.

In planar graphs there is a relation between the numbers of vertices, edges, and faces. In a plane drawing of a graph, the plane is divided into regions called faces. One face is the region external to the graph. Figure 37.44 shows a planar graph with five vertices and seven edges, and with four faces: A, B, C, and the external face D.

A remarkable formula, due to Euler, links the numbers of vertices, edges, and faces of a graph.

Theorem (Euler). Suppose that the graph G has a planar drawing, and let v be the number of vertices, e the number of edges, and f the number of faces of G. Then

$$v - e + f = 2.$$

Proof. For the graph G, define a spanning tree (see, for example, Fig. 37.45). The spanning tree must have n vertices and $n-1$ edges (see Section 37.5). It must also have just one face. Since

$$n - (n-1) + 1 = 2,$$

Euler's formula holds for the spanning tree. Successively replace the other edges in the graph. Each time an extra edge is added, a face is divided and one extra face is added. However, algebraically, this cancels the additional edge in the accumulation to Euler's formula for the spanning tree. Hence

$$v - e + f = 2$$

for the reconstructed graph G.

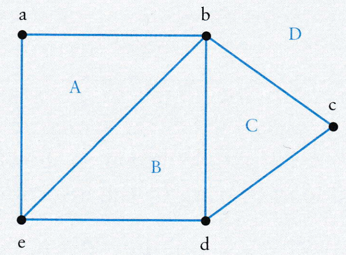


Fig. 37.44 A planar graph with fiver vertices, seven edges, and four faces.

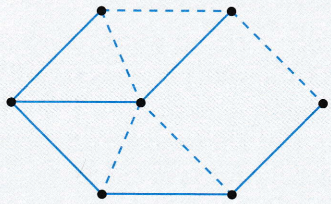


Fig. 37.45 A graph with a spanning tree.

The complete graph with n vertices is denoted by K_n. Since every vertex is joined to $n - 1$ vertices, K_n has $\frac{1}{2}n(n - 1)$ edges. The graphs of K_2, K_3, K_4, and K_5 are shown in Fig. 37.46. Of these graphs, K_2, K_3, and K_4 are planar, but K_5 and all succeeding complete graphs are not.

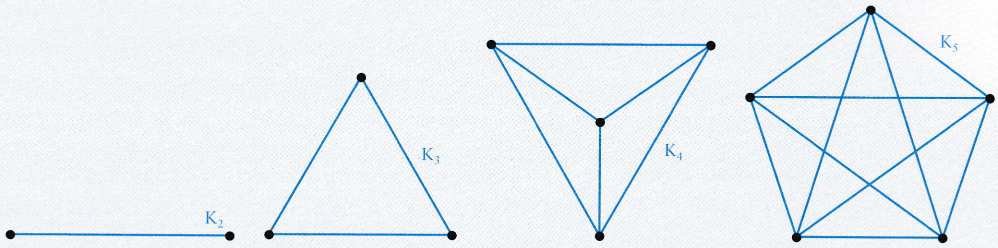


Fig. 37.46 The complete graphs K_n for $n = 2, 3, 4, 5$.

The graphs $K_{3,3}$ and K_5 are the keys to tests for planarity of graphs, and whether it is possible to design, for example, a plane printedcircuit board to make the required connections between electronic components. It was proved by Kuratowski in 1930 that every non-planar graph contains subgraphs which are either $K_{3,3}$ or K_5, or $K_{3,3}$ or K_5 with additional vertices on their edges.

Further discussion of graph theory with many applications can be found in the introductory text by Wilson and Watkins (1990).

Self-test 37.5

(a) A regular dodecahedron has 12 faces (pentagons) and 30 edges. How many vertices does it have?
(b) An icosahedron has 20 faces (triangles) and 12 vertices. How many edges does it have?

37.9 Further applications

Braced frameworks

Consider a frame which consists of four struts in the shape of a rectangle (Fig. 37.47a) with pin joints at each corner. Without a diagonal tie the structure will not support a vertical load, but will collapse into a parallelogram as shown in Fig. 37.47b. The structure can be made rigid and load bearing by the insertion of a diagonal strut as in Fig. 37.48.

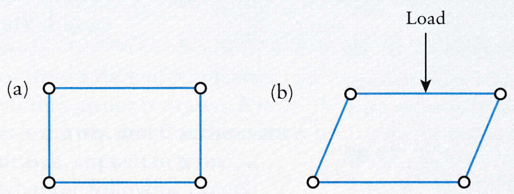


Fig. 37.47 Single unbraced pin-jointed frame.

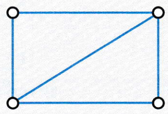


Fig. 37.48 Braced frame.

Consider now a pinjointed framework with $m \times n$ rectangular frames with some individual frames braced. How can we decide whether a particular framework is braced, that is no part of it can be sheared? And if it is braced, how many ties could be removed to leave a **minimum bracing**? The framework is similar to a vertical section of scaffolding or a steel-framed building, although in both cases the joins are bolted but can still need bracing to ensure rigidity.

Figure 37.49 shows a 5×6 framework with 11 braces as shown (braces can be diagonal struts in either direction). Label the *cell* rows $r_1, r_2, \ldots, r_5$ and the cell columns $c_1, c_2, \ldots, c_6$ as shown in Fig. 37.49. The framework will be represented by a *bipartite graph* (see Section 37.8) with the cell rows and columns as vertices. Arrange them in rows as shown in Fig. 37.50.

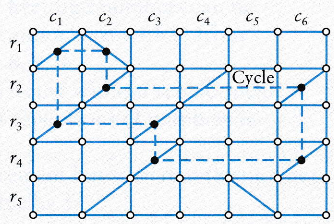


Fig. 37.49 5×6 framework.

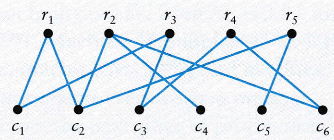


Fig. 37.50

If a particular rectangular cell is braced then the identifying row and column vertices are joined by an edge. Thus the cell r_1c_1 is braced so that an edge joins r_1 and c_1 in the bipartite graph. No edge joins r_1 and c_3 since this cell is not braced. The bipartite graph representing the framework is shown in Fig. 37.50. If the graph is *connected*, then the framework is braced since the shearing of any cell or group of cells is not then possible. The graph is connected in this case, and the framework is braced. Can any braces be removed in such a way that the framework is still braced? Any brace which is removed must not *disconnect* the graph. If the graph contains a *cycle* (Section 37.4) then any edge removed from the cycle will not disconnect the graph. This removal rule can be applied to each cycle in the graph. If, at the end of this process, there are no cycles remaining and the graph remains connected, then the framework is said to have a **minimum bracing**. The framework graph in Fig. 37.50 contains just one cycle, namely $r_1c_1r_3c_3r_4c_6r_2c_2r_1$ (see Fig. 37.49). *Any* edge can be removed from this cycle leaving a minimum bracing. The removal of any further edges will disconnect the graph.

If every cell is braced in a framework then the bipartite graph will be complete, and the framework will be seriously overbraced. You might note that a complete bipartite graph $K_{m,n}$ has mn edges but a minimum bracing for an $m \times n$ framework has $m + n - 1$ edges: for example, if $m = 5$ and $n = 6$ then $mn = 30$ whilst $m + n - 1 = 10$.

Figure 37.51 shows an unbraced 4×5 framework, its (disconnected) graph, and the same framework sheared.

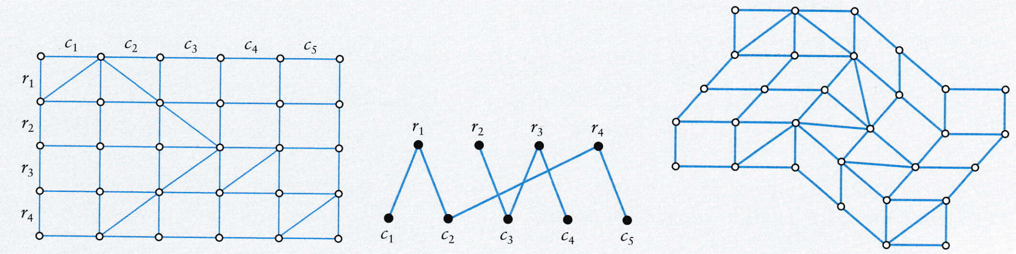


Fig. 37.51 An unbraced framework.

Phasing of traffic signals

Figure 37.52 shows a road junction with eight incoming lanes of traffic and a one-way exit. Suppose that each lane can be controlled by its own individual signal.

One solution for traffic management would be to allow each lane to have a green signal in sequence with the remaining all on red, but this would be inefficient since obviously several lanes of traffic can move simultaneously without risk. How can an efficient phasing of the signals be designed?

Label each incoming lane $a, b, c, \ldots, h$ as shown, and let these be vertices of a graph (Fig. 37.53). Starting, say, with a we decide which traffic lanes are *compatible* with a; that is, which lanes can also have green lights simultaneously without risk of a collision. Thus a and b are compatible, and we therefore join a and b by

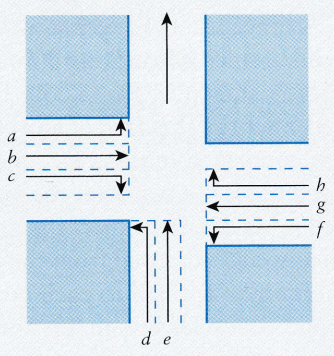


Fig. 37.52 Road junction.

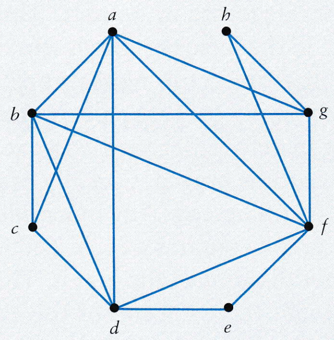


Fig. 37.53

an edge. Lanes *a* and *c* are also compatible, and we therefore join *a* and *b* by an edge. Lanes *a* and *c* are also compatible, but *a* and *e* are not, and so on. The graph *G* in Fig. 37.53 shows which lanes are compatible, and is known as the compatibility graph for this junction.

We now look for *complete subgraphs* (Section 37.8) in *G*. An edge is a complete subgraph (K_2), a triangle (K_3) is a complete subgraph with three vertices, K_4 with four vertices, and so on. We try to use the largest subgraphs in any *covering* of *G*, that is a list of subgraphs which includes *all* vertices. In *G*, *abcd*, *abdf*, and *abfg* are K_4 subgraphs, and there are a large number of triangles. For example, we can cover *G* by the set of subgraphs

$$\{abcd, abfg, def, fgh\}.$$

Generally, we include as many large subgraphs as possible. In this list it is better to use *fgh* rather than just *gh*: this could be chosen since *f* is included in other subgraphs.

Suppose that the period of the traffic signal sequence is *T* seconds with each lane having a green light for at least $\frac{1}{5}T$. There are four different traffic flows represented by the subgraphs. Suppose that each subgraph list of lanes has a green light for $\frac{1}{4}T$. The green/red phasing sequence is shown in Table 37.2.

The actual phasing lane by lane is shown in Fig. 37.54 where the solid line indicates the green light for a lane. For example, between $\frac{1}{4}T$ and $\frac{1}{2}T$, lanes *a*, *b*, *e*, *f* are on green with the others on red.

Table 37.2

	Subgraph			
Time	*abcd*	*abfg*	*def*	*fgh*
$0-\frac{1}{4}T$	green	red	red	red
$\frac{1}{4}T-\frac{1}{2}T$	red	green	red	red
$\frac{1}{2}T-\frac{3}{4}T$	red	red	green	red
$\frac{3}{4}T-T$	red	red	red	green

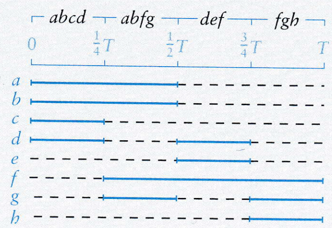


Fig. 37.54 Traffic phasing.

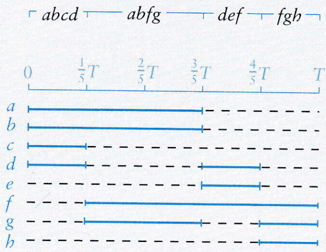


Fig. 37.55

The total waiting time for the traffic at the junction is a measure of the efficiency of the timings and phases. Let $t_a, t_b, t_c, \ldots$ be the waiting times of the lanes so that, from Fig. 37.54, we can see that $t_a = \frac{1}{2}T$, $t_b = \frac{1}{2}T$, $t_c = \frac{3}{4}T$, etc. Hence the total waiting time W_T is given by

$$W_T = t_a + t_b + \cdots + t_h = \tfrac{1}{2}T + \tfrac{1}{2}T + \tfrac{3}{4}T + \tfrac{1}{2}T + \tfrac{3}{4}T + \tfrac{1}{4}T + \tfrac{1}{2}T + \tfrac{3}{4}T = \tfrac{9}{2}T.$$

Can the waiting time be reduced within the time constraints by choosing either a different set of subgraphs to cover G, or a different sequence of timings? Figure 37.55 shows the same choice of subgraphs but with different timings. The result is a slightly shorter waiting time of $\frac{22}{5}T$.

Problems

37.1 (Section 37.2). Write down the degree of each vertex in the graph in Fig. 37.56.

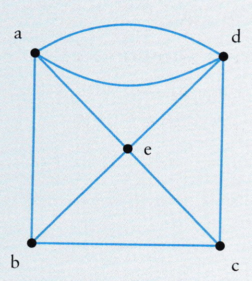


Fig. 37.56

37.2 (Section 37.2). Draw the complete graph with six vertices. How many edges does it have?

37.3 (Section 37.2). Sketch the 21 connected unlabelled graphs with five vertices. How many of them are planar?

37.4 Sketch the eight regular graphs with six vertices. How many of them are connected?

37.5 The adjacency matrix of a graph G with no loops is a vertex–vertex matrix, in which the element in the ith row and jth column is 0 if vertices i and j are not joined by an edge, and r if i and j are joined by r edges. Thus, if we list the vertices a, b, c, d as 1, 2, 3, 4 respectively, then the adjacency matrix of the graph in Fig. 37.1 is

$$A = \begin{bmatrix} 0 & 2 & 0 & 1 \\ 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

Note that the leading diagonal has zeros if there are no loops. The adjacency matrix is a formula for the graph.

Evaluate A^2. What is the interpretation of the matrix in terms of the edges of G?

37.6 Draw the graphs defined by the following adjacency matrices:

$$\text{(a)}\ A=\begin{bmatrix}0&1&1&1&1\\1&0&1&1&1\\1&1&0&1&1\\1&1&1&0&1\\1&1&1&1&0\end{bmatrix},\quad \text{(b)}\ A=\begin{bmatrix}0&2&0&0\\2&0&1&1\\0&1&0&1\\0&1&1&0\end{bmatrix}.$$

37.7 Write down the adjacency matrices of the graphs in Fig. 37.7. Note that a single loop introduces an element 1 into the appropriate position on the leading diagonal. What characterizes the matrix of a disconnected graph?

37.8 (Section 37.4). How many different cycles pass through a single vertex in a complete graph with four vertices?

37.9 (Section 37.4). List all trails between vertices a and f in the graph shown in Fig. 37.57. Identify which trails in the list are also paths.

37.10 (Section 37.4). Is the graph in Fig. 37.57 eulerian? If it is find an eulerian closed trail. Is it hamiltonian?

37.11 (Section 37.5). Construct a spanning tree for the graph shown in Fig. 37.57. Draw its cotree. Show that there is a spanning tree in which no vertex has degree more than two.

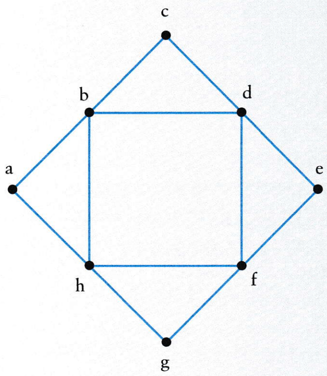


Fig. 37.57

37.12 Figure 37.58 shows a graph with seven vertices.
(a) Decide whether the graph is eulerian.
(b) Construct a spanning tree for the graph. How many branches does the tree have?
(c) Draw a cutset which disconnects the vertices a, b, g, f from the vertices c, d, e.

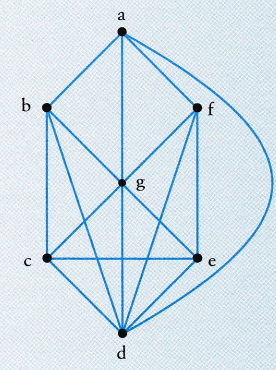


Fig. 37.58

37.13 Figure 37.59 shows a digraph. How many trails are there between a and e? Which of them are also paths? Can you find a four-edge cycle?

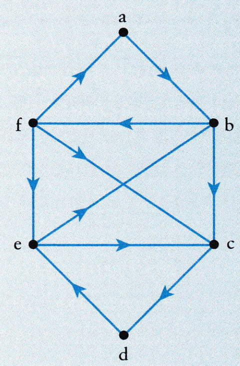


Fig. 37.59

37.14 (Section 37.6). Figure 37.60 shows a circuit with an independent current source i_0. Represent the circuit by a graph. How many vertices does the graph have?

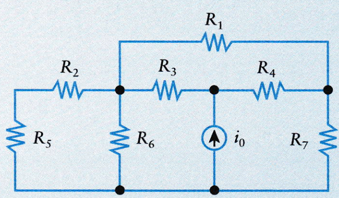


Fig. 37.60

37.15 (Section 37.6). A circuit is represented by the graph shown in Fig. 37.61. The current i_0 is from an independent source, and all other edges contain a resistor in which the current i_1 passes through a resistor R_1 and so on. Define a spanning tree for the

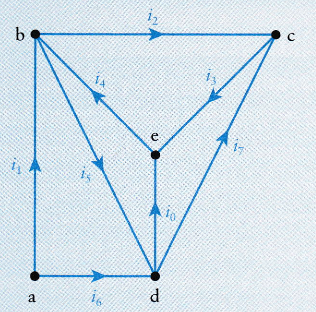


Fig. 37.61

graph. How many fundamental cutsets are required? Write down the current equations associated with each of the cutsets. If $i_1 = 2$ A, a maintained current, $v_c = 0$ (earthed), and $R_k = 1\ \Omega$ for $k = 0, 1, 2, \ldots, 7$, find the remaining voltages v_a, v_b, v_d, v_e.

37.16 (Section 37.6). Figures 37.62a,b show two circuits with current sources and resistors. Use the cutset method to find the modal voltages and currents through the resistors.

37.17 Complete the block-reduction method for the multi-feedback control system shown in Fig. 37.28.

37.18 (Section 37.5). Figure 37.63 shows a positive-feedback control system. If $P(s)$ is the system input, find its output $Q(s)$, and the transfer function of a single equivalent device.

37.19 (Section 37.7). Find the outputs in the systems shown in Figs 37.64a,b by progressively replacing parts of the system by equivalent devices until just one device remains. Find the transfer function of the resulting equivalent single device.

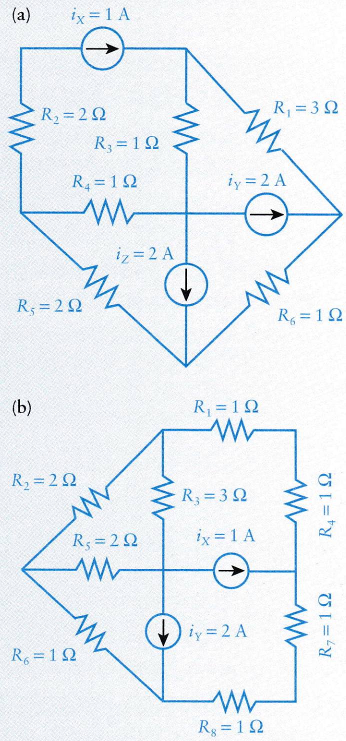


Fig. 37.62

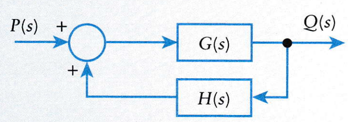


Fig. 37.63

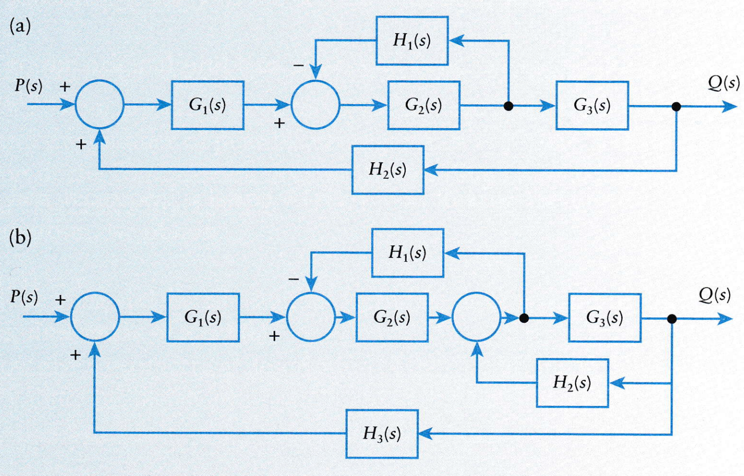


Fig. 37.64

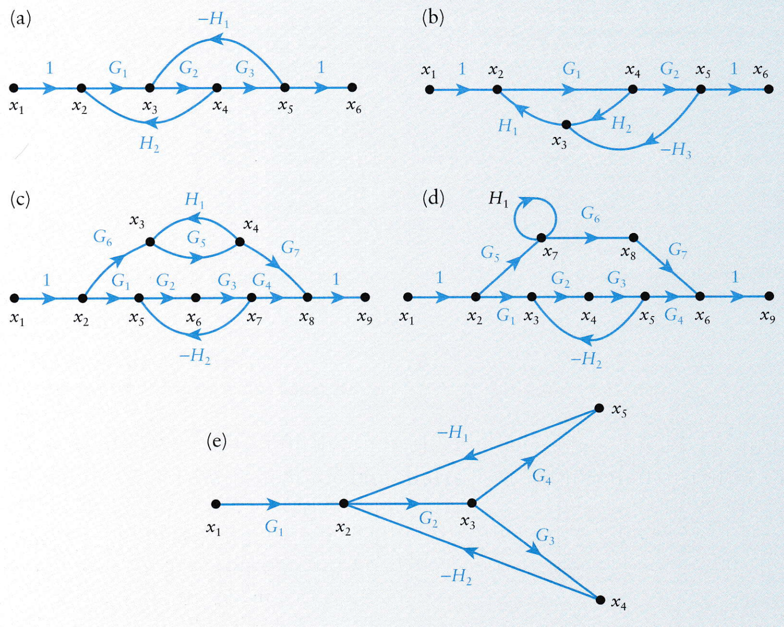


Fig. 37.65

37.20 (Section 37.7). Reduce each of the signal-flow graphs in Figs 37.65a,b,c,d to an equivalent single edge, and (e) to a stem, and find the transfer function in each case.

37.21 (Section 37.8). Label the edges, vertices, and faces of the graphs shown in Figs 37.66a,b and verify Euler's formula.

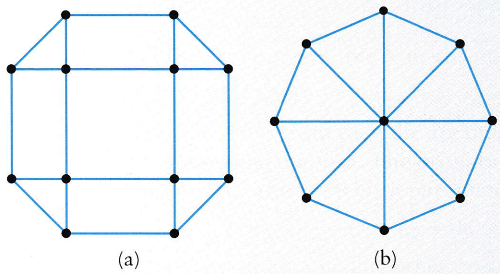


Fig. 37.66

37.22 (Section 37.8). Show that the bipartite graph $K_{2,3}$ has a planar representation.

37.23 (Section 37.8). The complete graph K_5 does not have a plane drawing. What is the minimum number of edge crossings in a plane representation of the graph?

37.24 (Section 37.1). List all the paths between S and T in the network given in Fig. 37.6, and hence find the shortest and longest paths. (This method of simply listing all paths can become very extensive for larger networks: efficient algorithms are really required to reduce the number of calculations.)

37.25 (Section 37.9). Show that the framework in Fig. 37.67 is overbraced. How many ties can be removed to leave a minimum bracing?

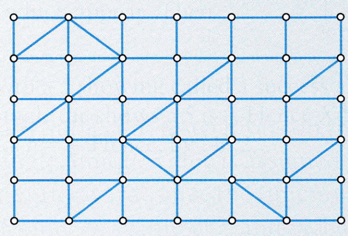


Fig. 37.67

37.26 (Section 37.9). How many ties will be needed to secure a minimum bracing for the framework shown in Fig. 37.68? Draw in a suitable set of ties for a minimum bracing.

37.27 (Section 37.9). Decide whether the frameworks shown in Fig. 37.69 are overbraced, have a minimum bracing, or are not braced.

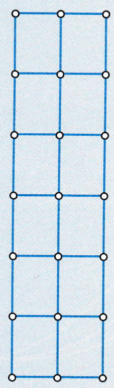


Fig. 37.68

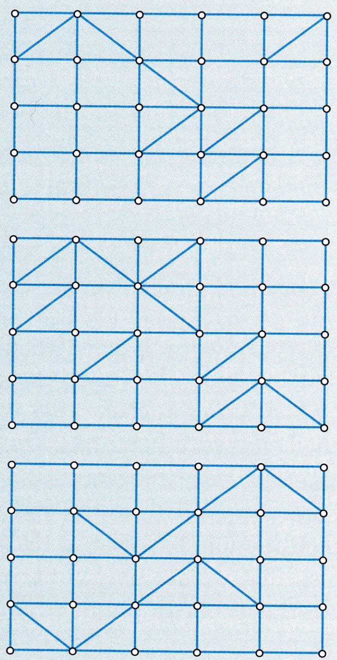


Fig. 37.69

37.28 (Section 37.9). The framework in Fig. 37.69c is required to be strengthened so that it is overbraced with each diagonal tie as an edge in at least one cycle in the associated bipartite graph. What is the minimum number of ties which must be added?

37.29 Figure 37.70 shows a junction with eight distinct lanes of traffic each controlled by a separate traffic signal. This is really a 'design and solve' problem. Here is one model: of the doubtful cases assume that lane a is compatible with both c and e, and that e is compatible with h. Draw the compatibility graph for this junction. List all complete subgraphs with four and three vertices. If the period of the traffic signal cycle is T and the subgraphs

$$\{abef, cdg, aeh\}$$

are chosen with each allowed green for $\frac{1}{3}T$, calculate the total waiting time. Suppose that the subgraph *abef* runs for $\frac{1}{2}T$ and the others for $\frac{1}{4}T$ each. How does this affect the total waiting time?

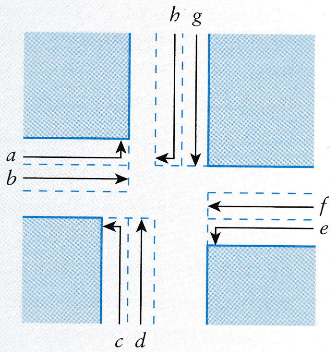


Fig. 37.70

38 Difference equations

CONTENTS

In many applications, functions can only take discrete values – that is, they cannot (for various reasons) take a continuous spectrum of values. It is reasonable to model the temperature in a room by a function which varies continuously with time – most of the calculus in this book is concerned with such functions. On the other hand, the population size of a country can only take integer values. As births and deaths occur, the population size is *discontinuous* in time, and the graph of population size against time will be a step function. Between births and deaths the population number will be constant so that we are only concerned with changes which take place at these events. In this problem jumps occur at variable time intervals.

We can obtain discrete data from a continuous signal or function by sampling the signal at regular time steps rather than keeping a continuous record. This is often the situation in microprocessor-driven operations.

The progress of events is often described in the form of equations linking several successive events: so-called difference equations. The reader may notice analogies between the solutions of these and the solutions of differential equations.

38.1 Discrete variables

Let us start by considering a simple financial application which generates discrete values. In compound interest the sum of £P_0 is invested in an account to which interest accrues annually at a compound rate of $100I\%$. If £P_1 is the amount in the account at the end of the first year, then

$$P_1 = (1 + I)P_0. \quad (38.1)$$

Let £P_n be the sum after n years. Then, similarly

$$P_n = (1 + I)P_{n-1}. \tag{38.2}$$

This is an example of a difference equation or recurrence relation. It gives the values of P_n at the integer values 1, 2, … in terms of the immediately preceding value. Treating the variable as n, the difference in this case is 1. The notation $P(n)$ instead of P_n is often used to emphasize the function aspect of P but we have chosen the more economical subscript form P_n.

It is fairly easy to solve (38.2) by repeated application of the formula starting with (38.1). Thus

$$P_2 = (1 + I)P_1 = (1 + I)^2P_0,$$

$$P_3 = (1 + I)P_2 = (1 + I)^3P_0,$$

and so the formula

$$P_n = (1 + I)^nP_0 \tag{38.3}$$

holds at least for values of n up to 3. Suppose that (38.3) holds for $n = k$. Then (38.2) implies that

$$P_{k+1} = (1 + I)P_k = (1 + I)^{k+1}P_0.$$

So the same formula holds for P_{k+1}. Hence, if the result is true for k then it is also true for $k + 1$. Equation (38.1) confirms that it is true for $k = 1$. It follows sequentially that it is true for $n = 2$, $n = 3$, and so on. (This method of proof is known as induction.)

Example 38.1 *£1000 is invested for 5 years at the following rates:* (a) 5% *annually;* (b) $\frac{5}{12}$% *calendar monthly;* (b) $\frac{5}{365}$% *daily (ignoring leap years).* (c) *Calculate the final amount in the account in each case.*

In each case the formula is

$$P_n = (1 + I)^nP_0,$$

with $P_0 = 1000$, but the I and n differ.

(a) This is the original problem with $n = 5$ and $I = 0.05$. Hence

$$P_5 = (1 + 0.05)^5 \times 1000 = 1.05^5 \times 1000 = 1276.28$$

(in £, to the nearest penny).

(b) This account has 12 *compounding periods* each year, giving a total of 60 over the 5 years. Hence we require

$$P_{60} = \left(1 + \frac{0.05}{12}\right)^{60} \times 1000 = 1283.36.$$

(c) For the daily rate, there are $365 \times 5 = 1825$ compounding periods. Thus we require

$$P_{1825} = \left(1 + \frac{0.05}{365}\right)^{1825} \times 1000 = 1284.00.$$

There is a slight gain with increasing number of compounding periods.

The following financial application of a loan repayment leads to a difference equation.

The general mortgage problem is as follows. An amount £P is borrowed for a period of N years, at an interest rate of i% per annum (as a *fraction* I this is equivalent to £(i/100) per year per pound of debt). Repayment is made by N *equal* payments £A, one at the end of every year, starting at the end of of the first year. There are two constituents of each payment A. One part goes to pay the *interest* on the debts that was carried during the previous year. The rest is used for *capital repayment* to reduce future debt. Given P, N and I, we want to know the regular annual repayment A required to exactly clear the debt at the end of year N. (There are other mortgage models that are used which calculate interest daily or monthly: the above method can be adapted by changing N to handle these cases.)

The nth payment A is made at the end of year n, after which the debt outstanding is denoted by u_0. The payment A comprises:

(interest owed on u_{n-1} through year n) + (a capital repayment)

Therefore

$$A = Iu_{n-1} + (u_{n-1} - u_n) \quad \text{or} \quad u_n = -A + (1 + I)u_{n-1} \tag{38.4}$$

where $n = 1, 2, \ldots, N$, $u_0 = P$, and the constant A is to be chosen so that the final payment clears the debt, that is $u_N = 0$.

The difference equation can be solved by step-by-step employment of the recurrence relation (38.4). In general

$u_n = -A + \beta u_{n-1}$ (we put $1 + I = \beta$, for brevity).

Start with $u_0 = P$, and calculate the sequence $u_1, u_2, u_3, \ldots, u_N$:

$$\begin{aligned} u_1 &= -A + \beta P, \\ u_2 &= -A + \beta u_1 = -A + \beta(A + \beta P) = -A(1 + \beta) + \beta^2 P, \\ u_3 &= -A + \beta u_2 = -A + \beta\{-A(1 + \beta) + \beta^2 P\} = -A(1 + \beta + \beta^2) + \beta^2 P, \end{aligned}$$

and so on – the rule for subsequent terms of the sequence is clear. Use eqn (1.36) to sum N terms of the emerging geometric series in β; then

$$u_N = -A(1 + \beta + \beta^2 + \cdots + \beta^{N-1}) + \beta^N P = -\frac{A(\beta^N - 1)}{(\beta - 1)} + \beta^N P.$$

Using the condition $u_N = 0$, we find that

$$A = \frac{I(1 + I)^N P}{(1 + I)^N - 1}.$$

Example 38.2 *The sum of £50 000 is borrowed over 25 years to be repaid in equal instalments, the interest on the outstanding balance in any year being 8%. Find the annual repayments over the term of the loan.*

In the notation above, $P = £50\,000$, $I = 0.08$, $N = 25$. Therefore the annual repayment to the nearest penny is

$$A = \frac{I(1 + I)^N P}{(1 + I)^N - 1} = \frac{0.08 \times 1.08^{25} \times 50\,000}{1.08^{25} - 1} = £4683.94$$

Example 38.2 continued

The total repayment over 25 years is

$NA = 25A = £117\,098.47.$

The capital repayment included in A at the end of year 1 is only

$u_0 - u_1 = A - IP = £683.94,$

which indicates how interest payment predominates in the early years of the mortgage.

38.2 Difference equations: general properties

Any equation of the form

$$u_n = f(u_{n-1}, u_{n-2}, \dots, u_{n-m}) \quad (38.5)$$

(where m is an integer $\geqslant 1$) for consecutive sequence of integers n, which may or may not terminate, is known as a difference equation. The term discrete dynamical system is also frequently used. Thus

$$u_n = 2u_{n-1} + 2, \quad (38.6)$$

$$u_n = 3u_{n-1} + 2u_{n-2} + n^2, \quad (38.7)$$

$$u_{n+1} = ku_n(1 - u_n) \quad (38.8)$$

are examples of difference equations.

The number m in (38.5) is known as the order of the difference equation: it is the difference between the largest and smallest subscripts attached to u, namely

$$n - (n - m) = m.$$

Thus (38.6) and (38.8) are first-order difference equations, while (38.7) is second-order. The sequence of integers attached to u can be translated (i.e. any integer can be added to the index n) without affecting the difference equation. The difference equation

$$u_{n+2} = 3u_{n+1} + 2u_n + (n + 2)^2,$$

is the same as (38.7): n has been replaced by $n + 2$ throughout, although the limits on n change.

Given initial conditions, the successive terms are very easy to compute. For a first-order difference equation, we can assume that u_0 is given, but it could be any term, say u_r, which is taken as the initial condition. Generally, our aim is to find a sequence $\{u_n\}$ and a formula for u_n for $n \geqslant r$ which satisfies the difference equation.

The difference equation (38.8) (which is known as the logistic equation) with $k = 2$ is

$$u_{n+1} = 2u_n(1 - u_n). \quad (38.9)$$

Suppose that we put $u_0 = \frac{1}{4}$; then the sequence

$$u_1 = \frac{3}{8}, \quad u_2 = \frac{15}{32}, \quad u_3 = \frac{255}{512}, \quad u_4 = \frac{65\,535}{131\,072}, \quad \dots,$$

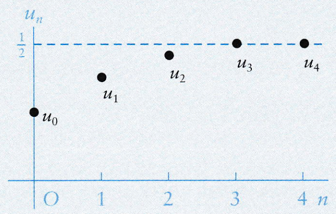


Fig. 38.1 Iterations of the sequence $u_{n+1} = 2u_n(1 - u_n)$ with $u_0 = \frac{1}{4}$.

follows by successive substitution. This sequence of numbers is actually approaching the value $\frac{1}{2}$ as n increases. We can sketch the sequence by discrete values at integer values of x in the usual cartesian axes. The series of dots in Fig. 38.1 is a graphical representation of the sequence.

The implied limiting value of u_n as $n \to \infty$ for this particular sequence suggests that $u_n = \frac{1}{2}$ is a constant solution of the difference equation (38.9), and this can be confirmed. We can find all constant solutions by simply putting $u_n = u$ for all n. From (38.9), the constant solutions are given by

$$u = 2u(1 - u), \qquad \text{or } 2u^2 - u = 0,$$

which implies that $u = 0$ and $u = \frac{1}{2}$ are solutions. These are also known as the fixed points or equilibrium values of the difference equation.

Fixed points or equilibrium values

For any first-order difference equation $u_{n+1} = f(u_n)$, its fixed points are given by solutions of

$$u = f(u). \tag{38.10}$$

You might notice, by trial computation, that the solutions of (38.9) vary quantitatively with the initial value, u_0. If $0 < u_0 < 1$, then u_n appears to approach $\frac{1}{2}$ as n becomes large, but, if $u_0 > 1$ or $u_0 < 0$, then u_n becomes unbounded for large n. We shall discuss the logistic equation further in Section 38.5.

For the second-order difference equations, the same process gives equilibrium values. For example, if

$$u_{n+2} - 2u_{n+1} + 4u_n = 6,$$

then this equation has an equilibrium value obtained by putting $u_{n+2} = u_{n+1} = u_n = u$, so that

$$u - 2u + 4u = 6 \quad \text{or} \quad u = 2.$$

On the other hand, the second-order difference equation (38.7) has no equilibrium values since

$$u - 3u - 2u - n^2 = -4u - n^2$$

can never be zero for constant u and all n.

Self-test 38.1

(38.2) The sum of £100 000 is borrowed over a 25 year term at an annual interest rate of 6.5%. Find the annual repayment assuming that the interest rate remains the same throughout. At the end of 5 years, the interest rate is increased to 7%. What should the annual repayments be increased to repay the outstanding loan over the remaining 20 years?

38.3 First-order difference equations and the cobweb

An alternative method of representing solutions of difference equations graphically is the cobweb construction. Consider the first-order difference equation

$$u_{n+1} = f(u_n) = \tfrac{1}{2}u_n + 1.$$

The equation has a fixed point or equilibrium value where

$$u = \tfrac{1}{2}u + 1 \quad \text{so that} \quad u = 2.$$

With this in view plot the lines $y = x$ and $y = \frac{1}{2}x + 1$ (Fig. 38.2) in the x,y plane. These straight lines intersect at $x = y = 2$, which corresponds to the fixed point.

Select an initial value, say, $u_0 = \frac{1}{2}$, and represent it by the point $P_0 : (u_0, 0) = (\frac{1}{2}, 0)$ in the x,y plane. From the difference equation

$$u_1 = \tfrac{1}{2}u_0 + 1 = \tfrac{1}{2} \cdot \tfrac{1}{2} + 1 = \tfrac{5}{4}.$$

We can represent this by the point $P_1 : (u_0, u_1) = (\frac{1}{2}, \frac{5}{4})$ in Fig. 38.2. Join P_0 to P_1, and then to $Q_1 : (u_1, u_1) = (\frac{5}{4}, \frac{5}{4})$ on the line $y = x$. Now join Q_1 to $P_2 = (u_1, u_2) = (\frac{5}{4}, \frac{13}{8})$ on $y = \frac{1}{2}x + 1$. Repeat the process by drawing lines between $y = x$ and $y = \frac{1}{2}x + 1$ using the same rules.

The usefulness of the method is that a graphical representation and interpretation of the solutions can be achieved by simple line drawings as shown in Fig. 38.2 and in the following example. It is particularly helpful for finding fixed points and assessing their stability. The connected lines are known as cobwebs for obvious reasons. We can observe that this difference equation has only one fixed point at (2, 2), which is stable, since all cobwebs approach the point form *any* initial point.

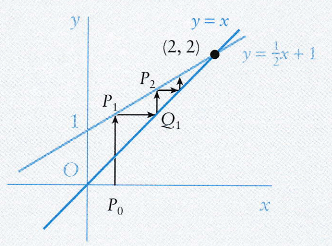


Fig. 38.2

For a general difference equation $u_{n+1} = f(u_n)$, the cobweb construction takes place between the straight line $y = x$ and the curve $y = f(x)$.

Example 38.3 *Sketch a cobweb solution for*

$$u_{n+1} = -ku_n + k,$$

for (a) $k = \frac{1}{2}$, (b) $k = \frac{3}{2}$, (c) $k = 1$, *using the initial value* $u_0 = \frac{3}{4}$ *in each case.*

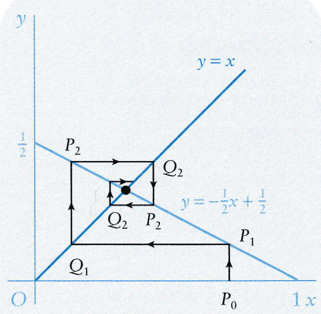


Fig. 38.3 Cobweb for $u_{n+1} = -\frac{1}{2}u_n + \frac{1}{2}$ with $u_0 = \frac{3}{4}$.

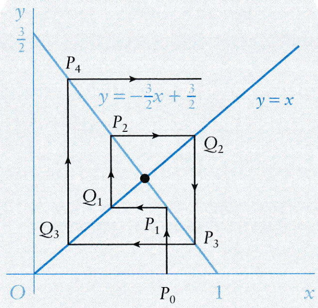


Fig. 38.4 Cobweb for $u_{n+1} = -2u_n + 2$ with $u_0 = \frac{3}{4}$.

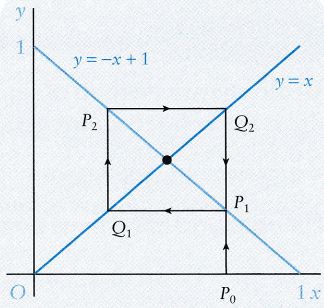


Fig. 38.5 Cobweb for $u_{n+1} = -u_n + 1$ with $u_0 = \frac{3}{4}$.

(a) Plot the lines $y = x$ and $y = -\frac{1}{2}x + \frac{1}{2}$. They intersect at the fixed point $(\frac{1}{3}, \frac{1}{3})$. Starting from $P_0 : (\frac{3}{4}, 0)$, the cobweb traces $P_0P_1Q_1P_2Q_2P_3 \ldots$ in Fig. 38.3. Evidently it approaches the fixed point as $n \to \infty$, indicating stability.

(b) The lines are $y = x$ and $y = -\frac{3}{2}x + \frac{3}{2}$. The fixed point is at $(\frac{3}{5}, \frac{3}{5})$, and the cobweb path is $P_0P_1Q_1P_2Q_2 \ldots$ in Fig. 38.4. The path moves away from the fixed point implying its instability.

(c) The lines are $y = x$ and $y = -x + 1$ with fixed point $(\frac{1}{2}, \frac{1}{2})$. The path starting at P_0: $(\frac{3}{4}, 0)$ follows the rectangle $P_1Q_1P_2Q_2$, indicating periodicity (Fig. 38.5). This is true for any starting value except that of the fixed point itself.

Graphs of the sequences u_n versus n are shown in Fig. 38.6.

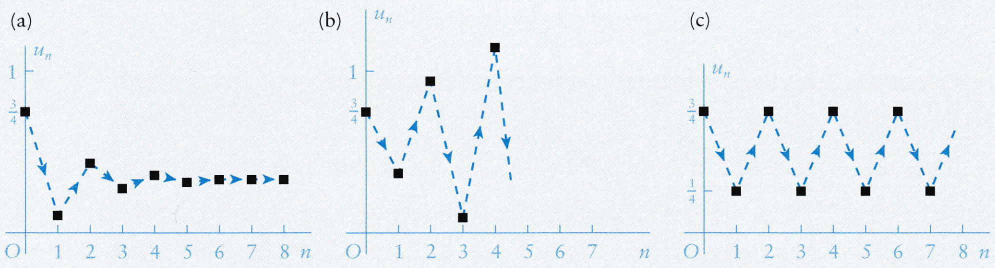


Fig. 38.6 Solutions of $u_{n+1} = -k(u_n + 1)$ for (a) $k = \frac{1}{2}$, (b) $k = 2$, (c) $k = 1$.

The stability of the fixed point of the general first-order linear difference equation can be summarized as follows.

Stability

The first-order difference equation $u_{n+1} = -ku_n + a$ has a fixed point at $u = a/(1+k)$, $(k \neq -1)$. The fixed point is stable if $|k| < 1$, unstable if $|k| > 1$, and periodic if $k = 1$.

If $k = -1$, the equation has no fixed point unless $a = 0$. (38.11)

Self-test 38.2

Consider the difference equation $u_{n+1} = f(u_n) = \frac{1}{2} - u_n^2$. Plot the curve $y = f(x) = \frac{1}{2} - x^2$ and the straight line $y = x$. What are the coordinates of the fixed point in the x,y plane? Given $u_1 = 0.2$, compute u_2, u_3, u_4, u_5. Draw the corresponding cobweb. Does it indicate stability of the fixed point?

38.4 Constant-coefficient linear difference equations

Any difference equation of the form

$$u_n + a_{n-1}u_{n-1} + \cdots + a_{n-m}u_{n-m} = f(n),$$

where the a_i $(i = n - m, \ldots, n-1)$ are constants, is a constant-coefficient linear difference equation. We shall look in detail at the second-order case

$$u_{n+2} + 2au_{n+1} + bu_n = f(n), \tag{38.12}$$

where a and b are constants and $f(n)$ is a given function. The methods generalize in a fairly obvious way to higher-order systems.

There are many parallels between the difference equation (38.12) and second-order constant-coefficient equations (Chapters 18–19). The equation is said to be **homogeneous** if $f(n) = 0$, and **inhomogeneous** otherwise, just as in the case of second-order differential equations. However, this section is self-contained and reference back is not necessary. The general solution of the inhomogeneous case requires that of the homogeneous case: hence we start with the latter.

Homogeneous equations

We can see how to proceed by looking at the first-order constant-coefficient equation

$$u_{n+1} - cu_n = 0. \tag{38.13}$$

As can be seen from (38.2) or verified directly, the general solution of this equation is

$$u_n = Ac^n, \tag{38.14}$$

where A is any constant. Notice that we could equally well write

$$u_n = Ac^{n-1}, \quad \text{or} \quad u_n = Ac^{n+1}:$$

the result would be equally correct, although A would take different values for the same initial condition. The significant property of (38.13) and its solution (38.14) is that u_{n+1} is a constant multiple of u_n.

With this in view, we attempt to find solutions of

$$u_{n+2} + 2au_{n+1} + bu_n = 0 \tag{38.15}$$

in the form $u_n = p^n$, where p is a constant. Thus

$$u_{n+2} + 2au_{n+1} + bu_n = p^{n+2} + 2ap^{n+1} + bp^n = (p^2 + 2ap + b)p^n = 0,$$

for all n, if $p = 0$ or

$$p^2 + 2ap + b = 0. \tag{38.16}$$

The case $p = 0$ leads to the self-evident solution $u_n = 0$. We are interested in solutions of (38.16), which is known as the characteristic equation of (38.15).

There are various cases to consider. Suppose that the roots of (38.16) are the distinct numbers p_1 and p_2. Hence $u_n = p_1^n$ and $u_n = p_2^n$ are solutions of (38.15). Since this equation is homogeneous and *linear*, it follows that any linear combination of p_1^n and p_2^n is also a solution. We state this as follows.

Distinct roots

The general solution of $u_{n+2} + 2au_{n+1} + bu_n = 0$ for distinct roots p_1 and p_2 of $p^2 + 2ap + b = 0$ is

$$u_n = Ap_1^n + Bp_2^n, \quad \text{for any constants } A \text{ and } B. \tag{38.17}$$

Example 38.4 *Find the general solution of*

$$u_{n+2} - u_{n+1} - 6u_n = 0. \tag{38.18}$$

The characteristic equation of (38.18) is

$$p^2 - p - 6 = 0, \quad \text{or} \quad (p-3)(p+2) = 0.$$

The roots are $p_1 = 3$, $p_2 = -2$. Hence the general solution is

$$u_n = A \cdot 3^n + B(-2)^n.$$

Example 38.5 *Find the solution of*

$$u_{n+2} + 2u_{n+1} - 3u_n = 0$$

that satisfies $u_0 = 1$, $u_1 = 2$.

The characteristic equation is

$$p^2 + 2p - 3 = 0, \quad \text{or} \quad (p+3)(p-1) = 0.$$

The roots are $p_1 = -3$, $p_2 = 1$. Hence the general solution is

$$u_n = A(-3)^n + B \cdot 1^n = A(-3)^n + B.$$

From the initial conditions,

$$u_0 = 1 = A + B, \quad u_1 = 2 = -3A + B.$$

Hence $A = -\frac{1}{4}$ and $B = \frac{5}{4}$. The required solution is

$$u_n = -\tfrac{1}{4} \cdot (-3)^n + \tfrac{5}{4}.$$

The characteristic equation can have equal roots, which is a special case. Consider the difference equation

$$u_{n+2} - 2au_{n+1} + a^2u_n = 0,$$

where $a \neq 0$. Its characteristic equation is

$$p^2 - 2ap + a^2 = 0, \quad \text{or} \quad (p-a)^2 = 0,$$

which has the repeated root $p = a$. One solution is Aa^n; but we require a second independent solution. Consider the expression $u_n = na^n$. Then

$$\begin{aligned} u_{n+2} - 2au_{n+1} + a^2u_n &= (n+2)a^{n+2} - 2(n+1)a^{n+2} + na^{n+2} \\ &= a^{n+2}(n + 2 - 2(n+1) + n) = 0. \end{aligned}$$

Hence a further independent solution is $u_n = Bna^n$.

Equal roots

The general solution of $u_{n+2} - 2au_{n+1} + a^2u_n = 0$ is

$$u_n = (A + Bn)a^n. \tag{38.19}$$

Roots can also be complex. Consider the difference equation

$$u_{n+2} + 2u_{n+1} + 2u_n = 0.$$

Its characteristic equation is

$$p^2 + 2p + 2 = 0$$

with roots $p_1 = -1 + \mathrm{i}$, $p_2 = -1 - \mathrm{i}$. The method still works and the general solution becomes

$$u_n = A(-1+\mathrm{i})^n + B(-1-\mathrm{i})^n.$$

For a real-valued problem, the constants A and B will be complex conjugates which ensure that u_n is real. The solution can be cast in real form by using the polar forms (Section 6.3) of the complex numbers. In this case

$$-1 \pm \mathrm{i} = \sqrt{2}\,\mathrm{e}^{\pm\frac{3}{4}\pi\mathrm{i}}.$$

Hence

$$\begin{aligned} u_n &= A2^{\frac{1}{2}n}\mathrm{e}^{\frac{3}{4}\pi\mathrm{i}n} + B2^{\frac{1}{2}n}\mathrm{e}^{-\frac{3}{4}\pi\mathrm{i}n} \\ &= 2^{\frac{1}{2}n}[A(\cos\tfrac{3}{4}\pi n + \mathrm{i}\sin\tfrac{3}{4}\pi n) + B(\cos\tfrac{3}{4}\pi n - \mathrm{i}\sin\tfrac{3}{4}\pi n)] \\ &= 2^{\frac{1}{2}n}(C\cos\tfrac{3}{4}\pi n + D\sin\tfrac{3}{4}\pi n), \end{aligned}$$

where $C = A + B$ and $D = (A - B)\mathrm{i}$.

Complex roots, $\alpha \pm \mathrm{i}\beta = r\,\mathrm{e}^{\pm\theta\mathrm{i}}$

The general complex solution of

$$u_{n+2} + 2au_{n+1} + bu_n = 0,$$

where $a^2 < b$, is

$$u_n = A(\alpha + \mathrm{i}\beta)^n + B(\alpha - \mathrm{i}\beta)^n.$$

The general real solution is

$$u_n = r^n(C\cos n\theta + D\sin n\theta). \tag{38.20}$$

Example 38.6 *Obtain the general solution of*

$$u_{n+2} + u_n = 0.$$

The characteristic equation is

$$p^2 + 1 = 0,$$

giving roots $p_1 = \mathrm{i}$, $p_2 = -\mathrm{i}$. Hence

$$u_n = A\mathrm{i}^n + B(-\mathrm{i})^n.$$

In polar form, $\mathrm{i} = e^{\frac{1}{2}\pi\mathrm{i}}$, $-\mathrm{i} = e^{-\frac{1}{2}\pi\mathrm{i}}$. Hence the real form of the solution is

$$u_n = C\cos\tfrac{1}{2}\pi n + D\sin\tfrac{1}{2}\pi n.$$

Inhomogeneous equations

The general inhomogeneous equation is

$$u_{n+2} + 2au_{n+1} + bu_n = f(n) \tag{38.21}$$

(see (38.12)). Let $u_n = v_n + q_n$, where v_n is the *general solution* of the corresponding *homogeneous equation*. Substitute this form of u_n into (38.21):

$$(v_{n+2} + q_{n+2}) + 2a(v_{n+1} + q_{n+1}) + b(v_n + q_n) = f(n),$$

or

$$(v_{n+2} + 2av_{n+1} + bv_n) + (q_{n+2} + 2aq_{n+1} + bq_n) = f(n).$$

Since v_n satisfies the homogeneous equation, it follows that

$$q_{n+2} + 2aq_{n+1} + bq_n = f(n),$$

which means that q_n must be a particular solution of the inhomogeneous equation. As in differential equations, v_n is known as the complementary function.

We construct particular solutions by appropriate choices of functions usually containing adjustable parameters which are suggested by the form of the function $f(n)$. If a particular choice fails, then we reject it and try something else.

Example 38.7 *Obtain the general solution of*

$$u_{n+2} - u_{n+1} - 6u_n = 4.$$

From Example 38.4, the complementary function is

$$v_n = 3^nA + (-2)^nB.$$

For the particular solution, we try $q_n = C$, since $f(n) = 4$. Then

$$q_{n+2} - q_{n+1} - 6q_n - 4 = C - C - 6C - 4 = -6C - 4 = 0,$$

if $C = -\frac{2}{3}$. Hence $q_n = -\frac{2}{3}$, and the general solution is

$$u_n = 3^nA + (-2)^nB - \tfrac{2}{3}.$$

Example 38.8 *Obtain the general solution of*

$$u_{n+2} + 2u_{n+1} - 3u_n = 4.$$

From Example 38.5, the complementary function is

$$v_n = (-3)^nA + B.$$

Example 38.8 continued

In this case we expect the choice $q_n = C$ to fail, since it must make the left-hand side of the difference equation vanish. When this happens, we try

$$q_n = Cn.$$

Then

$$q_{n+2} + 2q_{n+1} - 3q_n - 4 = C(n+2) + 2C(n+1) - 3Cn - 4 = 2C + 2C - 4 = 4C - 4 = 0,$$

if $C = 1$. Hence the general solution is

$$u_n = (-3)^n A + B + n.$$

Table 38.1 lists some simple forcing terms $f(n)$ with suggested forms of particular solution and alternatives containing parameters to be determined by direct substitution.

Table 38.1

$f(n)$	Trial solution q_n
k (a constant)	C; or Cn, if C fails; or Cn^2, if C and Cn fail; etc.
k^n	Ck^n; or Cnk^n, if Ck^n fails; etc.
n	$C_0 + C_1 n$
n^p (p an integer)	$C_0 + C_1 n + \cdots + C_p n^p$ (may need higher powers of n in special cases)
$\sin kn$ or $\cos kn$	$C_1 \cos kn + C_2 \sin kn$

Example 38.9 *Find the general solution of*

$$u_{n+2} - 4u_n = n.$$

The characteristic equation is

$$p^2 - 4 = 0, \quad \text{or} \quad (p-2)(p+2) = 0.$$

The roots are $p_1 = 2$, $p_2 = -2$. Hence the complementary function is

$$v_n = 2^n A + (-2)^n B.$$

For the particular solution, try (choosing from Table 38.1)

$$q_n = C_0 + C_1 n.$$

Then

$$q_{n+2} - 4q_{n+1} - n = C_0 + C_1(n+2) - 4C_0 - 4C_1 n - n = (-3C_0 + 2C_1) + n(-3C_1 - 1).$$

The right-hand side vanishes for all n if

$$-3C_0 + 2C_1 = 0, \qquad -3C_1 - 1 = 0.$$

Hence $C_1 = -\frac{1}{3}$, $C_0 = 2C_1/3 = -\frac{2}{9}$, and the general solution is

$$u_n = 2^n A + (-2)^n B - \tfrac{2}{9} - \tfrac{1}{3}n.$$

Self-test 38.3

Find the general solution of $u_{n+2} - 4u_{n+1} + 4u_n = 2^n$.

38.5 The logistic difference equation

Consider again the logistic difference equation

$$u_{n+1} = \alpha u_n(1 - u_n), \tag{38.22}$$

where α is a parameter which will take various values. This *nonlinear* equation can model population growth of generations. If u_n represents the population size of generation n and α is the birthrate, then we might expect the population size of the next generation to be αu_n in the absence of any inhibiting factors such as lack of resources or overcrowding. If $\alpha > 1$, then the population model given by the first-order difference equation $u_{n+1} = \alpha u_n$ would imply that the population would grow to infinity, since the equation has the solution $u_n = \alpha^n u_0$. To counter this possibility, we can introduce a feedback term $-\alpha u_n^2$ which will tend to reduce population growth when the population is large.

Fixed points of the equation (38.22) occur where

$$u = \alpha u(1 - u);$$

that is, for $u = 0$ and $u = 1 - 1/\alpha$. We can adapt the cobweb method of Section 38.3 to this nonlinear difference equation by plotting graphs of the parabola $y = f(x) = \alpha x(1 - x)$ and the straight line $y = x$. Fixed points of the difference equation occur where the line and the parabola intersect. The values of x at these points are given by the solutions of

$$\alpha x(1 - x) = x, \quad \text{or} \quad x(\alpha x - 1 - \alpha) = 0.$$

In the cobweb, the fixed points have coordinates $(0, 0)$ and $P : (1 - (1/\alpha), 1 - (1/\alpha))$. We shall only look at values of $\alpha > 1$, so that one fixed point is in the first quadrant, $x > 0$, $y > 0$. A cobweb solution starting at for the case $\alpha = 2.8$ is shown in Fig. 38.7.

Notice that, for this choice of α and u_0, the fixed point P appears to be *stable*; that is, the cobweb solution approaches P. The slope of the graph of $y = \alpha x(1 - x)$ at P determines the stability or instability of the solutions. The slope at P is

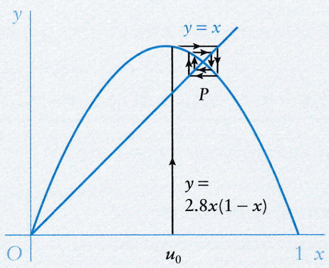


Fig. 38.7 Cobweb solution for $u_{n+1} = 2.8u_n(1 - u_n)$ showing a solution starting from $x = u_0$ approaching the fixed point at P.

$$mf'(1-(1/\alpha)) = \alpha - 2\alpha(1-(1/\alpha)) = -\alpha + 2.$$

As with the cobweb for two intersecting lines for the linear difference equation in Section 38.3, the fixed point P is locally stable if $m = 2 - \alpha > -1$, in that all cobweb paths starting close to $(\alpha - 1)/\alpha$ approach the fixed point P as $n \to \infty$. This inequality implies that $\alpha < 3$. Notice also that, if $1 < \alpha < 2$, then $y = x$ intersects the parabola $y = \alpha x(1-x)$ between the origin and its maximum value. This follows since the maximum occurs at $x = \frac{1}{2}$ and $0 < 1 - 1/\alpha < \frac{1}{2}$ implies $1 < \alpha < 2$.

For $\alpha \geqslant 3$ the solutions become more complicated. The fixed point at the origin is unstable: hence there is no *stable* fixed point to which solutions can approach. We can obtain a clue as to what happens if we look at the function of a function given by

$$y = f(f(x)) = \alpha[\alpha x(1-x)][1 - \alpha x(1-x)] = \alpha^2 x(1-x) - \alpha^3 x^2(1-x)^2.$$

When $\alpha = 3$, this curve intersects $y = x$ at $x = 0$ and at P only. This can be checked by noting that fixed points can be found from

$$x = 9x(1-x) - 27x^2(1-x)^2$$

which can be written as

$$x(27x^3 - 54x^2 + 36x - 8) = 0, \quad \text{or} \quad x(3x-2)^3 = 0.$$

Graphs of the curves $y = f(x)$ and $y = f(f(x))$ for $\alpha = 3$ are shown in Fig. 38.3a. The fixed point P on $y = f(x)$ is at $(\frac{2}{3}, \frac{2}{3})$. As α increases two additional fixed points develop on the line $y = x$. Further graphs of the two functions $y = f(x)$ and $y = f(f(x))$ for $\alpha = 3.4$ are shown in Fig. 38.8b, together with the line $y = x$. The graph indicates that there are now four fixed points at O, A, B, C.

For general α, the fixed points of $y = f(f(x))$ occurs where

$$x = \alpha^2 x(1-x) - \alpha^3 x^2(1-x)^2,$$

or

$$x(1 - \alpha - \alpha x)[\alpha^2 x^2 - \alpha(1+\alpha)x + 1 + \alpha] = 0,$$

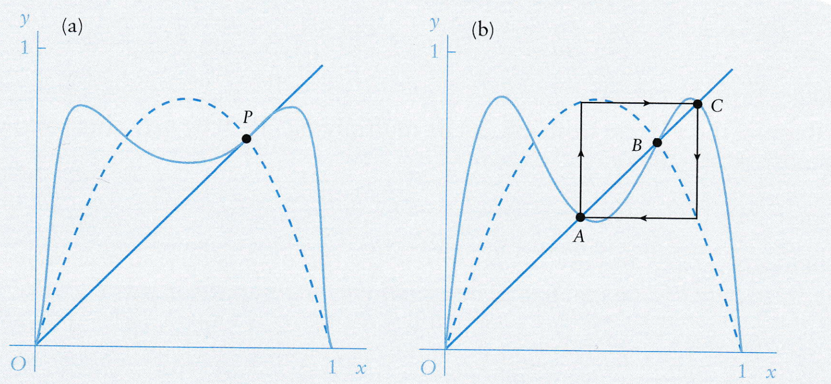


Fig. 38.8 (a) Graph of $y = f(f(x))$ for the critical case $\alpha = 3$. (b) Graph of $y = f(f(x))$ for $\alpha = 3.4$ showing fixed points O, A, B, C. The dashed curve shows $y = f(x)$ in both cases.

where we could have predicted the solution $x = (1 - \alpha)/\alpha$ corresponding to the point B in Fig. 38.8b. The solutions of

$$\alpha^2 x^2 - \alpha(1+\alpha)x + 1 + \alpha = 0 \tag{38.23}$$

are

$$\left.\begin{matrix} x_1 \\ x_2 \end{matrix}\right\} = \frac{1}{2\alpha}[1+\alpha \mp \surd\{(\alpha+1)(\alpha-3)\}] \quad (\alpha > 3)$$

which determine, respectively, the coordinates of A and C.

From (38.23)

$$x_1 + x_2 = (1+\alpha)/\alpha. \tag{38.24}$$

Also

$$\begin{aligned} f(x_1) &= \alpha x_1(1-x_1) = \alpha x_1 - \alpha x_1^2 \\ &= \alpha x_1 - (1/\alpha)[\alpha(1+\alpha)x_1 - 1 - \alpha] \quad \text{(using (38.23))} \\ &= (1/\alpha)(-\alpha x_1 + 1 + \alpha) = x_2 \end{aligned}$$

by eqn (38.24). Similarly $f(x_2) = x_1$.

It follows that

$$f(f(x_1)) = f(x_2) = x_1 \quad \text{and} \quad f(f(x_2)) = f(x_1) = x_2.$$

Hence if $x = x_1$ initially then subsequently x alternates between x_1 and x_2 shown by the square in Fig. 38.8b. This phenomenon is known as period doubling. The values $x = x_1$ and $x = x_2$ are fixed points of $y = f(f(x))$, and their stability is determined by the slopes of $y = f(f(x))$ at the points.

The critical slopes for stability at A and C are both (−1); we now find the value of α at which this occurs. We have

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}x} f(f(x)) = \alpha^2 - 2\alpha^2 x - \alpha^3(2x - 6x^2 + 4x^3) \\ &= \alpha^2 - 2\alpha^2(1+\alpha)x + 6\alpha^3 x^2 - 4\alpha^3 x^3. \end{aligned} \tag{38.25}$$

We require the value of α given by

$$\frac{\mathrm{d}}{\mathrm{d}x} f(f(x)) = -1 \quad \text{or} \quad 4\alpha^3 x^3 - 6\alpha^3 x^2 + 2\alpha^2(1+\alpha)x - \alpha^2 = 1, \tag{38.26}$$

when x satisfies (38.23).

Remove the x^3 term from (38.26) by multiplying (38.3) by $4\alpha x$, and subtracting it from (38.26). Then

$$-2\alpha^2(\alpha-2)x^2 + 2\alpha(\alpha-2)(\alpha+1)x - (1+\alpha^2) = 0. \tag{38.27}$$

Equations (38.26) and (38.27) must have the *same* roots in x. In each case, make the coefficient of x^2 equal to 1. The equations for comparison are

$$x^2 - \frac{(\alpha+1)}{\alpha}x + \frac{(\alpha+1)}{\alpha^2} = 0,$$

$$x^2 - \frac{(\alpha+1)}{\alpha}x + \frac{(\alpha^2+1)}{2\alpha^2(\alpha-2)} = 0.$$

These equations have the same roots if

$$\frac{\alpha+1}{\alpha^2} = \frac{(\alpha^2+1)}{2\alpha^2(\alpha-2)},$$

or

$$\alpha^2 - 2\alpha - 5 = 0. \quad (38.29)$$

We are interested in values of $\alpha > 3$, so that the required root of (38.29) is $\alpha = 1 + \surd 6 = 3.449\ldots$. In fact the slopes at both A and C both become -1 for this value of α. Thus, for

$$3 < \alpha < 1 + \surd 6,$$

the 2-cycle solution is stable.

As α in creases from 1, the stable fixed point at $x = (\alpha - 1)/\alpha$ becomes unstable at $\alpha = 3$. This **bifurcates** into a stable period 2 solution.

At $\alpha = 1 + \surd 6$, the system bifurcates again into a 4-cycle or period-4 solution, which corresponds to the set of stable fixed points of $y = f(f(f(f(x))))$. A graph of this function for $\alpha = 3.54$ is shown in Fig. 38.9 together with the eight fixed points. The cycle doubles again at about $\alpha = 3.544, \ldots$ and so on. The intervals between the bifurcations of the period doubling rapidly decrease, until a limit is reached at about $\alpha = 3.570, \ldots$ beyond which **chaos** occurs. The iterations are no longer periodic for most values of α beyond this point, although there are some brief intervals of periodicity.

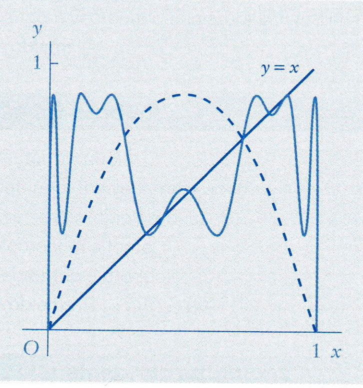


Fig. 38.9 Fixed points of $y = f(f(f(f(x))))$ for $\alpha = 3.54$, given by the intersection of the curve and the line $y = x$.

Logistic equation

$$u_{n+1} = f(u_n) = \alpha u_n(1 - u_n).$$

Fixed point for $\alpha > 0$, $x > 0$ at $x_0 = (\alpha - 1)/\alpha$.
Fixed point x_0 stable if

$$f'(x_0) = \alpha - 2\alpha x_0 = -\alpha + 2 > -1, \text{ that is if } \alpha < 3.$$

Period-2 solution: fixed points ($\alpha > 3$)

$$x_1, x_2 = \{1 + \alpha \mp \surd[(\alpha+1)(\alpha-3)]\}/(2\alpha).$$

Period-2 solution stable if $3 < \alpha < 1 + \surd 6$. (38.28)

The sequence of period-doubling bifurcations is known as the Feigenbaum sequence, and it has certain universal aspects in that it is not just a consequence of the logistic equation, but has common features with other difference equations which generate period doubling.

The simplest way to view the progressively complex behaviour is through a computer-drawn picture of the iterations of

$$u_{n+1} = \alpha u_n(1 - u_n)$$

for stepped increases in α starting at $\alpha = 2.8$ up to $\alpha = 3.8$, which covers the main area of interest. The result is shown in Fig. 38.10. The series of single dots for each α in $2.8 \leqslant \alpha \leqslant 3$ indicates the fixed point, which then bifurcates into a stable 2-cycle attractor for $3 < \alpha \leqslant 1 + \sqrt{6}$. This in turn bifurcates into a stable 4-cycle attractor at $\alpha = 1 + \sqrt{6}$ and so on. The effect of infinite period doubling is that the solution is ultimately non-periodic. The generally chaotic and noisy behaviour of the difference equation can clearly be seen in the large number of dots for larger values of α. These non-periodic sets are known as strange attractors. The successive iterates of the logistic equation wander about in a seemingly random but bounded manner, and never settle into a periodic solution. However, within the chaotic band of α values, there appear windows of periodic cycles. Problem 38.26, for example, confirms that there is a 3-cycle around $\alpha = 3.83$.

The logistic equation can be thought of as a relatively simple model example. Many similar nonlinear difference equations also exhibit similar period-doubling bifurcations and strange attractors.

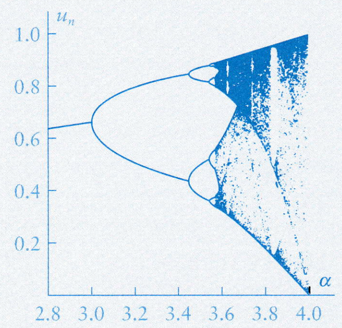


Fig. 38.10 Period doubling for the logistic equation for increasing α, followed by chaotic iterations beyond about $\alpha = 3.57$.

Problems

38.1 £1000 is invested over 10 years at an interest rate of 6% annually. Find the final total investment. What should the *monthly* interest rate be to achieve the same final total?

38.2 The sum of £50 000 is borrowed over 25 years and the money is repaid in equal annual instalments. The interest rate on the outstanding balance in any year is 10%. Find what the annual repayments would be. After 5 years, the interest rate is reduced to 9%.
(a) Find the required adjustment to the annual repayments for the loan to be repaid over the original term.
(b) If the repayments are not changed, by how much will the mortgage term be reduced?

38.3 Find the fixed points of the following difference equations:
(a) $u_{n+1} = u_n(2 - u_n)$;
(b) $u_{n+1} = u_n(1 + u_n)(2 - 3u_n)$;
(c) $u_{n+1} = \sin u_n$; (d) $u_{n+1} = \frac{1}{2}\sin u_n$;
(e) $u_{n+1} = e^{u_n} - 1$.

38.4 Given the initial value u_0 in each case, calculate the sequence of terms up to u_5 for each of the following first-order difference equations:
(a) $u_{n+1} = 2u_n(3 - u_n)$, $u_0 = 1$;
(b) $u_{n+1} = 2u_n(1 - u_n)$, $u_0 = \frac{1}{2}$;
(c) $u_{n+1} = 3.2u_n(1 - u_n)$, $u_0 = \frac{1}{2}$;
(d) $u_{n+1} = 4u_n(1 - u_n)$, $u_0 = \frac{1}{2}$.

38.5 (Section 38.3). Sketch the cobweb solutions for the following first-order equations with the stated initial conditions, and discuss the stability of the fixed point:
(a) $u_{n+1} = \frac{1}{2}u_n + \frac{1}{2}$, $u_0 = \frac{1}{2}$ and $u_0 = \frac{3}{2}$;
(b) $u_{n+1} = 2u_n - 2$, $u_0 = \frac{1}{2}$ and $u_0 = \frac{3}{2}$;
(c) $u_{n+1} = -u_n + 2$, $u_0 = \frac{1}{2}$ and $u_0 = \frac{3}{4}$;
(d) $u_{n+1} = -\frac{1}{2}u_n + \frac{3}{2}$, $u_0 = \frac{1}{2}$ and $u_0 = \frac{3}{2}$;
(e) $u_{n+1} = -2u_n + 3$, $u_0 = \frac{1}{2}$ and $u_0 = \frac{3}{2}$.

38.6 The function $f(n)$ satisfies

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

Put $n = 2^m$ and $g(m) = f(2^m)$, and show that

$$g(m) = g(m - 1) + 1.$$

Hence find $f(n)$ given that $f(1) = 0$.

38.7 Use the method suggested in the previous problem to solve

$$f(n) = f(\tfrac{1}{3}n) + \tfrac{5}{8},$$

given the initial condition $f(1) = 0$.

38.8 (Section 38.3). Find the general solutions of the following difference equations:
(a) $u_{n+2} + 2u_{n+1} - 3u_n = 0$;
(b) $u_{n+2} - 9u_n = 0$;
(c) $u_{n+2} + 9u_n = 0$;
(d) $u_n - 4u_{n-1} + 5u_{n-2} = 0$;
(e) $u_{n+2} - 4u_{n+1} + 4u_n = 0$;
(f) $u_{n+3} - u_{n+2} + u_{n+1} - u_n = 0$;
(g) $u_{n+3} - u_n = 0$;
(h) $u_{n+3} - 3u_{n+2} + 3u_{n+1} - u_n = 0$;
(i) $u_{n+2} - u_{n+1} - u_n + u_{n-1} = 0$.

38.9 Express the solution of the initial-value problem

$$u_{n+2} - 6u_{n+1} + 13u_n = 0, \qquad u_0 = 0, \quad u_1 = 1,$$

in real form.

38.10 Find the difference equation satisfied by

$$u_n = A\cdot 2^n + B\cdot(-5)^n,$$

for all A and B.

38.11 Obtain particular solutions of the following inhomogeneous difference equations:
(a) $u_{n+2} + 2u_{n+1} - 3u_n = f(n)$, where
(i) $f(n) = 2^n$; (ii) $f(n) = n$; (iii) $f(n) = 2$
(iv) $f(n) = (-3)^n$.
(b) $u_{n+2} + 2u_{n+1} + 2u_n = f(n)$, where
(i) $f(n) = 1$; (ii) $f(n) = n + 3$;
(iii) $f(n) = \cos\frac{3}{4}\pi n$.
(c) $u_{n+3} - 3u_{n+2} + 3u_{n+1} + u_n = f(n)$, where
(i) $f(n) = 1$; (ii) $f(n) = n$; (iii) $f(n) = n^2$.
(d) $u_{n+2} - 6u_{n+1} + 9u_n = f(n)$, where (i) $f(n) = 2^n$;
(ii) $f(n) = 3$; (iii) $f(n) = 3^n$; (iv) $f(n) = n3^n$.

38.12 A ball bearing is dropped from a height $z = h_0$ on to a metal plate, and the coefficient of restitution between the ball and the plate is ε, where $0 < \varepsilon < 1$. Set up a difference equation for the maximum height reached after n impacts. Solve the equation. (Assume that a ball dropped from a height h hits the plate with speed $v = \sqrt{(2gh)}$, where g is the acceleration due to gravity. The rebound speed of the ball is εv.) Instead of being stationary, the plate now oscillates so that it is moving upwards at a speed u (a constant) at the moment of each impact with the ball. Find the difference equation for h_n. Show that the difference equation has a fixed point and interpret its meaning.

38.13 $D_n(x)$ is the $n \times n$ determinant defined by

$$D_n(x) = \begin{vmatrix} 2x & 1 & 0 & \dots & 0 \\ 1 & 2x & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2x \end{vmatrix} \quad (n > 2),$$

$$D_2(x) = \begin{vmatrix} 2x & 1 \\ 1 & 2x \end{vmatrix}, \quad D_1(x) = 2x.$$

Show that

$$D_n(x) = 2xD_{n-1}(x) - D_{n-2}(x).$$

Solve the difference equation for $x \neq 1$ and $x = 1$.

38.14 Let $\{u_n\}$ $(n = 0, 1, \dots)$ be a sequence. The power series

$$f(u_n, x) = \sum_{n=0}^{\infty} u_n x^n$$

is known as the generating function of the sequence. Thus, for example, if $u_n = (-1)^n/n!$, then

$$f(u_n, x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} x^n = e^{-x},$$

which means that e^{-x} is the generating function of $\{u_n\}$.

The generating function of $\{u_{n+1}\}$ is

$$f(u_{n+1}, x) = \sum_{n=0}^{\infty} u_{n+1}x^n = \frac{1}{x}\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(n+1)!} x^{n+1}$$

$$= \frac{1}{x}\left(\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} x^n - 1\right) = \frac{1}{x}[f(u_n, x) - 1].$$

Consider the difference equation

$$u_{n+2} + u_{n+1} - 2u_n = 0, \qquad u_0 = 1, \quad u_1 = -2.$$

By taking the generating function of the equation, show that

$$f(u_n, x) = \frac{1}{1 + 2x}.$$

Using the binomial theorem find u_n.

38.15 A Fibonacci sequence is defined as a sequence in which any term is the sum of the two preceding terms. For the Fibonacci sequence starting with $u_1 = 1$, $u_2 = 2$, find and solve the difference equation for u_n.

38.16 Solve the initial-value difference equation

$$3u_{n+2} - 2u_{n+1} - u_n = 0, \qquad u_1 = 2, \quad u_2 = 1,$$

and show that $u_n \to \frac{5}{4}$ as $n \to \infty$.

38.17 A symmetric random walk takes place on the integer steps on the line between $x = 0$ and $x = N$. At any position $x = r$ $(1 \leqslant r \leqslant N - 1)$, the probability that the walker moves to either $x = r + 1$ or $x = r - 1$ at any stage is $\frac{1}{2}$. The probability u_k that the walker reaches $x = 0$ first, given an initial position $x = k$, satisfies the difference equation

$$u_k = \tfrac{1}{2}u_{k-1} + \tfrac{1}{2}u_{k+1}, \qquad u_0 = 1, \quad u_N = 0,$$

for $1 \leqslant k \leqslant N - 1$. Find u_k. What is the probability that the walker reaches $x = N$ first?

If d_k is the expected number of steps in the walk before it reaches 0 or N, then d_k satisfies

$$d_k = \tfrac{1}{2}(1 + d_{k+1}) + \tfrac{1}{2}(1 + d_{k-1}), \qquad d_0 = d_N = 0$$

for $1 \leqslant k \leqslant N - 1$. Find the expected duration of the walk.

38.18 Show that $u_n = n!$ is a solution of the second-order difference equation

$$u_{n+2} = (n + 2)(n + 1)u_n.$$

By using the substitution $u_n = v_n n!$, find a second independent solution.

38.19 Given that

$$s_n = \sum_{k=1}^{n} k^3,$$

find a first-order difference equation for s_n. Solve the equation to find a formula for the sum s_n.

38.20 Show that the difference equation

$$u_{n+2} + 2au_{n+1} + bu_n = 0$$

can be expressed as

$$z_{n+1} = Az_n,$$

where

$$z_n = \begin{bmatrix} u_n \\ v_n \end{bmatrix}, \qquad A = \begin{bmatrix} -2a & b \\ -1 & 0 \end{bmatrix}.$$

Deduce that

$$z_n = A^n z_0.$$

Consider the case with $a = 1$ and $b = -8$. Find the eigenvalues of A and use the methods of Section 13.5 to find a formula for A^n. Hence solve the difference equation for u_n in terms of u_0 and u_1.

38.21 (Section 38.5). Consider the logistic equation

$$u_{n+1} = \alpha u_n(1 - u_n).$$

Draw cobweb solutions starting at $u_0 = \frac{1}{2}$ for the cases $\alpha = 2.7$, $\alpha = 2.9$, and $\alpha = 3.3$. What do you infer about the stability of the fixed point in the first quadrant?

38.22 (Section 38.5). In the logistic equation $u_{n+1} = \alpha u_n(1 - u_n)$, for what positive values of α is the origin a stable fixed point?

38.23 (Section 38.5). Find the two stable values between which u_n ultimately oscillates in the logistic equation $u_{n+1} = 3.25u_n(1-u_n)$.

38.24 Consider the difference equation

$$u_{n+1} = \alpha(\tfrac{1}{2} - |u_n - \tfrac{1}{2}|).$$

Sketch the function $y = f(x) = \alpha(\frac{1}{2} - |x - \frac{1}{2}|)$ for $\alpha = \frac{3}{2}$. Where are the equilibrium points of the difference equation for $\alpha > 1$? Show that the origin is stable if $\alpha < 1$, and unstable if $\alpha > 1$. What happens if $\alpha = 1$?

Sketch the graph of $y = f(f(x))$ for $\alpha = 2$. Show that there exists a 2-cycle and locate the periodic values of u_n.

38.25 Find the fixed points of

$$u_{n+1} = \alpha u_n(1 - u_n^3),$$

for all α. Determine the slope of $y = f(x) = \alpha x(1 - x^3)$ at the nonzero fixed point. Confirm that this fixed point is stable if $\alpha < \frac{5}{3}$ and unstable if $\alpha > \frac{5}{3}$. Sketch cobweb solutions for $\alpha = 1.2, 1.4, 1.8$.

38.26 By starting from $u_0 = 0.957\,417$, compute $u_1, u_2, \ldots, u_5$ for the difference equation

$$u_{n+1} = \alpha u_n(1-u_n), \quad \alpha = 3.83,$$

and confirm that the logistic equation appears to have a 3-cycle for this value of α.

38.27 Find the fixed points of the difference equation

$$u_{n+1} = \alpha u_n(1-u_n)^2,$$

in the three cases (a) $\alpha = 9$, (b) $\alpha = 4$, (c) $\alpha = \frac{9}{4}$. Discuss the stability of the fixed points in each case.

38.28 Show that the special logistic equation

$$u_{n+1} = 4u_n(1-u_n)$$

has the solution

$$u_n = \sin^2(2^n C\pi)$$

where C is any constant. This general solution includes closed-form chaotic solutions. For example, if $C = 1/\pi$, then

$$u_n = \sin^2(2^n)$$

which never repeats itself for $n = 0, 1, 2, \ldots$.