Discrete Mathematics

Contents

Discrete Mathematics

Assessment statements

10.1 Strong induction.
Pigeon-hole principle.

10.2 $a|b \Rightarrow b = na$ for some $n \in \mathbb{Z}$.
The theorem $a|b$ and $a|c \Rightarrow a|(bx \pm cy)$ where $x, y \in \mathbb{Z}$.
Division and Euclidean algorithms.
The greatest common divisor, gcd(a, b), and the least common multiple, lcm(a, b), of integers a and b.
Prime numbers; relatively prime numbers and the fundamental theorem of arithmetic.

10.3 Linear Diophantine equations $ax + by = c$.

10.4 Modular arithmetic.
The solution of linear congruences.
Solution of simultaneous linear congruences (Chinese remainder theorem).

10.5 Representation of integers in different bases.

10.6 Fermat's little theorem.

10.7 Graphs, vertices, edges, faces. Adjacent vertices, adjacent edges.
Degree of a vertex, degree sequence.
Handshaking lemma.
Simple graphs; connected graphs; complete graphs; bipartite graphs; planar graphs; trees; weighted graphs, including tabular representation.
Subgraphs; complements of graphs.
Euler's relation: $v - e + f = 2$; theorems for planar graphs including $e \leqslant 3v - 6$, $e \leqslant 2v - 4$, leading to the results that K_5 and $K_{3,3}$ are not planar.

10.8 Walks, trails, paths, circuits, cycles.
Eulerian trails and circuits.
Hamiltonian paths and cycles.

10.9 Graph algorithms; Kruskal's; Dijkstra's.

10.10 Chinese postman problem.
Not required: Graphs with more than four vertices of odd degree.
Travelling salesman problem.
Nearest-neighbour algorithm for determining an upper bound.
Deleted vertex algorithm for determining a lower bound.

10.11 Recurrence relations. Initial conditions, recursive definition of a sequence.
Solution of first- and second-degree linear homogeneous recurrence relations with constant coefficients.
The first-degree linear recurrence relation $u_n = au_{n-1} + b$.
Modelling with recurrence relations.

1 Number Theory I

1.1 Introduction

This option deals with two 'relatively' separate topics: number theory and graph theory. The name Discrete Mathematics is actually not a well-defined subject in the mathematics community. In some cases it includes number theory and in some it does not. However, your syllabus contains ideas from both, and that is what we will focus on. A common thread between the two parts is the requirement for relatively 'rigorous' proofs. We will start with **number theory**.

Number theory

Elementary number theory deals with the study of **integers** *in general* and the **positive integers** 1, 2, 3, … *in particular*. The set of positive integers is denoted by $\mathbb{Z}^+$, and that of integers is denoted by $\mathbb{Z}$, where

$\mathbb{Z}^+ = \{1, 2, 3, \ldots\}$ (This is an IBO notation. In several mathematics sources, you will see that this set is called the set of natural numbers and is denoted by $\mathbb{N}$. Since you are preparing for an IB exam, we will follow this notation from this point onwards.)

$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$

Of course, the integers are familiar to you from your primary school. You have worked with them hundreds of times and have formed an intuitive sense of many of their laws. This intuition carries some danger with it. It becomes hard to see the necessity to prove laws that we have become used to. However, we will assume some of the axioms we considered earlier as 'obvious' and will use them in the rest of the course.

Properties/axioms

On the set of integers, we can define the operations of addition and multiplication. As usual, we denote the sum and product of a and b by $a + b$ and $a \cdot b$, respectively. Following convention, we will also write ab for $a \cdot b$. Important properties of integers with respect to these two operations are mentioned below.

Closure property of addition: If $a, b \in \mathbb{Z}$, then $a + b \in \mathbb{Z}$.

Closure property of multiplication: If $a, b \in \mathbb{Z}$, then $ab \in \mathbb{Z}$.

Commutative property of addition: If $a, b \in \mathbb{Z}$, then $a + b = b + a$ for all $a, b \in \mathbb{Z}$.

Commutative property of multiplication: If $a, b \in \mathbb{Z}$, then $ab = ba$ for all $a, b \in \mathbb{Z}$.

Associative property of addition: If $a, b, c \in \mathbb{Z}$, then $(a + b) + c = a + (b + c)$ for all $a, b, c \in \mathbb{Z}$.

Associative property of multiplication: If $a, b, c \in \mathbb{Z}$, then $(ab)c = a(bc)$ for all $a, b, c \in \mathbb{Z}$.

Distributive property of multiplication over addition: If $a, b, c \in \mathbb{Z}$, then $a(b + c) = ab + ac$ for all $a, b, c \in \mathbb{Z}$.

Additive identity property: For all $a \in \mathbb{Z}$, $a + 0 = a$.

Multiplicative identity property: For all $a \in \mathbb{Z}$, $a \cdot 1 = 1 \cdot a = a$.

Additive inverse property: For all $a \in \mathbb{Z}$, $a + (-a) = (-a) + a = 0$. Thus, $a + \left(-b\right)$ is written as $a - b$.

Cancellation property of multiplication: If $a, b, c \in \mathbb{Z}$, $a \neq 0$, then $ab = ac$ implies $b = c$.

These properties are also called axioms. An **axiom**, as you will recall, is a universally accepted principle, rule, or a proposition that is assumed without proof and serves as a starting point from which other statements are logically derived.

Here are some more properties, some of which can be proved by using the axioms mentioned before.

Cancellation property of addition: If $a, b, c \in \mathbb{Z}$, $a \neq 0$, then $a + b = a + c$ implies $b = c$.

This can be easily proved:

Given $a + b = a + c$, we add $-a$ to both sides. We get $(-a) + (a + b) = (-a) + (a + c)$.

By associative property of addition we get $((-a) + a) + b = ((-a) + a) + c$.

Now, by the additive inverse property, $0 + b = 0 + c$.

Using the additive identity property, we get $b = c$.

Ordering relation: On the system of integers $\mathbb{Z}$, there is an order relation 'less than', denoted by '$<$', on the basis of which we have the following law:

Law of trichotomy: If $a \in \mathbb{Z}$ then exactly one of the following statements is true:

(i) $a < 0$ (ii) $a = 0$ (iii) $a > 0$.

Properties of inequality:

(i) If $a, b, c \in \mathbb{Z}$, and $a < b$, then $a + c < b + c$.

(ii) If $a, b, c \in \mathbb{Z}$, $a < b$, and $c > 0$, then $ac < bc$.

(iii) If $a, b, c \in \mathbb{Z}$, $a < b$, and $c < 0$, then $ac > bc$.

The following is an important property of positive numbers:

Well-ordering property: Every non-empty set of positive integers contains a least element.

The well-ordering property is a fundamental axiom of the system of positive integers. We can quickly verify that this property is quite an obvious one if we consider a finite set of positive integers like the ones mentioned below:

1 $S_1 = \{2, 5, 7, 9, 14, 21\}$

2 $S_2 = \{4, 29, 17, 3, 101\}$

In **1**, the least element is 2, because it is smaller than every other element in S_1.

In **2**, the least element is 3.

In this publication, we expect that you are familiar with these properties of integers from your earlier work with numbers. What we have mentioned here are a set of axioms which describe the properties of integers. We have neither tried to make these axioms independent of each other nor to mention a minimal number of axioms to develop the system of integers.

Next, we will demonstrate a few proofs for you to refresh your knowledge and to get started with proving statements yourself. Recall that a rational number is expressed as a ratio of two integers. Real numbers that are not rational are irrational. The sets $\mathbb{Q}$ and $\mathbb{R}$ denote the set of all rational numbers and real numbers, respectively.

Proofs

Most statements you will prove in this option are **implications**, i.e. assertions of the form 'if P, then Q', where P and Q are themselves statements. P is called the **hypothesis** and Q is called the **conclusion**. This is also written as $P \Rightarrow Q$. An example is:

S: If I have a free moment, then I will call you.

Here P is the statement 'I have a free moment' and Q is the statement 'I will call you'.

The implication 'if P, then Q' is considered to be true unless P is true and Q is false. Thus, my statement is truthful in each of the following cases:

- I have a free moment and I call you.
- I do not have a free moment and I do not call you.
- I do not have a free moment, but I call you anyway!

I would lie only if I have a free moment and I don't call you.

The meaning of 'if P, then Q' is summarized in the truth table right (where T is for true, and F for false).

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Notice that 'if P, then Q' is not the same as 'if Q, then P'; and one of them could be true while the other false. The statement 'if Q, then P' is called the **converse** of 'if P, then Q'.

In the previous example, the converse would be 'If I call you, then I have a free moment'.

To disprove a statement we ordinarily use a counter example. For example, consider the statement:

If $a > b$, then $ac > bc$.

This could easily be disproved by letting $a = 1$, $b = -1$, and $c = -2$; obviously $a > b$, but $ac = -2 \ngtr bc = 2$.

In this option, you will be dealing mainly with two types of proofs:

- **Direct proof** is a proof in which logical arguments lead directly from the hypothesis to the conclusion. To prove $P \Rightarrow Q$ by direct proof, assume P holds, and show that Q must follow (see Example 1).
- **Indirect proof** is itself of two types: proof by **contradiction** and proof by **contrapostive**. In a proof by **contradiction**, we assume the statement is false and show that this leads to a contradiction, thereby showing that it is impossible for the statement to fail. A proof by **contrapositive** uses the fact that the implication $P \Rightarrow Q$ is logically equivalent to its contrapositive $\neg Q \Rightarrow \neg P$, thus proving the contrapositive will prove the statement itself (see Examples 2 and 3).

$\neg$ is a negation symbol.
'$\neg$' is read as 'not'.

Mathematical induction

In Section 4.7 of the textbook you worked extensively with one form of the principle of **mathematical induction** (MI). In this part of the option, we will prove the principle and introduce you to another form, which is called **strong mathematical induction**.

Proof of the mathematical induction principle (MI)

Statement 1: Let S be the set of positive integers such that

1 $1 \in S$

2 Whenever the integer $k \in S$, then $k + 1 \in S$.

Then S is $\mathbb{Z}^+$.

Proof

Let T be the set of integers not in S. Assume T to be non-empty. The well-ordering principle implies that T has a least element. Call the least element a. Since, by hypothesis, $1 \in S$, then $1 \notin T$ and hence $a > 1$.

Now, $a - 1 > 0$ and hence $a > a - 1 > 0$.

Since a is the smallest element in T, then $a - 1$ cannot be in T, and therefore $a - 1 \in S$. Now, if $a - 1 \in S$, by (**2**) above, $(a - 1) + 1 \in S$, i.e. $a \in S$. This contradicts the fact that $a \in T$. Therefore, we conclude that T must be empty and that S contains all positive integers.

Strong mathematical induction

A second version of the MI principle called '**second principle of MI**', or '**strong MI**', has the same structure except in the induction step:

Statement 2: Let S be the set of positive integers such that

1 $1 \in S$

2 Whenever the integers $1, 2, 3, \ldots, k \in S$, then $k + 1 \in S$.

Then S is $\mathbb{Z}^+$.

Proof

Let T be the set of integers not in S. Assume T to be non-empty. The well-ordering principle implies that T has a least element. Call the least element a. Since, by hypothesis, $1 \in S$, then $1 \notin T$ and hence $a > 1$.

Also, $1, 2, 3, \ldots, a - 1$ are all in S by hypothesis, and hence if $a - 1 \in S$, then by (**2**) above, $(a - 1) + 1 \in S$, i.e. $a \in S$. This contradicts the fact that $a \in T$. Therefore, we conclude that T must be empty and that S contains all positive integers.

Before we demonstrate how to use strong induction in specific examples, let us summarize the steps you need to follow:

To prove $S(n)$ true for all positive integers $n \geqslant n_0$, we complete the following two steps.

Basis Step: Verify that $S(n_0)$ is true. (In many cases $n_0 = 1$)

Inductive Step: Show that the implication $(S(n_0) \wedge S(n_0 + 1) \ldots \wedge S(k)) \rightarrow S(k + 1)$ is true for all positive integers k.

Conclude: $S(n)$ is true for all positive integers larger than or equal to n_0.

Example 1

For any integer $n \geqslant 2$, n is divisible by a prime number.

Proof

Basis step:
$S(2)$ is true, since 2 is divisible by 2 and 2 is a prime number.

Inductive step:
Assume the statement is true for all $n = i$ with $2 \leqslant i \leqslant k$, i.e. $S(2) \wedge \ldots \wedge S(k)$ is true. (This is called the inductive hypothesis.)

Show that it is true for $n = k + 1$.

We must show that $n = k + 1$ is divisible by a prime number.

We consider two cases:

(i) $k + 1$ is prime, and in this case is divisible by itself, or,

(ii) $k + 1$ is composite, and hence

$k + 1$ can be written as a product of two integers x and y such that $2 \leqslant x \leqslant k$ as well as $2 \leqslant y \leqslant k$. However, with the assumption that all numbers between 2 and k are divisible by a prime, then x and y are divisible by a prime and hence by transitive property, $k + 1$ is also divisible by a prime.

Therefore $S(n)$ is true for all positive integers by the principle of strong induction.

Example 2

A sequence $\{a_n\}$ is defined by

$$\begin{cases} a_0 = 1, a_1 = 2, a_2 = 3 \\ a_n = a_{n-1} + a_{n-2} + a_{n-3}\ \forall n \in \mathbb{Z}, n \geqslant 3 \end{cases}$$

Show that $S(n)$: $a_n \leqslant 2^n$ for all non-negative integers.

Proof

Basis step:
$S(0)$ is true since $a_0 = 1 \leqslant 1 = 2^0$, $S(1)$ is true since $a_1 = 2 \leqslant 2 = 2^1$, and $S(2)$ is true since $a_2 = 3 \leqslant 4 = 2^2$.

Inductive step:
Assume the statement is true for all $n = i$ with $0 \leqslant i \leqslant k$, i.e. $S(0) \wedge \ldots \wedge S(k)$ is true, i.e., $a_0 \leqslant 2^0, \ldots, a_k \leqslant 2^k$.

Show that it is true for $n = k + 1$.

We must show that $a_{k+1} \leqslant 2^{k+1}$.

$a_{k+1} = a_k + a_{k-1} + a_{k-2} \leqslant 2^k + 2^{k-1} + 2^{k-2}$ which is based on the assumption above.

This leads to $a_{k+1} \leqslant 2^k + 2^{k-1} + 2^{k-2} \leqslant 2^k + 2^{k-1} + 2^{k-2} + 2^{k-3} + \ldots + 1$

But $2^k + 2^{k-1} + 2^{k-2} + 2^{k-3} + \ldots + 1 = \dfrac{1 - 2^{k+1}}{1 - 2} = 2^{k+1} - 1$ since it is a geometric series with $k + 1$ terms, first term equal to 1 and a common ratio of 2.

Hence, $a_{k+1} \leqslant 2^{k+1}$ and therefore $S(n)$ is true for all positive integers by the principle of strong induction.

Example 3

Fibonacci sequences are defined recursively by

$$\begin{cases} u_1 = 1, u_2 = 1 \\ u_n = u_{n-1} + u_{n-2}, n > 2. \end{cases}$$

Show that the closed form for the nth term of Fibonacci sequence is given by

$$u_n = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n\sqrt{5}} \text{ for } n > 2.$$

Proof

Basis step:

$S(1)$ is true, since $u_1 = \frac{(1+\sqrt{5})^1 - (1-\sqrt{5})^1}{2\sqrt{5}} = \frac{2\sqrt{5}}{2\sqrt{5}} = 1.$

$S(2)$ is true, since $u_2 = \frac{(1+\sqrt{5})^2 - (1-\sqrt{5})^2}{2^2\sqrt{5}} = \frac{1+2\sqrt{5}+5-1+2\sqrt{5}-5}{4\sqrt{5}}$

$$= \frac{4\sqrt{5}}{4\sqrt{5}} = 1.$$

Inductive step:

Assume the statement is true for all $n = i$ with $1 \leqslant i \leqslant k$, i.e. $S(1) \wedge \ldots \wedge S(k)$ is true.

Show that it is true for $n = k + 1$.

We must show that $u_{k+1} = \frac{(1+\sqrt{5})^{k+1} - (1-\sqrt{5})^{k+1}}{2^{k+1}\sqrt{5}}$

We know that $u_{k+1} = u_k + u_{k-1}$ by definition of Fibonacci sequence.

By assumption, we know that $u_k = \frac{(1+\sqrt{5})^k - (1-\sqrt{5})^k}{2^k\sqrt{5}}$ and

$$u_{k-1} = \frac{(1+\sqrt{5})^{k-1} - (1-\sqrt{5})^{k-1}}{2^{k-1}\sqrt{5}}.$$

$$\begin{aligned}
\text{Hence, } u_{k+1} &= \frac{(1+\sqrt{5})^k - (1-\sqrt{5})^k}{2^k\sqrt{5}} + \frac{(1+\sqrt{5})^{k-1} - (1-\sqrt{5})^{k-1}}{2^{k-1}\sqrt{5}} \\
&= \frac{(1+\sqrt{5})^k - (1-\sqrt{5})^k + 2(1+\sqrt{5})^{k-1} - 2(1-\sqrt{5})^{k-1}}{2^k\sqrt{5}} \\
&= \frac{\left((1+\sqrt{5})^k + 2(1+\sqrt{5})^{k-1}\right) - \left((1-\sqrt{5})^k + 2(1-\sqrt{5})^{k-1}\right)}{2^k\sqrt{5}} \\
&= \frac{(1+\sqrt{5})^k\left(1+\frac{2}{1+\sqrt{5}}\right) - (1-\sqrt{5})^k\left(1+\frac{2}{1-\sqrt{5}}\right)}{2^k\sqrt{5}}
\end{aligned}$$

By more algebraic manipulation we have:

$$u_{k+1} = \frac{(1+\sqrt{5})^k\left(1+\frac{2}{1+\sqrt{5}}\right)-(1-\sqrt{5})^k\left(1+\frac{2}{1-\sqrt{5}}\right)}{2^k\sqrt{5}}$$

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

Therefore, by the principle of strong induction, the closed form for the nth term of Fibonacci sequence is given by $u_n = \dfrac{(1+\sqrt{5})^n-(1-\sqrt{5})^n}{2^n\sqrt{5}}, n > 2.$

Example 4

Fibonacci sequences are defined recursively as in example 3.

Prove that $\sum_{i=1}^{n} u_i = u_{n+2} - 1$ for every $n \in \mathbb{Z}^+$.

Proof

Basis step:

$S(1)$ is true since for $n = 1$

$$\sum_{i=1}^{1} u_i = u_1 = 1 = u_3 - 1 = 2 - 1, \text{ which is true.}$$

As a check, we also know that $S(2)$ is true since

$$\sum_{i=1}^{2} u_i = u_1 + u_2 = 1 + 1 = u_4 - 1 = 3 - 1$$

Inductive step:

Assume the statement is true for $n = k$, show that it is true for $n = k + 1$.

We must show that $\sum_{i=1}^{k+1} u_i = u_{k+3} - 1$

$$\sum_{i=1}^{k+1} u_i = \sum_{i=1}^{k} u_i + u_{k+1} = u_{k+2} - 1 + u_{k+1}$$
$$= u_{k+3} - 1$$

Therefore, by the principle of mathematical induction, the statement is true for all positive integers.

Note: We could have used strong induction here in the following manner.

Assume the statement is true for all $n = i$ with $1 \leqslant i \leqslant k$, i.e. $S(1) \wedge \dots \wedge S(k)$ is true, i.e.,

Show that it is true for $n = k + 1$

$$\sum_{i=1}^{k+1} u_i = \sum_{i=1}^{k-1} u_i + u_k + u_{k+1} = u_{k+1} - 1 + u_{k+2}$$
$$= u_{k+3} - 1$$

We suggest that you use the following format for proofs by strong (or mathematical) induction.

- Say what you are proving.
- Say that the proof is by strong mathematical induction, and make it clear what is playing the role of *n*.
- In the induction case, state the induction hypothesis (IH) and what you need to show (NTS).
- Divide the argument into cases, as needed.
- Indicate clearly where and how you use the inductive hypothesis.

Other methods of proofs – examples

Example 5

Prove that the product of two odd integers is an odd integer.

Proof

Given that a and b are odd integers, we need to prove that ab is an odd integer.

Let a and b be odd integers. Then you can find two integers m and n such that

$a = 2m + 1$ and $b = 2n + 1$.

The product ab is

$$\begin{aligned} ab &= (2m+\ 1)(2n+\ 1) = 4mn + 2m + 2n + 1 \\ &= 2(2mn + m + n) + 1 = 2k + 1 \end{aligned}$$

where $k = 2mn + m + n \in \mathbb{Z}$. Thus ab is odd.

Example 6

Prove that $\sqrt{2}$ is an irrational number.

Proof

Assume that $\sqrt{2}$ is rational. Then, by definition of rational numbers, $\sqrt{2}$ can be written as a **reduced** fraction $\frac{m}{n}$ where the two integers m and n, with $n \neq 0$, have no common divisor except 1.

$$\sqrt{2} = \frac{m}{n} \Rightarrow \left(\sqrt{2}\right)^2 = \frac{m^2}{n^2} \Rightarrow m^2 = 2n^2$$

This tells us that m^2 is even.

If m^2 is even, m must also be even (this is assumed true, but can be proved true easily), thus $m = 2k$ for some integer k.

This leads us to $m^2 = 4k^2$, and hence

$2n^2 = m^2 = 4k^2 \Rightarrow n^2 = 2k^2$, and thus n^2 is even, which in turn leads to n being even.

Thus both m and n are even, and hence they have another common factor, 2, which contradicts the definition of a rational number.

Therefore, assuming $\sqrt{2}$ to be rational leads us to a contradiction and so $\sqrt{2}$ cannot be rational.

Example 6 is a proof by contradiction. The next example will demonstrate the use of contrapostive in the proof.

Example 7

Let a be a positive real number. Prove that if a is an irrational number then $\sqrt{a}$ is also irrational.

Proof

Stated differently, we need to prove: $a \notin \mathbb{Q} \Rightarrow \sqrt{a} \notin \mathbb{Q}$.

We will use the contrapositive and attempt to prove $\sqrt{a} \in \mathbb{Q} \Rightarrow a \in \mathbb{Q}$.

Suppose $\sqrt{a} \in \mathbb{Q}$, then there are two integers m and n, with $n \neq 0$, such that $\sqrt{a} = \frac{m}{n}$ by definition of rational numbers. Thus $a = \left(\sqrt{a}\right)^2 = \frac{m^2}{n^2}$, and since m and n, with $n \neq 0$, are integers, then m^2 and n^2, with $n^2 \neq 0$, are also integers.

So, a can be written as the quotient of two integers, and hence it is a rational number, by definition.

By proving the contrapositive, the statement itself is true.

Note: There is a convention that is well known in mathematics and that is the use of the 'iff'. This word stands for 'if and only if,' which in turn means a logical equivalence. That is, if we say P iff Q, we mean P implies Q and Q implies P. Hence, in some proofs, we will have to prove both statements. In this publication, we will indicate the two-way process by using ($\Rightarrow$) for the first and ($\Leftarrow$) for the second.

Pigeonhole principle

As the name indicates, the idea stems from the following situation:

A flock of pigeons flies into a set of pigeonholes. If there are more pigeons than pigeonholes, then there must be at least one pigeonhole with more than one pigeon (at least two pigeons).

Theorem: The pigeonhole principle

If $n + 1$ objects or more are placed into n positions, then there is at least one position that contains at least two of the objects.

Proof

Assume that no position has more than one object. Then there will be at most n objects. This is a contradiction since there are $n + 1$ or more objects.

Note: The pigeonhole principle is sometimes called the **Dirichlet drawer principle**, after the German mathematician Dirichlet.

Example 8

What is the minimum number of people in a room where at least two of them have the same birth month?

Solution

There should be at least 13 as there are only 12 possible months.

Example 9

True or false: In a HL IB class of 10 students, there will be at least two students with the same score.

Solution

True, since there are only seven grades possible in the mathematics examination.

Example 10

True or false: In a 5-digit number code situation given to a group larger than 10, there will be at least two codes that start with the same digit, end with the same digit, etc.

Solution

True, since there are only 10 digits possible!

1.2 Division algorithm

The sum, difference, and product of two integers is always an integer, but the quotient may not be. The concept of divisibility of one integer by another is central in number theory. We are not only interested to know the underlying reason for an integer to be divisible by another integer, but also interested to see how this concept is applied in different situations.

If a is a divisor of b so is $-a$, since $b = ac$ implies $b = (-a)(-c)$. So, the divisors of an integer at all times happen in pairs. To obtain all the divisors of a given integer, it is enough to get the positive divisors and then tag on to them the matching negative integers. In this book, we will usually limit our listing of divisors to the positive ones.

Definition 1

If a and b are integers with $a = 0$, then b is divisible by a if there exists an integer c such that $b = ac$.

In this case we say a divides b and denote this by $a|b$. a is called a divisor or factor of b and b is called a **dividend** or a **multiple** of a. If a does not divide b then we write $a \nmid b$.

Example

The following statements illustrate the concept of divisibility of integers:

$11 | 143$, $-4 | 28$, $19 | 133$, $5 | 0$, $3 \nmid 2$, and $15 \nmid 47$.

Example

The divisors of 8 are ± 1, ± 2, ± 4, and ± 8. The divisors of 11 are ± 1 and ± 11.

In subsequent sections, we will need some simple properties of divisibility which we now state and prove as theorems.

Theorem 1

If a, b, and c are integers with $a | b$ and $b | c$, then $a | c$.

Proof

Since $a | b$ and $b | c$, there exist integers m and n such that $b = am$ and $c = bn$. Hence, $c = (am)n = (mn)a$. Now, since mn is an integer, then, by definition, this shows that $a | c$.

Example

$3 | 6$ and $6 | 216$, then $3 | 216$; $5 | 15$ and $15 | 3375$, then $5 | 3375$; $11 | 44$ and $44 | 308$, then $11 | 308$.

Theorem 2

If $a | b$ and $a | c$, then $a | (b \pm c)$.

Proof

Since $a | b$ and $a | c$, then there exist integers m and n such that $b = ma$ and $c = na$.

Hence, $b \pm c = ma \pm na = (m \pm n)a$.

Now, since $m \pm n$ is an integer, $a | (b \pm c)$.

Corollary 1

If $a | b$ and $a | c$, then $a | (bx \pm cy)$, where a, b, x, and y are integers.

The corollary follows from Theorem 2 by recognizing that bx and cy are integers and can be substituted for b and c in the theorem.

This is to say that if a divides b and c then a divides any integer linear combination of b and c.

This property can be extended to sums of more than two integers. That is, if $a \mid b_j$ for $j = 1, 2, \ldots, n$, then

$a \mid (b_1x_1 + b_2x_2 + \ldots + b_nx_n)$ for all integers $x_1, x_2, \ldots, x_n$.

Example

$5 \mid 45$ and $5 \mid 60$, then $5 \mid (45 + 60)$, i.e. $5 \mid 105$; $5 \mid (7 \cdot 45 - 2 \cdot 60)$, i.e. $5 \mid 195$.

Theorem 3

If $a, b, c \in \mathbb{Z}$, then the following hold:

(i) $a \mid 0$, $1 \mid a$, and $a \mid a$.

(ii) $a \mid 1$ if and only if $a = \pm 1$.

(iii) If $a \mid b$, and $c \mid d$, then $ac \mid bd$.

(iv) $a \mid b$, and $b \mid a$, if and only if $a = \pm b$.

(v) If $a \mid b$, and $b \neq 0$, then $|a| \leqslant |b|$.

Proof

We will leave the proofs of parts (i)–(iv) as an exercise, and only prove (v) here.

If $a \mid b$, then there exists an integer c such that $b = ac$; moreover, $b \neq 0$ means that $c \neq 0$. Now, taking absolute values,

$$|b| = |ac| = |a|\,|c|.$$

Since $c \neq 0$, then $|c| \geqslant 1$, and therefore

$$|b| = |a|\,|c| \geqslant |a|.$$

Theorem 4: The division algorithm

If a and b are integers such that $b > 0$, then there exist unique integers q and r such that $a = bq + r$, with $0 \leqslant r < b$.

Note: We call q the **quotient** and r the **remainder**; we also call a the **dividend** and b the **divisor**.

Note that a is divisible by b if and only if the remainder in the division algorithm is zero.

Before we prove the division algorithm, let us consider some examples.

Example

If $a = 183$ and $b = 31$, then $q = 5$ and $r = 28$, since $183 = 31 \cdot 5 + 28$.

Also, $a = -183$ and $b = 31$, then $q = -6$ and $r = 3$, since $-183 = 31(-6) + 3$.

Note: It is natural for us to ask, given two numbers a and b, how can we find the quotient q and the remainder r mentioned in the division algorithm? As an illustration, let us take $a = 94$ and $b = 13$. In order to find the quotient q, multiply 13 successively by $\{1, 2, 3, \ldots\}$ until you reach a number larger than or equal to 91.

$13 \cdot 1 = 13, 13 \cdot 2 = 26, 13 \cdot 3 = 39, \ldots, 13 \cdot 7 = 91, 13 \cdot 8 = 104$

So, $q = 7$, and the remainder $r = 94 - 13 \cdot 7 = 3$.

This process is a result of the division algorithm itself:

$a = bq + r$, with $0 \leqslant r \leqslant b \Leftrightarrow \frac{a}{b} = q + \frac{r}{b}$, with $0 \leqslant \frac{r}{b} \leqslant 1$. This in turn can be interpreted as follows:

q is the integer part of the quotient of a by b, and $\frac{r}{b}$ is the decimal part, and hence q is nothing but the greatest integer function of $\frac{a}{b}$. So,

$q = \left[\frac{94}{13}\right] = [7.23] = 7$ and $r = 94 - 13(7) = 3$.

For instance, in the example above, we have

$q = \left[\frac{183}{31}\right] = [5.9] = 5$, and hence the remainder r is $183 - 31 \cdot 5 = 28$; also

$q = \left[\frac{-183}{31}\right] = [-5.9] = -6$, and $r = -183 - 31(-6) = 3$.

Example

$a = 121$ and $b = 9$, then $q = \left[\frac{121}{9}\right] = [13.4] = 13$, and $r = 121 - 13 \cdot 9 = 4$, and so $121 = 9 \cdot 13 + 4$. Also, if $a = -148$ and $b = 12$, then

$q = \left[\frac{-148}{12}\right] = [-12.3] = -13$, and $r = -148 - 12(-13) = 8$, and so

$-148 = 12(-13) + 8$.

We now present a proof of the division algorithm.

Proof of the division algorithm

This is an existence and uniqueness proof. First we have to prove that q and r exist, and then, if they exist, they are the only numbers that satisfy the division algorithm.

Existence:

Suppose the real number a/b is $q + k$, where q is an integer and $0 \leqslant k < 1$. Then

$$a = b(q + k) = bq + bk.$$

Now, since a is an integer and bq is an integer (product of two integers), it follows that bk is also an integer. Moreover, since $b > 0$, multiplying it with all sides of $0 \leqslant k < 1$ gives us $0 \leqslant bk < b$. With this in mind, we set $r = bk$, and thus we have

$$a = bq + r \text{ with } 0 \leqslant r < b.$$

Uniqueness:

Next we show that q and r are unique. Using an indirect proof, suppose they are not unique, then there exists at least another pair q_1 and r_1 that satisfy the division algorithm, and now we have

$$a = bq + r \text{ with } 0 \leqslant r < b, \text{ and}$$

$$a = bq_1 + r_1 \text{ with } 0 \leqslant r_1 < b.$$

Subtract the two equations and simplify:

$$r - r_1 = b(q - q_1) \text{ ....(1)}$$

Add the two inequalities $0 \leqslant r < b$ and $-b \leqslant -r_1 < 0$, and thus

$-b < r - r_1 < b.$

Divide all sides by b and we have

$-1 < \dfrac{r - r_1}{b} < 1.$

Since $\dfrac{r - r_1}{b} = q - q_1$ from equation (1), and since

$q - q_1$ is an integer, and the only integer between -1 and $+1$ is zero, then

$q - q_1 = 0$, which implies that $q = q_1$. Also, $\dfrac{r - r_1}{b} = 0 \Rightarrow r - r_1 = 0 \Rightarrow r = r_1$.

Therefore, q and r are unique.

Note: The result we established can also be applied when $b < 0$. For if $b < 0$, then $-b > 0$, and hence we can say that according to Theorem 4, there exist two integers q_1 and r such that

$a = (-b)q_1 + r$ with $0 \leqslant r < -b$, which can be rewritten as $a = b(-q_1) + r$ with $0 \leqslant r < -b$. Now take

$q = -q_1$, and we get $a = bq + r$ with $0 \leqslant r < -b$ and $q \in \mathbb{Z}$. This is the existence part of the theorem. Uniqueness follows the same approach as in the main theorem.

Combining this observation with the statement from Theorem 4, we obtain:

Corollary 2

If a and b are integers and $b \neq 0$, then there are unique integers q and r such that $a = qb + r$ with $0 \leqslant r < |b|$.

Example

a) Let $a = 51$ and $b = -9$, then $51 = (-9)(-5) + 6$. Here too, we can use the largest integer function in the following manner:

$$q = \left[\frac{a}{-b}\right] = \left[\frac{51}{9}\right] = [5.67] = 5, \text{ and } r = a - 9 \cdot 5 = 6.$$

b) Let $a = -51$ and $b = -9$, then $-51 = (-9)(6) + 3$. Here too, we can use the largest integer function:

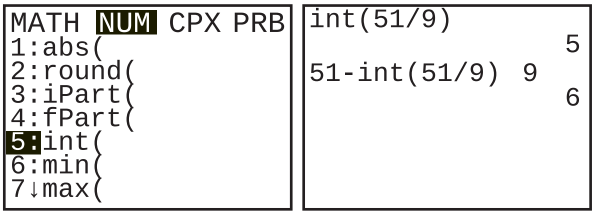
$$q = \left[\frac{a}{-b}\right] = \left[\frac{-51}{9}\right] = [-5.67] = -6, \text{ and } r = a - 9 \cdot (-6) = 3.$$

Division algorithm with a GDC

The calculation we made above can also be performed with your GDC. Here are the solutions, i.e. q and r for the previous examples.

First, you go to the MATH menu, then to the ‘NUM’ submenu, then to the ‘int(’ function, which is the greatest integer function.

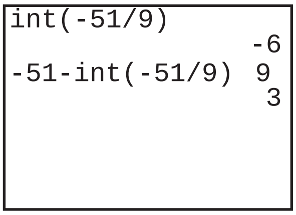
a)



```
MATH NUM CPX PRB
1:abs(
2:round(
3:iPart(
4:fPart(
5:int(
6:min(
7↓max(
```

```
int(51/9)
                5
51-int(51/9)  9
                6
```

b)



```
int(-51/9)
                   -6
-51-int(-51/9)  9
                    3
```

Example 11

Prove that if $a \in \mathbb{Z}$, then a^2 leaves a remainder of 0 or 1 when divided by 4.

Solution

By the division algorithm, $a = 4q + r$, where $0 \leqslant r < 4$. Thus,

$$a^2 = (4q + r)^2 = 16q^2 + 8qr + r^2.$$

Now the possible values of r are 0, 1, 2, or 3.

If $r = 0$, then $a^2 = 16q^2$, which is divisible by 4, so $r = 0$.

If $r = 1$, then $a^2 = 16q^2 + 8q + 1 = 4(4q^2 + 2q) + 1$, so $r = 1$.

If $r = 2$, then $a^2 = 16q^2 + 16q + 2^2 = 4(4q^2 + 4q + 1)$, which is divisible by 4, so $r = 0$.

If $r = 3$, then $a^2 = 16q^2 + 24q + 9 = 4(4q^2 + 6q + 2) + 1$, so $r = 1$.

Therefore, in all cases, $r = 0$ or 1.

Example 12

Show that the square of an odd integer is of the form $8k + 1$ for some integer k.

Solution

By the division algorithm, any integer is of the form $4q$, $4q + 1$, $4q + 2$, or $4q + 3$.

Hence, an odd integer can be of the form $4q + 1$ or $4q + 3$. When we square, we get

$$(4q+1)^2 = 16q^2 + 8q + 1 = 8(2q^2 + q) + 1 = 8k + 1, \text{ where } k = 2q^2 + q.$$

If the odd integer is of the form $4q + 3$, we have

$$\begin{aligned}(4q+3)^2 &= 16q^2 + 24q + 9\\ &= 8(2q^2 + 3q + 1) + 1 = 8k + 1, \text{ where } k = 2q^2 + 3q + 1.\end{aligned}$$

Example 13

Show that for all integers $a \geqslant 1$, $\dfrac{a(a^2+2)}{3}$ is an integer.

Solution

By the division algorithm, a is of the form $3q$, $3q + 1$ or $3q + 2$ for $q \in \mathbb{Z}$.

If $a = 3q$, then $\dfrac{a(a^2+2)}{3} = q(9q^2+2) \in \mathbb{Z}$.

If $a = 3q + 1$, then

$$\frac{a(a^2+2)}{3} = \frac{(3q+1)(9q^2+6q+3)}{3} = (3q+1)(3q^2+2q+1) \in \mathbb{Z}.$$

If $a = 3q + 2$, then

$$\frac{a(a^2+2)}{3} = \frac{(3q+2)(9q^2+12q+6)}{3} = (3q+2)(3q^2+4q+2) \in \mathbb{Z}.$$

Combining all three possibilities gives $\dfrac{a(a^2+2)}{3} \in \mathbb{Z}$ for $a \geqslant 1$.

Exercise 1.1–1.2

1 Find $a > 0$ where $a \mid 18$, $a \nmid 12$, and $\frac{36}{a} \nmid 10$.

2 Find $a > 0$ where $a \nmid 1000$, $5 \mid a$, $a \mid 60$, and $\frac{a}{2} \mid 75$.

3 Prove: If $m \neq 0$, then $a \mid b$ if and only if $ma \mid mb$.

4 Prove: $a \mid b$ and $b \mid a$ if and only if $a = \pm b$.

5 Prove: If $d \mid a$ and $a \neq 0$, then $|d| \leqslant |a|$.

6 Prove: If $c \mid a$ and $c \mid b$, then $c \mid (au + bv)$ for all $u, v \in \mathbb{Z}$.

7 Find the unique quotient and remainder when

- **a** 1028 is divided by 34
- **b** -380 is divided by 75
- **c** 180 is divided by -31.

8 Show that the sum of an even integer and an odd integer is odd.

9 Show that the sum of two even integers or two odd integers is even.

10 Show that if a and b are odd integers and $b \nmid a$, then there exists k and l such that $a = bk + l$, where l is odd and $|l| < b$.

11 Show that if a, b, and c are integers with $b > 0$ and $c > 0$, such that when a is divided by b the quotient is q and the remainder is r, and when q is divided by c the quotient is u and the remainder is v, then when a is divided by bc, the quotient is u and the remainder is $bv + r$.

12 Show that if a and b are integers, then there are integers q, r, and $s = \pm 1$ such that $a = bq + sr$, where $\frac{-b}{2} < r \leqslant \frac{b}{2}$.

13 Prove that if u and v are integers with $v > 0$, then there exist unique integers s and t such that $u = sv + t$, where $2v \leqslant t < 3v$.

14 Use the division algorithm to prove that the cube of any integer has one of the following forms: $9k$, $9k + 1$, $9k + 8$ for some $k \in \mathbb{Z}$.

15 Use the division algorithm to prove that the fourth power of any integer is either of the form $5k$ or $5k + 1$ for $k \in \mathbb{Z}$.

16 Let a and b be non-zero integers.

- **a** Prove that there exists unique integers q and r such that $a = bq + r$ with $\frac{-|b|}{2} < r \leqslant \frac{|b|}{2}$.
- **b** Find the unique q and r given in **a** for $a = 49$ and $b = -6$.

17 For all odd integers m and n, if $mn = 4k + 1$, then m or n is of the form $4j - 1$.

18 Prove parts (i)–(iv) of Theorem 3.

19 Find positive integers x and y such that $x|y$ and $x2^x|y^2$, but $2^x > y$.

20 Find positive integers x and y such that $x|y$ and $2^x \leqslant y$, but $x2^x \nmid y^2$.

21 Find positive integers x and y such that $x2^x|y^2$ and $2^x \leqslant y$, but $x \nmid y$.

22 Prove that if $a|b$, and $b|c$, then $a|(ax + by + cz)$ for all $x, y, z \in \mathbb{Z}$.

In questions 23–29, prove each statement if it is true, or show that it is false either by reasoning or by finding a counter example.

23 For all integers a and b, $a + b$ is odd if and only if (iff) one of the numbers is odd and the other is even.

24 For all integers a and b, ab is even iff at least one of the numbers is even.

25 For all integers a and b, $a^3 - b^3$ is even iff $a - b$ is even.

26 For all integers n, $n^2 + n + 3$ is odd.

27 For all integers a, b, and c, $a|(b + c)$ iff $a|b$ and $a|c$.

28 For all integers a, b, and c, $a|(bc)$ iff $a|b$ and $a|c$.

29 For all integers a and b, $a^2|b^2$ iff $a|b$.

30 **a** If a group of eight students are chosen, what is the probability that two of them will be born on the same day of the week?

b Show that if any 11 numbers are chosen from the set of numbers $\{1, 2, 3, \ldots, 20\}$, then one of them will be a multiple of another.

c Show that if any five points are chosen on or inside an equilateral triangle with side 1 cm, then two of them must be no more than 0.5 cm apart.

d Show that if any of seven points are chosen inside a hexagon with 1 cm sides, then two of them must be no more than 1 cm apart.

31 If Fibonacci numbers are denoted by F_n, and the golden ratio by $\varphi = \frac{1+\sqrt{5}}{2}$, prove that $\varphi^n = F_n\varphi + F_{n-1}$

32 Prove that $4 \mid 3^{2n-1} + 1$ for any integer $n \geq 1$.

33 Prove that $\sum_{i=1}^{n} \frac{1}{\sqrt{i}} \geqslant \sqrt{n}, n \geqslant 1.$

34 Show that for all $n \in \mathbb{N}$, $n(n^2 + 5)$ is a multiple of 6.

1.3 Greatest common divisor/ Euclidean algorithm

If a, b, and c are integers and $c \neq 0$, then c is called a **common divisor** of a and b if $c \mid a$ and $c \mid b$. (In some cases, it is called a divisor of a and b.)

Let S be the set of all common divisors of a and b. S is a non-empty set, because ± 1 belong to the set.

If a and b are both non-zero, then the number of divisors of a and b is finite.

Hence, it makes sense to speak of the largest member of the set S.

Definition 2

If a and b are integers with at least one of them different from zero, then we define the **greatest common divisor** of a and b, denoted by gcd(a, b), as the largest positive integer which divides a and b.

Stated differently, the gcd(a, b) is a number d that satisfies the two conditions:

1 $d \mid a$ and $d \mid b$.

2 If c is a divisor of a and b, then $c \leqslant d$.

Example

- gcd(30, 80) = 10. The positive divisors of 30 are: 1, 2, 3, 5, 6, 10, 15, 30. The divisors of 80 are: 1, 2, 4, 5, 8, 10, 20, 40, 80. Divisors of 30 and 80 are {1, 2, 5, 10}, and thus gcd(30, 80) = 10. Notice that any other divisor must be less than 10.
- gcd(−30, 80) = 10
- gcd(−30, 60) = 30
- gcd(60, −75) = 15
- gcd(25, 14) = 1
- gcd(0, 23) = 23

In defining the gcd, we can go as far as saying $d \mid |a|$ and $d \mid |b|$, i.e. in finding the gcd, we can ignore the sign!

The next theorem indicates that gcd(a, b) can be represented as a linear combination of a and b. That is, we can find two integers, x and y, such that gcd(a, b) $= ax + by$.

For example, gcd(−24, 60) = 12 implies that we can find two numbers x and y such that $12 = -24x + 60y$, and indeed $12 = -24 \cdot 2 + 60 \cdot 1$.

Theorem 5

If a and b are integers which are not both zero, then the greatest common divisor, $\gcd(a, b)$, of a and b is the smallest positive integer such that

$$\gcd(a, b) = ax + by$$

for $x, y \in \mathbb{Z}$.

Proof

Let S be the set of all positive integers of the form $ax + by$:
$S = \{ax + by \mid ax + by > 0;\ x, y \in \mathbb{Z}\}$.

S is non-empty, since $aa + bb = a^2 + b^2 > 0$. Hence, there is a smallest positive integer g such that

$$g = ax_1 + by_1$$

(by the well-ordering principle).

If either a or b is zero, the proof that $\gcd(a, b) = g$ is simple. For example, if $a = 0$, then $g = 0 + by_1 = b$ by taking $y_1 = 1$, and since $\gcd(0, b) = b$, thus $\gcd(a, b) = g$.

Assume that $a \neq 0$ and $b \neq 0$.

By the division algorithm,

$$a = gq + r \text{ with } 0 \leqslant r < g$$

and so

$$r = a - gq.$$

Hence,

$$r = a - (ax_1 + by_1)q = a(1 - x_1) + b(-qy_1).$$

Since $1 - x_1$ is an integer and $-qy_1$ is also an integer then r is of the form $ax + by$, which qualifies it to be a member of S. But r cannot be a member of S since $r < g$ and g is the smallest element in S, and therefore r must be zero.

This implies that $r = a - gq = 0$, and thus $a = gq$, or equivalently $g \mid a$. In a similar manner, we can show that $g \mid b$. Hence, g is a common divisor of a and b.

Let g_1 be any other common divisor of a and b, then Corollary 1 of Theorem 2 allows us to conclude

$$g_1 \mid (ax + by).$$

That is, $g_1 \mid g$, and by Theorem 3, part (v),

$$g_1 = |g_1| \leqslant |g| = g.$$

Thus, g is greater than any common divisor of a and b.

Finally, we can now claim that $g = \gcd(a, b)$.

The preceding theorem proved that the gcd exists and that it can be written as a linear combination of a and b. The theorem did not attempt to prove that g as found is unique. Below is a theorem that proves uniqueness.

Theorem 6

The greatest common divisor of two integers which are not both zero is unique.

Proof

Assume that g is not unique, then there is at least another integer g_1 that is also a gcd for a and b.

If g is the gcd, then any common divisor of a and b is a divisor of g, and hence $g_1 | g$, similarly $g | g_1$, and therefore $g_1 = g$.

Example

Let $a = 12$ and $b = 18$. Set S as described in the proof of Theorem 5 is

$$\begin{aligned} S &= \{ax + by | ax + by > 0; x, y \in \mathbb{Z}\} \\ &= \{12x + 18y\} \\ &= \{12(4) + 18(-2), 12(4) + 18(-1), 12(5) + 18(-3), \ldots\} \\ &= \{12, 30, 6, \ldots\}. \end{aligned}$$

The smallest element in this set is 6, which is the gcd of 12 and 18.

Now we know that $\gcd(a, b)$ is unique, and we know too that it is the smallest integer in the form $ax + by$. We have to decide how to efficiently calculate the $\gcd(a, b)$.

Theorem 7

If $a = bq + r$, then $\gcd(a, b) = \gcd(b, r)$.

Proof

Any common divisor of b and r also divides $bq + r = a$. Similarly, $r = a - bq$ implies that any common divisor of a and b also divides r. Thus, the two pairs of integers (a, b) and (b, r) have the same common divisors. So, they have the same greatest common divisor.

Example

- Let $a = 748$ and $b = 143$.

  We can write $748 = 143 \cdot 5 + 33$.

Now $\gcd(748, 143) = 11$, and $\gcd(143, 33) = 11$.

- Let $a = 954$ and $b = 216$.

  $954 = 216 \cdot 4 + 90$

  $\gcd(954, 216) = 18$, and $\gcd(216, 90) = 18$.

The Euclidean algorithm

Let a and b be two integers not both zero. Since $\gcd(|a|, |b|) = \gcd(a, b)$ there is no harm in assuming $a \geqslant b > 0$. By the division algorithm,

$$a = bq_1 + r_1, \text{ where } 0 \leqslant r_1 < b.$$

If $r_1 = 0$, then $b | a$ and $\gcd(a, b) = b$. If $r_1 \neq 0$, divide b by r_1 to produce integers q_2 and r_2 such that

$$b = r_1 q_2 + r_2, \text{ where } 0 \leqslant r_2 < r_1.$$

If $r_2 = 0$, then we stop and write $\gcd(a, b) = r_1$. If $r_2 \neq 0$, we continue the process. This results in the system of equations:

$$\begin{aligned}
a &= bq_1 + r_1,\ 0 < r_1 < b \\
b &= r_1 q_2 + r_2,\ 0 < r_2 < r_1 \\
r_1 &= r_2 q_3 + r_3,\ 0 < r_3 < r_2 \\
&\vdots \\
&\vdots \\
r_{n-1} &= r_{n-1} q_n + r_n,\ 0 < r_n < r_{n-1} \\
r_{n-1} &= r_n q_{n+1} + 0
\end{aligned}$$

Now, r_n, the last non-zero remainder, is the greatest common divisor of a and b by Theorem 7.

Example 14

Find the greatest common divisor of 306 and 657.

Solution

$$\begin{aligned}
657 &= 306 \cdot 2 + 45 \\
306 &= 45 \cdot 6 + 36 \\
45 &= 36 \cdot 1 + 9 \\
36 &= 9 \cdot 4 + 0
\end{aligned}$$

Thus, $\gcd(306, 657) = 9$.

Example 15

Find the greatest common divisor of 7469 and -2387.

Solution

We know that $\gcd(-2387, 7469) = \gcd(2387, 7469)$.

$$\begin{aligned}
7469 &= 3287 \cdot 3 + 308 \\
2387 &= 308 \cdot 7 + 321 \\
308 &= 231 \cdot 1 + 77 \\
231 &= 77 \cdot 3 + 0
\end{aligned}$$

Hence, $\gcd(-2387, 7469) = 77$.

Application

Euclid's algorithm may be used to find integers x and y such that $\gcd(a, b) = ax + by$.

Example 16

Find $x, y \in \mathbb{Z}$ such that $\gcd(4147, 10\,672) = 4147x + 10\,672y$.

Solution

Using the Euclidean algorithm, we have

$$\begin{aligned}
10\,672 &= 4147\cdot 2 + 2378 \ldots\ldots\ldots\ldots (0) \\
4147 &= 2378\cdot 1 + 1769 \ldots\ldots\ldots\ldots (1) \\
2378 &= 1769\cdot 1 + 609 \ldots\ldots\ldots\ldots (2) \\
1769 &= 609\cdot 2 + 551 \ldots\ldots\ldots\ldots (3) \\
609 &= 551\cdot 1 + 58 \ldots\ldots\ldots\ldots (4) \\
551 &= 58\cdot 9 + 29 \ldots\ldots\ldots\ldots (5) \\
58 &= 29\cdot 2 + 0
\end{aligned}$$

Thus, $\gcd(4147, 10\,672) = 29$. Now,

- From (5), $29 = 551 - 9(58)$.
- From (4), $29 = 551 - 9(609 - 551) = 10(551) - 9(609)$.
- From (3), $29 = 10(1769 - 2(609)) - 9(609) = 10(1769) - 29(609)$.
- From (2), $29 = 10(1769) - 29(2378 - 1769) = 39(1769) - 29(2378)$.
- From (1), $29 = 39(4147 - 2378) - 29(2378) = 39(4147) - 68(2378)$.
- From (0), $29 = 39(4174) - 68(10\,762 - 2(4147))$
  $= 175(4174) - 68(10\,762)$.

The last statement gives us the required expression, i.e. $29 = 175(4174) - 68(10\,672)$.

In this case, $x = 175$ and $y = -68$.

Example 17

Find $x, y \in \mathbb{Z}$ such that $\gcd(-180, 252) = -180x + 252y$.

Solution

Using the Euclidean algorithm, we have

$$252 = 180 \cdot 1 + 72$$
$$180 = 72 \cdot 2 + 36$$
$$72 = 36 \cdot 2 + 0$$

Hence, $\gcd(-180, 252) = \gcd(180, 252) = 36$. Now,

$$36 = 180 - 2(72) = 180 - 2(252 - 180) = 3(180) - 2(252).$$

So, $36 = -3(-180) - 2(252)$.

In this case, $x = -3$ and $y = -2$.

Example 18

Find $x, y \in \mathbb{Z}$ such that $\gcd(143, 252) = 143x + 252y$.

Solution

Using the Euclidean algorithm, we have

$$252 = 143 \cdot 1 + 109$$
$$143 = 109 \cdot 1 + 34$$
$$109 = 34 \cdot 3 + 7$$
$$34 = 7 \cdot 4 + 6$$
$$7 = 6 \cdot 1 + 1$$
$$6 = 1 \cdot 6 + 0$$

Hence, $\gcd(143, 252) = 1$ (143 and 252 are said to be relatively prime). Now,

$$\begin{aligned} 1 &= 7 - 6 = 7 - (34 - 7(4)) = 5(7) - 34 \\ &= 5(109 - 3(34)) - 34 = 5(109) - 16(34) \\ &= 5(109) - 16(143 - 109) = 21(109) - 16(143) \\ &= 21(252 - 143) - 16(143) = 21(252) - 37(143). \end{aligned}$$

So, $1 = 21(252) - 37(143)$ or $1 = -37(143) + 21(252)$.

Here, $x = -37$ and $y = 21$.

Example 18 triggers a new definition and a new theorem.

Definition 3

Two integers a and b, not both zero, are said to be relatively prime if $\gcd(a, b) = 1$.

So, 143 and 252 are relatively prime. 12 and 25 are relatively prime because $\gcd(12, 25) = 1$; however, 18 and 24 are not relatively prime because $\gcd(18, 24) = 6$.

Theorem 8

Let a and b be integers, not both zero. Then a and b are relatively prime if and only if there exist integers x and y such that $ax + by = 1$.

Proof

If a and b are relatively prime, so that $\gcd(a, b) = 1$, then Theorem 5 guarantees the existence of x and y satisfying $1 = ax + by$.

Now, suppose on the other hand, $1 = ax + by$ for some integers x and y. Let $g = \gcd(a, b)$. Since $g | a$ and $g | b$, then $g | (ax + by)$ by Corollary 1 of Theorem 2. This means that $g | 1$, which is only possible if $g = 1$, since g has to be positive.

Therefore, if a and b are relatively prime, then there exist two integers x and y such that $ax + by = 1$.

Example 19

Find $\gcd(14, 75)$ and write it in the form $14x + 75y$.

Solution

$$75 = 14 \cdot 5 + 5$$
$$14 = 5 \cdot 2 + 4$$
$$5 = 4 \cdot 1 + 1$$

So, $\gcd(14, 75) = 1$.

Now,

$$\begin{aligned} 1 &= 5 - 4 = 5 - (14 - 5(2)) = 3(5) - 14 = 3(75 - 14(5)) - 14 \\ &= 3(75) - 16(14) \\ &= -16(14) + 3(75). \end{aligned}$$

Example 20

Find $\gcd(49, 60)$ and write it in the form $49x + 60y$.

Solution

$$60 = 49 \cdot 1 + 11$$
$$49 = 11 \cdot 4 + 5$$
$$11 = 5 \cdot 2 + 1$$

So, $\gcd(49, 60) = 1$.

Now,

$$1 = 11 - 5 \cdot 2 = 11 - (49 - 11 \cdot 4) \cdot 2 = 9 \cdot 11 - 2 \cdot 49$$
$$= 9(60 - 49) - 2 \cdot 49 = 9(60) - 11(49) = -11(49) + 9(60).$$

Corollary 3

If $\gcd(a, b) = g$, then $\gcd\left(\frac{a}{g}, \frac{b}{g}\right) = 1$.

Proof

Since $\gcd(a, b) = g$, then by Theorem 5, it is possible to find integers x and y such that

$$g = ax + by.$$

Dividing both sides of the equation by g, we obtain

$$1 = \left(\frac{a}{g}\right)x + \left(\frac{b}{g}\right)y.$$

Now, using Theorem 8, we conclude that $\left(\frac{a}{g}\right)$ and $\left(\frac{b}{g}\right)$ are relatively prime, and hence

$$\gcd\left(\frac{a}{g}, \frac{b}{g}\right) = 1.$$

Note: Even though $\left(\frac{a}{g}\right)$ and $\left(\frac{b}{g}\right)$ appear as fractions, they are, in fact, integers because g is a divisor of both a and b.

Example

- $\gcd(180, 252) = 36 \Rightarrow \gcd\left(\frac{180}{36}, \frac{252}{36}\right) = \gcd(5, 7) = 1$
- $\gcd(4147, 10\,672) = 29 \Rightarrow \gcd\left(\frac{4147}{29}, \frac{10\,672}{29}\right) = \gcd(143, 368) = 1$
- $\gcd(-2387, 7469) = 77 \Rightarrow \gcd\left(\frac{-2387}{77}, \frac{7469}{77}\right) = \gcd(-31, 97) = 1$

Corollary 4

If $\gcd(a, b) = 1$, and if $a \mid c$ and $b \mid c$, then $ab \mid c$.

Proof

$a \mid c \Rightarrow c = ma$, and $b \mid c \Rightarrow c = nb$, and

$\gcd(a, b) = 1 \Rightarrow 1 = ax + by$ for some $x, y \in \mathbb{Z}$.

Multiplying the last equation by c renders

$c = cax + cby$, and with appropriate substitution of the values for c on the right-hand side, we have

$c = nbax + maby = ab(nx + my)$, which leads to the conclusion that $ab \mid c$.

Example

$\gcd(9, 14) = 1$, $9 \mid 756$ and $14 \mid 756$, then $9 \cdot 14 = 126 \mid 756$. In fact, $756 = 6 \cdot 126$.

Two other theorems of interest are detailed below.

Theorem 9

This is sometimes called Euclid's lemma.

If $a \mid bc$, and if $\gcd(a, b) = 1$, then $a \mid c$.

Proof

Since $1 = ax + by$, then $c = acx + bcy$. Obviously $a \mid ac$ and $a \mid bc$ which is given, and thus $a \mid (acx + bcy)$; therefore $a \mid c$.

Theorem 10

Let $a, b \in \mathbb{Z}$ not both zero. For a positive integer d, $d = \gcd(a, b)$ iff:

1 $d \mid a$ and $d \mid b$.

2 If $c \mid a$ and $c \mid b$, then $c \mid d$.

This is sometimes considered as an alternative to Theorem 5.

Proof

($\Rightarrow$) If $d = \gcd(a, b)$, then obviously $d \mid a$ and $d \mid b$. Also, $d = ax + by$, and if $c \mid a$ and $c \mid b$, then $c \mid (ax + by)$, i.e. $c \mid d$.

($\Leftarrow$) If $d \mid a$ and $d \mid b$, then d is a common divisor of a, and b. If $c \mid a$ and $c \mid b$, then $c \mid d$, then $d \geqslant c$, which means that d is greater than any divisor of a and b, and thus it is the greatest common divisor of a and b.

Note: The gcd can be extended to more than two integers. We can define it in a similar manner:

Let $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ with $a_1, a_2, \ldots, a_n$ not all zero. The greatest common divisor of $a_1, a_2, \ldots, a_n$, denoted $\gcd(a_1, a_2, \ldots, a_n)$, is the greatest integer d such that d divides $a_1, a_2, \ldots, a_n$.

For example, to find the greatest common divisor of $(18, 36, 63)$, we can perform the process by taking $\gcd(18, 36) = 18$, and then $\gcd(18, 63) = 9$. Or for $\gcd(30, 42, 70)$, we find $\gcd(30, 70) = 10$, and then $\gcd(10, 42) = 2$. Or for $\gcd(36, 48, 54, 126)$, we find $\gcd(36, 48) = 12$, and $\gcd(54, 126) = 18$, and so $\gcd(12, 18) = 6$.

Note:

- If $g = \gcd(a, b)$, and if k is an integer, then $\gcd(ka, kb) = kg$.
- If $g = \gcd(a, b)$, and if k is an integer, then $\gcd(a, b + ka) = g$.

The proofs are left for you as exercises.

Least common multiple

In this section we will discuss the smallest integer which is divisible by two given integers a and b.

We call such an integer the **least common multiple** of a and b. We will also investigate its relation with $\gcd(a, b)$.

Definition 4

Let $a, b, c \in \mathbb{Z}$ with $a, b > 0$. Then a **common multiple** of a and b is a number c such that $a|c$ and $b|c$.

Example

36 is a common multiple of 12 and 18 since $12 | 36$ and $18 | 36$.

Definition 5a

Let $a, b \in \mathbb{Z}$ and $a, b > 0$. Then the smallest positive integer l such that l is a multiple of a and b is called the **least common multiple** of a and b. l is denoted by $\text{lcm}(a, b)$.

The existence of $l = \text{lcm}(a, b)$ follows from the well-ordering principle. To see this, let S be the set of all positive multiples of a and b with $a, b > 0$. S is a non-empty set, since $a, b \in S$. By the well-ordering principle, S has a least element, say l. l is the $\text{lcm}(a, b)$.

A slightly different definition of the lcm is given below. It may prove to be more appropriate for proofs later on.

Definition 5b

The **least common multiple** of two integers a and b, denoted by $\text{lcm}(a, b)$, is the positive integer m satisfying the following:

1. $a|m$ and $b|m$.
2. If $a|c$ and $b|c$, with $c > 0$, then $m \leqslant c$.

Note: Given non-zero integers a and b, $\text{lcm}(a, b)$ always exists and $\text{lcm}(a, b) \leqslant |ab|$.

Theorem 11

For positive integers a and b,

$$\gcd(a, b) \cdot \text{lcm}(a, b) = ab.$$

Proof (Optional)

Let $e = \dfrac{a}{g}$ and $f = \dfrac{b}{g}$. Then $\dfrac{ab}{g^2} = ef \Leftrightarrow \dfrac{ab}{g} = gef$. Since a, b, and g are positive integers, gef is also a positive integer.

We show now that $gef = \text{lcm}(a, b)$.

Since $gef = (ge)f = af$ and $gef = egf = e(gf) = eb$, gef is a common multiple of a and b.

Now, let $l = \dfrac{ab}{g}$ and c be another common multiple of a and b.

Let $c = au$ and $c = bv$, where u and v are positive integers.

Also, by Theorem 5, there are integers x and y such that $g = ax + by$.

Hence,

$$\frac{c}{l} = \frac{cg}{ab} = \frac{c(ax+by)}{ab} = \left(\frac{c}{b}\right)x + \left(\frac{c}{a}\right)y = vx + uy\cdot$$

Thus, $l|c$ and we conclude that $l \leqslant c$.

By the definition of $\text{lcm}(a, b)$, $l = \text{lcm}(a, b) = \dfrac{ab}{\gcd(a, b)}$.

Thus,

$$\gcd(a, b) \cdot \text{lcm}(a, b) = ab.$$

Example 21

Find

a) lcm(36, 63)

b) lcm(396, 756)

c) lcm(2387, 7469).

Solution

a) Since $\gcd(36, 63) = 9$, then $\text{lcm}(36, 63) = \dfrac{36 \cdot 63}{9} = 252$.

b) Since $\gcd(396, 756) = 36$, then $\text{lcm}(396, 756) = \dfrac{396 \cdot 756}{36} = 8316$.

c) Since $\gcd(2387, 7469) = 77$, then $\text{lcm}(2387, 7469) = \dfrac{2387 \cdot 7469}{77} = 231\,539$.

Note: If $\text{lcm}(a, b) = l$, and if k is an integer, then $\text{lcm}(ka, kb) = kl$. The proof is left for you as an exercise.

Exercise 1.3

In questions 1–6 find the greatest common divisor by Euclidean algorithm.

1 $a = 172, b = 64$ **2** $a = 167, b = 117$

3 $a = -323, b = 221$ **4** $a = 1292, b = 884$

5 $a = 7469, b = -2387$ **6** $a = 11\,143, b = 8749$

In questions 7–12 find integers x and y such that:

7 $2 = 32x + 78y$ **8** $13 = 91x + 104y$

9 $6 = 3054x + 12\,378y$ **10** $\gcd(-119, 272) = -119x + 272y$

11 $\gcd(1769, 2378) = 1769x + 2378y$

12 $\gcd(-2059, 2581) = -2059x + 2581y$

13 Do integers x and y exist such that $x + y = 100$ and $\gcd(x, y) = 8$?

14 Let a and b be relatively prime integers. Prove that $\gcd(a + b, a - b)$ is either 1 or 2.

15 Let $a, b \in \mathbb{Z}$ with a and b both non-zero. Prove that $\gcd(ca, cb) = |c|\gcd(a, b)$ for any non-zero integer c.

16 Let $a, b \in \mathbb{Z}$ with $\gcd(a, b) = 1$ and $c|(a + b)$. Prove that $\gcd(a, c) = 1$ and $\gcd(b, c) = 1$.

17 Find lcm(152, 236).

18 Find lcm(336, 746).

19 Find lcm(100, 105).

20 Find all pairs of positive integers whose greatest common divisor is 12 and the least common multiple is 360.

21 If two integers a and b have greatest common divisor 1, what can you say about $\text{lcm}(a, b)$? Give a reason for your answer.

22 You are given positive integers a, b, and c. If $\gcd(a, b, c) = g$, is it true that $\text{lcm}(a, b, c) = abc \div g$? If your answer is yes, find lcm(24, 42, 28).

23 Show that $\gcd(a, b) = \gcd(|a|,|b|)$.

24 Show that $\text{lcm}(a, b) = \text{lcm}(|a|,|b|)$.

25 Show why $\dfrac{ab}{\text{lcm}(a, b)}$ must be an integer when $a, b \neq 0$.

26 Prove that $\gcd(k, k + 2) = 2$ when k is even and $\gcd(k, k + 2) = 1$ when k is odd.

27 If $k \in \mathbb{Z}^+$, show that $\text{lcm}(k, k + 2) = \dfrac{k(k + 2)}{2}$ when k is even, and $\text{lcm}(k, k + 2) = k(k + 2)$ when k is odd.

28 If $k \in \mathbb{Z}^+$, show that $\gcd(a, a + k) = \gcd(a, k)$.

29 Let $a, b, c \in \mathbb{Z}\backslash\{0\}$. Show that if $a = bx + cy$, then $\gcd(b, c) \leqslant \gcd(a, b)$.

30 Let $a, b, c \in \mathbb{Z}\backslash\{0\}$. Show that if $a = bx + cy$, then $\gcd(b, c)|\gcd(a, b)$.

1.4 Fundamental theorem of arithmetic

Prime numbers

Consider the following numbers and their divisors:

Number	Divisors
2	1, 2
3	1, 3
4	1, 2, 4
5	1, 5
6	1, 2, 3, 6
7	1, 7
8	1, 2, 4, 8
15	1, 3, 5, 15

You can clearly see that 2, 3, 5, and 7 each have two divisors, 1 and the number itself. The numbers 4, 6, 8, and 15 have additional divisors other 1 and the number itself. This leads to the following definition.

Definition 6

Every integer, p, greater than one which has only p and 1 for its divisors is called a **prime number**. If an integer $n > 1$ is not prime, then it is called a **composite number**.

For instance, integers 2, 3, 5, and 7 are prime numbers, while 4, 6, 8, and 15 are composite numbers.

Note:

- By definition, 1 is neither prime nor composite!
- 2 is the only even integer that is prime, all other even integers are composite. Every even integer can be written in the form $2n$, where n is an integer. As such, every integer has at least two divisors, 2 and n, different from 1 and itself.

  For instance, $6 = 2 \cdot 3$ has 2 and 3 as divisors in addition to 1 and 6; $18 = 2 \cdot 9$ has several divisors, but at least two are immediately apparent, 2 and 9. The other divisors of 18 are 3 and 6.

Example

Prime numbers between 2 and 100 are: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, and 97.

Theorem 12

Every integer greater than 1 has a prime divisor.

Proof

We will present an indirect proof.

Suppose that not every integer greater than 1 has a prime divisor. Then there is an integer $n > 1$ which has no prime divisor. Let S be the collection of all integers greater than 1 with no prime divisors. Since, by assumption, $n > 1$ has no prime divisors, $n \in S$. S is a non-empty subset of natural numbers. By the well-ordering principle, S has a least element, say m. Since m has no prime divisors, m is not a prime. Hence, there exist $a, b \in \mathbb{Z}$ such that $m = ab$ with $1 < a < m$ and $1 < b < m$. Since $1 < a < m$, a has a prime divisor, say p. So $p | m$ which contradicts that m has no prime divisor. This proves that every integer greater than 1 has a prime divisor.

Example

Integers that are larger than 1 are even or odd. If a number m is even, then we can write it as $m = 2n$, and hence it has at least one prime divisor, 2. If the number is odd, then either it is a prime number, and that satisfies the theorem, or it has at least one of the following prime numbers as a divisor: 3, 5, 7, 11,…, and that satisfies the theorem too! Here are some numbers:

9 has 3 as a divisor, 11 is prime, 21 has 3 as a divisor, 143 has 11 as a divisor, 149 is prime.

Our next result shows that there are infinitely many primes. The proof of this result appears in Proposition 20 in Book IX of Euclid's *Elements*. This proof demonstrates a higher level of thinking and great mathematical ingenuity.

Theorem 13

There are infinitely many prime numbers.

Proof

Assume the result is not true. Then there are a finite number of primes. Let us label these primes $p_1, p_2, \ldots, p_n$. Let $N = p_1 \cdot p_2 \cdot \ldots \cdot p_n + 1$.

Here N is an integer greater than 1. By Theorem 12, N has a prime divisor p.

Since $p_1, p_2, \ldots, p_n$ are all the primes, p has to be one of these, say p_i for some $i = 1, 2, \ldots, N$. Since $p_i | N$ and $p_i \mid p_1 \cdot p_2 \cdot \ldots \cdot p_n$, then $p_i \mid N - p_1 \cdot p_2 \cdot \ldots \cdot p_n$, i.e. $p_i | 1$, a contradiction, since $p_i > 1$. Hence, there are infinitely many primes.

Example

Mathematicians still compete to find the largest prime number. The following are some of the numbers discovered.

- $48\,047\,305\,725 \cdot 2^{172\,403} - 1$
- $34\,790! + 1$
- $2^{43\,112\,609} - 1$

Theorem 14

Let n be a composite number. Then n has a prime divisor p with $p \leqslant \sqrt{n}$.

Proof

Given that n is a composite number, there exists $a, b \in \mathbb{Z}$ such that $n = ab$, with $1 < a < n$ and $1 < b < n$.

Without loss of generality, let us assume $a \leqslant b \cdot n = ab$ implies that $a \leqslant \sqrt{n}$ because if $a > \sqrt{n}$, then $n = ab > \sqrt{n}\sqrt{n} = n$ which is impossible. By Theorem 12, a has a prime divisor. Let this divisor be p. Hence, $p | a$ and $n = ab$ implies $p | n$.

Furthermore, $p \leqslant a \leqslant \sqrt{n}$.

Theorem 14 provides a method of finding all prime numbers less than or equal to n. This was first given by the Greek mathematician Eratosthenes of Cyrene (276 BC–194 BC).

Example

Suppose that we wish to find all prime numbers less than or equal to 50. By Theorem 14, every composite number less than or equal to 50 has a prime divisor less than or equal to $\sqrt{50} = 7.07106\ldots$.

Such prime numbers are 2, 3, 5, and 7. Hence, from the list of integers from 2 to 50, we delete all multiples of 2, 3, 5, and 7, excluding 2, 3, 5, and 7. Applying this, we have

2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~	11
~~12~~	~~13~~	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~	~~21~~
~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~	31
~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~	41
~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~	~~50~~	

Any number which is in this list after removing the multiples of 2, 3, 5, and 7 cannot be composite by Theorem 14.

This method is called the **sieve of Eratosthenes**.

Note: Theorem 14 also provides an algorithm for testing whether a given positive integer $n > 1$ is prime or composite. To do this, determine all prime numbers less than or equal to $\sqrt{n}$, then test out if n is divisible by those primes. If n is divisible, then it is composite, otherwise it is a prime number.

Example 22

Test if 227 is a prime or composite number. Repeat with 456.

Solution

227: $\sqrt{227} = 15.066$. Hence, prime numbers less than 15 are 2, 3, 5, 7, 11, and 13. A simple divisibility test shows that 227 is not divisible by any of these numbers and thus it is prime.

456: $\sqrt{457} = 21.38$. Hence, prime numbers less than 21 are 2, 3, 5, 7, 11, 13, 17, and 19. A simple divisibility test shows that 457 is not divisible by any of these numbers and thus it is prime.

Note: If two prime numbers differ by two, then such pairs of prime numbers are called **twin primes.**

Examples of some twin primes are 3, 5; 5, 7; 11, 13; 17, 19; 29, 31; etc.

The twin prime conjecture

There are infinitely many prime numbers p such that $p + 2$ is also a prime number.

This is still an unsolved conjecture. At the time of writing, the largest known pair of twin primes are $65\,516\,468\,355 \cdot 2^{333\,333} \pm 1$.

Many problems in number theory deal with integers that are expressible in certain forms. For example, the even numbers 4, 6, 8, 10, 12, and 14 are expressed as the sum of two prime numbers, not necessarily distinct:

$$4 = 2 + 2, \quad 6 = 3 + 3, \quad 8 = 3 + 5, \quad 10 = 55, \quad 12 = 5 + 7, \quad 14 = 7 + 7.$$

This led Christian Goldbach to make the following conjecture in 1742.

The Goldbach conjecture

Every even integer greater than 2 can be expressed as the sum of two (not necessarily distinct) prime numbers.

Some extra problems

In this section we solve some additional problems to gain a better understanding of the methods previously outlined.

Example 23

Prove that if p is a prime and $p \mid a^k$ for some positive integer k, then $p \mid a$ and $p^k \mid a^k$. Is this valid if p is a composite number?

Solution

Since $a^k = a \cdot a \cdot \ldots \cdot a$ (k times), $p \mid a^k$ implies $p \mid a$. Hence, there is an integer q such that $a = pq$. Then $a^k = p^k q^k$ and consequently $p^k \mid a^k$.

This does not hold for all composite numbers. For example, take $p = 4$ and $a = 2$: $4 \mid p^k$ for $k = 2$, $4 \mid 2^2$, but $4 \nmid 2$.

Example 24

If $2^m + 1$ is prime, then prove that $m = 2^n$ for some integer $n \geqslant 0$.

Solution

We shall prove this by showing that if m is not a power of 2, then $2^m + 1$ is not a prime. If m is not a power of 2, then m has the form $2^n q$ for some odd integer $q > 1$.

$f(t) = t^q + 1$ is divisible by $t + 1$ (since $t^q + 1 = (t + 1)(t^{q-1} - t^{q-2} + \ldots + 1)$). Substituting $t = x^{2^n}$, we find that $2^{2^n} + 1$ divides $g(2) = 2^m + 1$. This implies that $2^m + 1$ cannot be a prime. This argument proves that when m is not a power of 2, $2^m + 1$ is not a prime. By using equivalence of statements, $P \Rightarrow Q$ and $\neg Q \Rightarrow \neg P$, we complete the proof of the result.

The fundamental theorem of arithmetic

The fundamental theorem of arithmetic appeared in Proposition 14 in Book 1 of Euclid's *Elements*. This is the first big result in number theory and guarantees that any integer greater than 1 can be decomposed uniquely into a product of prime numbers.

Example

$12 = 2 \times 2 \times 3 = 2^2 \times 3, 56 = 2^3 \times 7, 124 = 2^2 \times 31, 11\,430 = 2 \times 3^2 \times 5 \times 127$

Theorem 15

Let a, b, $p \in \mathbb{Z}$, with p a prime number. If $p \mid ab$, then $p \mid a$ or $p \mid b$.

Proof

Suppose $p \nmid a$. Then $\gcd(a, p) = 1$. Then there are integers m and n such that $ma + np = 1$. Also $p \mid ab$ means that there is an integer c such that $ab = pc$. Now multiplying both sides of $ma + np = 1$ by b, we get $mab + npb = b$. Using $ab = pc$, $mab + npb = pc$ reduces to $p(mc + nb) = b$. So $p \mid b$. This can be repeated for the case $p \nmid b$, and the conclusion would be $p \mid a$.

One may wonder if it is necessary that p be a prime in Theorem 15. In fact, the theorem fails to hold when p is a composite number. For example, take $p = 6$ and $a = 9$ and $b = 8$: $6 \mid (8 \cdot 9)$, but $6 \nmid 8$ and $6 \nmid 9$.

We can show that if $a_1, a_2, \ldots, a_n, p \in \mathbb{Z}$, with p a prime, and $p \mid a_1 \cdot a_2 \cdot \ldots \cdot a_n$, then $p \mid a_k$ for some $1 \leqslant k \leqslant n$.

We are now in a position to state the most important theorem of this section.

Theorem 16 (The fundamental theorem of arithmetic)

Every integer n greater than 1 can be expressed in the form $n = p_1^{a_1} \cdot p_2^{a_2} \cdot \ldots \cdot p_n^{a_n}$ with distinct prime numbers $p_1, p_2, \ldots, p_n$ and positive integers $a_1, a_2, \ldots, a_n$.

Proof (Outline – optional)

We must prove two things:

1 Every positive integer can be expressed as a product of primes.

2 The expression in **1** is unique.

First, we use strong induction to prove that every positive integer n is a product of primes. As a base case, $n = 1$ is the product of the empty set of primes. (A standard convention: the product of an empty set of numbers is defined to be 1, much as the sum of an empty set of numbers is defined to be 0. Without this convention the theorem would not be true for $n = 1$. In that case we can choose another value.) For the inductive step, suppose that every $k < n$ is a product of primes. We must show that n is also a product of primes.

We must show that n is also a product of primes. If n is itself prime, then this is true trivially. Otherwise, $n = ab$ for some $a, b < n$. By the induction assumption, a and b are both products of primes. Therefore, $a \cdot b = n$ is also a product of primes. Thus, the claim is proved by induction.

Second, we use the well-ordering principle to prove that every positive integer can be written as a product of primes in a unique way. The proof is by contradiction: assume, contrary to the claim, that there exist positive integers that can be written as products of primes in more than one way. By the well-ordering principle, there is a smallest integer with this property. Call this integer n, and let

$$n = p_1 \cdot p_2 \cdot \ldots \cdot p_j = q_1 \cdot q_2 \cdot \ldots \cdot q_k$$

be two of the (possibly many) ways to write n as a product of primes. Now, $p_1 \mid n$ and so

$$p_1 \mid q_1 \cdot q_2 \cdot \ldots \cdot q_k.$$

By the previous theorem, this implies that p_1 divides one of the primes q_i. But since q_i is a prime, it must be that $p_1 = q_i$. Deleting p_1 from the first product and q_i from the second, we find that n/ p_1 is a positive integer *smaller* that n that can also be written as a product of primes in two distinct ways. But this contradicts the definition of n as the smallest such positive integer. Thus, the assumption is false and we have one way of writing the product of primes.

Example

Prime factorization of $132 = 2^2 \cdot 3 \cdot 11$.

Prime factorization of $3780 = 2^2 \cdot 3^3 \cdot 5 \cdot 7$.

We can use the fundamental theorem to find the gcd and lcm of two or more integers.

Example 25

Find gcd(132, 3780) and lcm(132, 3780).

Solution

We have from the previous example:

$132 = 2^2 \cdot 3 \cdot 11$ and $3780 = 2^2 \cdot 3^3 \cdot 5 \cdot 7$

For gcd(132, 3780), we compare the exponents appearing on like prime numbers and choose the minimum exponent appearing in prime factorizations of 132 and 3780 (since gcd(132, 3780) is the largest common divisor of 132, 3780).

So, $\gcd(132, 3780) = 2^2 \cdot 3 = 12$.

Similarly for lcm(132, 3780), we compare the exponents appearing on like prime numbers and choose the maximum exponent appearing in their prime factorization.

Since $132 = 2^2 \cdot 3 \cdot 11 = 2^2 \cdot 3^1 \cdot 5^1 \cdot 7^0 \cdot 11^1$ and $3780 = 2^2 \cdot 3^3 \cdot 5^1 \cdot 7^1 \cdot 11^0$, $\operatorname{lcm}(132, 3720) = 2^2 \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 = 4180$.

We can now state what we have done in Example 18 as a theorem (proof not included here).

Theorem 17

Let $a, b \in \mathbb{Z}$ with $a, b > 1$. Let $a = p_1^{a_1} \cdot p_2^{a_2} \cdot \ldots \cdot p_n^{a}$ and $b = p_1^{b_1} \cdot p_2^{b_2} \cdot \ldots \cdot p_n^{b_n}$, where $p_1, p_2, \ldots, p_n$ are distinct prime numbers and $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are non-negative integers (some of these may be 0). Let m_i be the smaller and M_i be the larger of a_i and b_i for $i = 1, 2, \ldots, n$. Then,

$\gcd(a, b) = p_1^{m_1} p_2^{m_2} \ldots p_n^{m_n}$, and

$\operatorname{lcm}(a, b) = p_1^{M_1} p_2^{M_2} \ldots p_n^{M_n}$.

Example 26

Using the fundamental theorem of arithmetic, find gcd(1176, 936) and lcm (1176, 936).

Solution

$1176 = 2^3 \cdot 3 \cdot 7^2$; $936 = 2^3 \cdot 3^2 \cdot 13$, and hence:

$\gcd(936, 1176) = 2^3 \cdot 3 = 24$

$\operatorname{lcm}(936, 1176) = 2^3 \cdot 3^2 \cdot 7^2 \cdot 13 = 45\,864$

This method of finding the gcd and lcm of two positive integers a and b is easily used to find the gcd and lcm of three or more positive integers. We consider the following as an illustration.

Example 27

Find gcd(132, 936, 1176) and lcm(132, 936, 1176).

Solution

$132 = 2^2 \cdot 3 \cdot 11, \quad 936 = 2^3 \cdot 3^2 \cdot 13, \quad 1176 = 2^3 \cdot 3 \cdot 7^2$

$\gcd(132, 936, 1176) = 2^2 \cdot 3 = 12$

$\operatorname{lcm}(132, 936, 1176) = 2^3 \cdot 3^2 \cdot 7^2 \cdot 11 \cdot 13 = 504\,504$

Exercise 1.4

1 Prove that there are infinitely many primes of the form $4q + 3$, $q = 0, 1, \ldots$.

2 Prove that every prime $p \neq 3$ has the form $3q + 1$ or $3q + 2$ for some integer q.

3 Prove that there are infinitely many primes of the form $3q + 2$.

4 Prove that only for the prime number $p = 3$, $p^2 + 2$ is a prime.

5 If $2^p - 1$ is a prime number, then show that $2^{p-1}(2^r - 1)$ is equal to the sum of its proper divisors.

6 From $5 = 2^2 + 1$, $17 = 4^2 + 1$, $37 = 6^2 + 1$, $101 = 10^2 + 1$, and $197 = 14^2 + 1$, what kind of conjecture can you propose for primes of the form $n^2 + 1$?

7 Find the prime factorization of each integer given below.

a 87 **b** 361 **c** 945 **d** 1001 **e** 6992

8 Using the fundamental theorem of arithmetic, find the following:

a gcd(87, 361) and lcm(87, 361)

b gcd(361, 1001) and lcm(361, 1001)

c gcd(87, 361, 1001) and lcm(87, 361, 1001)

d gcd(87, 945, 6992) and lcm(87, 945, 6992)

9 Find five integers that are relatively prime (when taken together) such that no two of the integers are relatively prime when taken separately.

10 Let a and b be positive integers.

a Prove that $\gcd(a, b) \mid \operatorname{lcm}(a, b)$.

b Find and prove a necessary and sufficient condition for $\gcd(a, b) = \operatorname{lcm}(a, b)$.

c Prove that $\operatorname{lcm}(ca, cb) = c \operatorname{lcm}(a, b)$.

11 Let $\gcd(a, b) = g$. Show that if $a \mid bc$, then $a \mid gc$.

12 Show that if a and b are relatively prime, then a^2 and b^2 are also relatively prime.

In questions 13–16, use prime factors to decide whether $x|y$, to find gcd(x, y), and to find lcm(x, y).

13 $x = 585$, $y = 14\,157$ **14** $x = 11\,500$, $y = 4232$

15 $x = 2277$, $y = 15\,939$ **16** $x = 1870$, $y = 2275$

In questions 17–22, prove each statement if it is true, or show that it is false either by reasoning or by finding a counter example.

17 For all integers x, $x > 2$, $x^3 - 8$ is composite.

18 If $m^2|n^2$ then $m|n$.

19 If $n|ab$ and $n \nmid a$, then $n|b$.

20 If $n|ab$ and gcd$(n, a) = 1$, then $n|b$.

21 gcd(a, b) = gcd$(a, b + ka)$ for all $k \in Z$.

22 gcd(a^n, b^n) = (gcd$(a, b))^n$.

23 What are the possible values of gcd$(a, a + 3)$?

24 If a and b are relatively prime, then what are the possible values of gcd$(a + b, a - b)$?

25 Under what conditions can we solve $ax + (a + 2)y = c$ for x and y?

2 Number Theory II

In Chapter 1 we dealt with all the theorems necessary to work on some applications of number theory. In this chapter we shall discuss a few of these applications.

2.1 Congruence

So far you have seen examples involving congruence for specific values. In this section we will discuss congruence in more general terms. This topic is important for this option, as well as for the abstract algebra option.

Definition 1

Let m be a positive integer. If a and b are integers, we say that a is congruent to b modulo m if $m \mid (a - b)$.

If a is congruent to b modulo m, then we write $a \equiv b \pmod{m}$. If a is not congruent to b modulo m, then we write $a \not\equiv b \pmod{m}$. The integer m is called the **modulus of congruence**.

Example

We have $24 \equiv 4 \pmod{5}$, since $5 \mid (24 - 4)$. Similarly, $5 \equiv -11 \pmod{8}$, since $8 \mid (5 - (-11))$. On the other hand, $4 \not\equiv 17 \left(\text{mod } 2\right)$, since $(4 - 17)$ is not divisible by 2.

Theorem 1

If $a, b \in \mathbb{Z}$, then $a \equiv b \pmod{m}$ for some positive integer m if and only if there exists an integer k such that $a = b + km$.

Proof

($\Rightarrow$) Since $m \mid (a - b)$ if and only if $a - b = km$ for some $k \in \mathbb{Z}$, then $a = b + km$.

($\Leftarrow$) If for some $k \in \mathbb{Z}$, $a = b + km$, $km = a - b$. Hence, $m \mid (a - b)$, and consequently $a \equiv b \pmod{m}$.

So, we can summarize this result by stating: Given a positive integer m and an integer b, integers which are congruent to b modulo m are obtained by adding integer multiples of m to b.

As an illustration, let $m = 2$ and $b = 0$. Then the integers congruent to 0 modulo 2 are given by $a = 0 + 2k$, $k \in \mathbb{Z}$, i.e. $\{\ldots, -4, -2, 0, 2, 4, \ldots\}$.

If $b = 1$, then the collection of all integers congruent to 1 are $\{\dots, -3, -1, 1, 3, \dots\}$. We can observe that these two classes of integers are distinct and each one is associated to a remainder when we divide an arbitrary integer n by 2.

This discussion leads us to the following important theorem which explains how congruence partitions the set of integers into different sets like the ones above. These are called **congruence classes modulo *m*.**

Theorem 2

$a \equiv b \pmod{m}$ if and only if a and b leave the same remainder when we divide them by m.

Proof

($\Rightarrow$) Let $a \equiv b \pmod{m}$. Then, by definition, $m \mid (a - b)$.

Now, by the division algorithm, if we divide a by m, we can find q_1 and r_1 such that

$$a = m \cdot q_1 + r_1, 0 \leqslant r_1 < m$$

and similarly, if we divide b by m, then we can find q_2 and r_2 such that

$$b = m \cdot q_2 + r_2, 0 \leqslant r_2 < m.$$

So, we now have

$$a - b = (m \cdot q_1 + r_1) - (m \cdot q_2 + r_2) = m(q_1 - q_2) + (r_1 - r_2).$$

However, $m \mid (a - b)$, and so m must divide the right-hand side, $m(q_1 - q_2) + (r_1 - r_2)$.

This leads to the fact that m must divide $(r_1 - r_2)$ too. But $0 \leqslant r_1 < m$ and $0 \leqslant r_2 < m$, and so $(r_1 - r_2)$ cannot divide m unless $r_1 - r_2 = 0$, i.e. $r_1 = r_2$.

Therefore, a and b leave the same remainder when we divide them by m.

($\Leftarrow$) Let a and b leave the same remainder when we divide them by m.

Then we have

$a = m \cdot q_1 + r$ and $b = m \cdot q_2 + r$, and consequently

$a - b = m(q_1 - q_2)$, which means that $m \mid (a - b)$ and therefore $a \equiv b \pmod{m}$.

Theorem 3

Let $m \in \mathbb{Z}^+$. Then congruence modulo m is an equivalence relation. (See Option 2 Chapter 2 for review.)

Proof

1 **Reflexive property:** $a \equiv a \pmod{m}$ since $m \mid (a - a)$ for all $a \in \mathbb{Z}$.

2 **Symmetric property:** Suppose $a \equiv b \pmod{m}$. Then there is an integer k such that $a - b = km$. Hence, $b - a = (-k)m$ and $m \mid (b - a)$ [$-k$ is also an integer]. Thus $b \equiv a \pmod{m}$.

3 **Transitive property:** If $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$, then $m \mid (a - b)$ and $m \mid (b - c)$. Hence, $m \mid \big((a-b)-(b-c)\big)$, i.e. $m \mid (a-c)$ and $a \equiv c \pmod{m}$.

Note: The two previous theorems enable us to generalize the structure of congruence classes modulo m. Since any two integers that leave the same remainder when divided by m, the remainder itself will represent the equivalence class. This is so because if a leaves a remainder r when divided by m, then as we showed before:

$a = m \cdot q_1 + r \Rightarrow a - r = m \cdot q_1 \Rightarrow m \mid (a - r) \Rightarrow a \equiv r \pmod{m}$.

Also, since $r < m$, then it takes on all the values $\{0, 1, 2, 3, \ldots, m - 1\}$, and hence the congruence classes modulo m are

$[0], [1], \ldots, [m - 1]$.

These classes are also called residue classes mod m. Also each *value of r* is called a least residue modulo m.

Example 1

List the congruence classes mod 7.

Solution

Since the possible remainders when dividing by 7 are 0, 1, 2, …, 6, then the congruence classes are:

$[0] = \{\ldots, -7, 0, 7, 14, \ldots\}$

$[1] = \{\ldots, -6, 1, 8, 15, \ldots\}$

$\vdots$

$[6] = \{\ldots, -1, 6, 13, 20, \ldots\}$

Given a positive integer m, the set of integers $\mathbb{Z}$ is partitioned into $m - 1$ congruence classes. If we pick two members of a congruence class then they are congruent modulo m. Further, $[a] = [b]$ if and only if $a \equiv b \pmod{m}$.

For a given $m \geqslant 1$, we denote the congruence classes by $\mathbb{Z}_m$, called the set of **residue classes** modulo m (also called the set of **integers modulo *m*** or the set of **least residues**). So, $\mathbb{Z}_5 = \{[0], [1], [2], [3], [4]\}$. For convenience purposes, once we make it clear that we are working with residue classes, we use the digits $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ to represent the classes.

Next, we show how to do arithmetic with these congruence classes, so that $\mathbb{Z}_m = \{k \mid k = 0, 1, \ldots, m-1\}$ behaves like a system of numbers.

For this purpose, we define arithmetic in the congruence classes as modular arithmetic.

First we know that an addition, subtraction or multiplication of both sides of a congruence preserves the congruence.

Theorem 4

If $a, b, c, m \in \mathbb{Z}$ and $m > 0$, such that $a \equiv b \pmod{m}$, then the following holds:

(i) $a + c \equiv b + c \pmod{m}$

(ii) $a - c \equiv b - c \pmod{m}$

(iii) $ac \equiv bc \pmod{m}$

Proof

$a \equiv b \pmod{m}$ implies that $m \mid (a - b)$.

Since $(a - b) = (a + c) - (b + c)$, $m \mid (a + c) - (b + c)$. Hence (i) holds.

In the same manner, (ii) follows from $(a - c) - (b - c)$.

To prove (iii), we use $ac - bc = c(a - b)$ and the fact that $m \mid (a - b)$ implies $m \mid (a - b)c$, i.e. $m \mid (ac - bc)$.

Example

Since $23 \equiv 7 \pmod{8}$, from Theorem 3,
$28 \equiv 23 + 5 \equiv 7 + 5 \pmod{8} \equiv 12 \pmod{8}$.

Also, $14 \equiv 23 - 9 \equiv (7 - 9) \bmod 8 \equiv -2 \pmod{8}$, and
$69 \equiv 23(3) \equiv 7(3) \bmod 8 \equiv 21 \pmod{8}$.

It is natural to ask if division upholds such a property – we see that it is not the case.

Example

$12 = 6 \cdot 2 \equiv 3 \cdot 2 \pmod{6}$. But $6 \not\equiv 3 \pmod{6}$. So we cannot cancel 2.

Similarly, $14 = 7 \cdot 2 \equiv 4 \cdot 2 \pmod{6}$. But $7 \not\equiv 4 \pmod{6}$.

Our next result is similar to Theorem 3. However, it generalizes the theorem.

Theorem 5

Let $a, b, c, d, m \in \mathbb{Z}$ and $m > 0$. Then $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$ imply the following:

(i) $a + c \equiv b + d \pmod{m}$

(ii) $a - c \equiv b - d \pmod{m}$

(iii) $ac \equiv bd \pmod{m}$

Proof

If $a \equiv b \pmod{m}$ and c $\equiv d \pmod{m}$, then $m \mid a - b$ and $m \mid c - d$.

These imply that $m \mid [(a-b)+(c-d)]$. But this is the same as $m \mid [(a+c)-(b+d)]$. This proves (i).

Proof of (ii) is similar.

To prove (iii), note that $m \mid (a-b)$ implies $m \mid c(a-b)$ and $m \mid (c-d)$ implies $m \mid b(c-d)$.

Thus, $m \mid [c(a-b)+b(c-d)]$, which is the same as $m \mid (ac-bd)$. This completes the proof.

Example

Since $31 \equiv 9 \pmod{11}$ and $15 \equiv 4 \pmod{11}$, by Theorem 4, we have

$31+15 \equiv 9+4 \pmod{11} \Rightarrow 46 \equiv 13 \pmod{11}$, and

$31 \times 15 \equiv 9 \times 4 \pmod{11} \Rightarrow 465 \equiv 36 \pmod{11}$.

Theorem 6

Let $a, b, c, m \in \mathbb{Z}$ with $m > 0$, and $d = \gcd(c, m)$, then

$ac \equiv bc \pmod{m} \Rightarrow a \equiv b \left(\bmod\ m/d\right)$.

Proof

If $ac \equiv bc \pmod{m}$, then we know $m \mid (ac - bc)$ or $m \mid c(a-b)$. Hence, there is an integer k such that $c(a - b) = km$. Divide both sides by d:

$$\frac{c}{d}(a-b) = k\frac{m}{d} \;\ldots\ldots\ldots\ldots(1)$$

Since, from Chapter 1 (Corollary 3), we know $\gcd\left(\frac{c}{d}, \frac{m}{d}\right) = 1$, then we know that $\frac{m}{d}$ divides the right-hand side of equation (1), so it has to divide the left-hand side, and since it is relatively prime to $\frac{c}{d}$, it should divide $(a - b)$ by Theorem 9 of Chapter 1. Therefore, $a \equiv b \left(\bmod\ m/d\right)$.

Example

$70 \equiv 40 \pmod{15}$, and $\gcd(10, 15) = 5$, then $7 \equiv 4 \pmod{3}$.

The following corollary is also helpful in solving congruence problems.

Corollary 1

Let $a, b, c, m \in \mathbb{Z}$ with $m > 0$, and $\gcd(c, m) = 1$, then

$ac \equiv bc \pmod{m} \Rightarrow a \equiv b \pmod{m}$.

The proof is a simple application of Theorem 6 when $d = 1$.

Example

$54 \equiv 24 \pmod 5$ implies that $\frac{54}{3} \equiv \frac{24}{3} \pmod 5$, i.e. $18 \equiv 8 \pmod 5$, since $\gcd(3, 5) = 1$.

Theorem 7

Let $a, b, c, m \in \mathbb{Z}$ with $c, m > 0$, then

$a \equiv b \pmod m \Rightarrow a^c \equiv b^c \pmod m$.

Proof

$a \equiv b \pmod m \Rightarrow m \mid (a - b)$. Also,

$a^c - b^c = (a - b)(a^{c-1} + a^{c-2}b + \ldots + ab^{c-2} + b^{c-1})$, then

$m \mid (a - b), (a - b) \mid (a^c - b^c) \Rightarrow m \mid (a^c - b^c)$.

Hence, $a^c \equiv b^c \pmod m$.

Example

$8 \equiv 3 \pmod 5$ implies $64 \equiv 9 \pmod 5$, or $512 \equiv 27 \pmod 5$, etc.

Theorem 8

If $a \equiv b \pmod{m_1}$, $a \equiv b \pmod{m_2}$, …, $a \equiv b \pmod{m_k}$,
where $a, b, m_1, \ldots, m_k \in \mathbb{Z}$ and $m_1, \ldots, m_k > 0$, then
$a \equiv b \pmod l$, where $l = \operatorname{lcm}(m_1, \ldots, m_k)$.

Proof

$a \equiv b \pmod{m_1}$, $a \equiv b \pmod{m_2}$, …, $a \equiv b \pmod{m_k}$ imply that $m_1 \mid (a - b)$, $m_2 \mid (a - b)$, …, $m_k \mid (a - b)$. This in turn implies that $\operatorname{lcm}(m_1, \ldots, m_k) \mid (a - b)$. (Proof is left as an exercise.)

Consequently,

$a \equiv b \pmod l$.

Note: A consequence of Theorem 8 is the situation where $m_1, \ldots, m_k$ are pairwise relatively prime. In such a case we will have

$a \equiv b \pmod{m_1 \cdot m_2 \cdot \ldots \cdot m_k}$.

Example

$342 \equiv 12 \pmod 5$, $342 \equiv 12 \pmod{10}$, $342 \equiv 12 \pmod{15}$, and $342 \equiv 12 \pmod 6$.

Since $\operatorname{lcm}(5, 10, 15, 6) = 30$, then we can conclude that $342 \equiv 12 \pmod{30}$, which is indeed true, as $342 - 12 = 330 = 30 \cdot 11$.

$342 \equiv 12 \pmod 5$, $342 \equiv 12 \pmod 2$, $342 \equiv 12 \pmod 3$, and $342 \equiv 12 \pmod{11}$. Since the moduli are pairwise relatively prime, then $342 \equiv 12 \pmod{5 \cdot 2 \cdot 3 \cdot 11}$, i.e. $342 \equiv 12 \pmod{330}$.

Exercise 2.1

1 Say whether each statement is true or false.

a $16 \equiv 49 \pmod{11}$ **b** $72 \equiv 24 \pmod 9$

c $87 \equiv 303 \pmod{16}$ **d** $-25 \equiv 215 \pmod{12}$

2 Find the least residue (mod 31) of $33 \cdot 26^2$.

3 Show that if $a \equiv b \pmod m$ and $d | m$, then $a \equiv b \pmod d$.

In questions 4–16, find the least residue of a modulo m.

4 $a = 114$, $m = 7$

5 $a = 85$, $m = 8$

6 $a = 67$, $m = 50$

7 $a = 60$, $m = 51$

8 $a = -62$, $m = 50$

9 $a = -81$, $m = 51$

10 $a = -114$, $m = 7$

11 $a = 72 \cdot 73 \cdot 74$, $m = 71$

12 $a = 80 \cdot 81 \cdot 85$, $m = 82$

13 $a = 100^6$, $m = 49$

14 $a = 49^4$, $m = 23$

15 $a = 50^{99}$, $m = 7$

16 $a = 50^{99}$, $m = 17$

17 If $x \equiv 2 \pmod{17}$, $y \equiv 4 \pmod{17}$, and $z \equiv 5 \pmod{17}$, find the least residue of $x + yz \pmod{17}$.

18 If $x \equiv 2 \pmod{17}$, $y \equiv 4 \pmod{17}$, and $z \equiv 5 \pmod{17}$, find the least residue of $x^2 + y^2 + z^2 \pmod{17}$.

19 Prove that $7^n \equiv 6n + 1 \pmod{36}$ for all $n \in \mathbb{Z}^+$.

20 Prove that $2 \cdot 7^n \equiv 2^n(5n + 2) \pmod{25}$ for all $n \in \mathbb{Z}^+$.

21 Prove that $2^n + 3^n \equiv 5^n \pmod 6$ for all $n \in \mathbb{Z}^+$.

22 Prove that $16^n \equiv 1 - 10n \pmod{25}$ for all $n \in \mathbb{Z}^+$.

23 Prove that $3 | (4^n - 1)$ for all $n \in \mathbb{Z}^+$.

24 Let f_n be the nth term of a Fibonacci sequence. Prove that
$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix} \text{ for all } n \in \mathbb{Z}^+.$$

25 Prove that $2^{2^n} + 1 \equiv 5 \pmod{12}$ for all $n \in \mathbb{Z}^+$.

26 Prove that $(-4)^n \equiv 1 - 5n \pmod{25}$ for all $n \in \mathbb{Z}^+$.

27 Prove that $5^n \equiv 1 + 4n \pmod{16}$ for all $n \in \mathbb{Z}^+$.

28 Prove that $8^n \mid (4n)!$ for all $n \in \mathbb{Z}^+$.

29 Show that $31 \mid 2^{5n} - 1$ for $n \in \mathbb{Z}^+$.

In questions 30–33, prove each statement if it is true, or show that it is false either by reasoning or by finding a counter example.

30 If a, b, and c are three consecutive integers, then $a + b + c \equiv 0 \pmod 3$.

31 If a and b are two even integers, then $ab \equiv 0 \pmod 4$.

32 If $n \in \mathbb{Z}$, $n > 1$, then $n^2 \not\equiv 3 \pmod 4$.

33 If $n \in \mathbb{Z}$, n is odd, then $n^4 - 1 \equiv 0 \pmod{16}$.

34 Find all values of a such that $a \equiv 307 \pmod{17}$, $0 \leqslant a \leqslant 33$.

35 Find all values of a such that $a \equiv 971 \pmod{23}$, $-20 \leqslant a \leqslant 50$.

36 Find all values of n such that $342 \equiv 573 \pmod{n}$.

37 Show that any integer is congruent modulo 17 to any multiple of 7.

38 Show that if $x^2 \equiv y^2 \pmod{p}$, where p is a prime, then $|a| \equiv |b| \pmod{p}$.

39 Show that $a \equiv b \pmod{n}$ implies that $\gcd(a, n) = \gcd(b, n)$.

40 The multiplicative inverse of a number a mod n is the number b such that $ab \equiv 1 \pmod{n}$. Find the multiplicative inverse, if any, of $7 \pmod{19}$, $39 \pmod{95}$ and $91 \pmod{191}$.

41 With p a prime number different from 2, show that $(p + 1)/2$ is an integer and that it is the multiplicative inverse of 2 modulo p.

42 With p a prime number different from 2, show that $(p + 1)^2/4$ is an integer and that it is the multiplicative inverse of 4 modulo p.

2.2 The Diophantine equation $ax + by = c$

The Greek mathematician Diophantus who lived in Alexandria around 250 AD considered linear equations with integer solutions. In honour of him, any equation with one or more unknowns, which is to be solved over the set of integers, is called a Diophantine equation. The simplest sort of Diophantine equation that we will consider is the linear Diophantine equation in two unknowns, $ax + by = c$.

Definition 2

A simple linear Diophantine equation in two unknowns is of the form $ax + by = c$, where a, b, and c are integers and a and b are not both zero.

A solution of the linear equation is a pair of integers x_0 and y_0 such that $ax_0 + by_0 = c$.

Before we consider the general method of solving such equations, let us consider the simple equation $2x + 4y = 16$. One solution is $x = 6$ and $y = 1$. Another solution is $x = 12$ and $y = -2$. In view of this, we expect that a linear Diophantine equation may have more than one pair of solutions. A fundamental question to ask is: Does every linear Diophantine equation have a solution? The equation $2x + 4y = 5$ does not have a solution for any integers x and y. This follows from Theorem 9 below.

Theorem 9

A linear Diophantine equation $ax + by = c$, where a, b, and c are integers and a and b are not both zero, has a solution if and only if $\gcd(a, b) \mid c$.

Proof

($\Rightarrow$) Suppose $g = \gcd(a, b)$. Then there are integers r and s such that $a = gr$ and $b = gs$.

If $ax + by = c$ has a solution (x_0, y_0), then $ax_0 + by_0 = c$. Thus, $c = ax_0 + by_0 = grx_0 + gsy_0 = g(rx_0 + sy_0)$.

This implies that $g|c$.

($\Leftarrow$) Conversely, assume that $g|c$, i.e. there exists an integer t such that $c = gt$.

By Theorem 5 of Chapter 1, there are integers u and v such that $au + bv = g$.

Hence, $atu + btv = tg = c$. Therefore, $x = tu$ and $y = tv$ form a particular solution of the equation $ax + by = c$. This completes the proof.

Our next result shows how to get all solutions of $ax + by = c$ when we know a particular solution (x_0, y_0).

Theorem 10

If $x = x_0$ and $y = y_0$ is a particular solution of the linear Diophantine equation $ax + by = c$, then other solutions are given by $x = x_0 + \left(\frac{b}{g}\right)t$ and $y = y_0 - \left(\frac{a}{g}\right)t$, where $g = \gcd(a, b)$ and t is an arbitrary integer.

Proof (Optional)

Suppose we have found a solution (x_0, y_0) of the equation $ax + by = c$. If (x_0', y_0') is any other solution of $ax + by = c$, then $ax_0 + by_0 = c = ax_0' + by_0'$, which is equivalent to $a(x_0' - x_0) = b(y_0 - y_0')$. We know that there are relatively prime integers r and s such that $a = gr$ and $b = gs$. Using these, we obtain

$$gr(x_0' - x_0) = gs(y_0 - y_0')$$

or

$$r(x_0' - x_0) = s(y_0 - y_0') \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

From (1), we see that $r \mid s(y_0 - y_0')$ with $\gcd(r, s) = 1$, and we have, by Euclid's lemma,

$r \mid (y_0 - y_0')$, and thus $(y_0 - y_0') = rl$ for some integer l.

Now substituting this in (1), we get

$x_0' - x_0 = sl$.

Thus, $x_0' = x_0 + sl = x_0 + \left(\frac{b}{g}\right)l$ and $y_0' = y_0 - rl = y_0 - \left(\frac{a}{g}\right)l$.

$$ax_0' + by_0' = a\left(x_0 + \left(\frac{b}{g}\right)l\right) + b\left(y_0 - \left(\frac{a}{g}\right)l\right)$$

$$= ax_0 + by_0 + \left(\frac{ab}{g} - \frac{ab}{g}\right)l = ax_0 + by_0 = c$$

since (x_0, y_0) is a solution of $ax + by = c$.

Thus, if a linear Diophantine equation has a solution, it has an infinite number of solutions.

The following is a direct result of Theorem 10.

Corollary 2

If a and b are relatively prime, then $ax + by = c$ has solutions given by

$x = x_0 + bt$ and $y = y_0 - at$,

where (x_0, y_0) is a particular solution of $ax + by = c$ and t is any integer.

Theorem 9 and 10 combined

Let $a, b, c \in \mathbb{Z}$. Consider the Diophantine equation

$ax + by = c$.

If $\gcd(a, b) \nmid c$, there are no solutions to the equation.

If $\gcd(a, b) \mid c$, there are infinitely many solutions of the form

$x = x_0 + \frac{b}{g}t$ and $y = y_0 - \frac{a}{g}t$,

where $g = \gcd(a, b)$, (x_0, y_0) is a particular solution, and t is any integer.

Theorems 9 and 10 are usually combined into one theorem which may be more meaningful. We used two separate theorems for the sake of easing up the proof!

Example 2

Solve $6x + 9y = 21$.

Solution

Since $\gcd(6, 9) = 3$, and $3 \mid 21$, there are an infinite number of solutions. To find them, we first attempt to find one by trial and error.
$x_0 = -4$ and $y_0 = 5$ is a particular solution.

Hence, the general solution is

$$x = -4 + \frac{9}{3}t = -4 + 3t \text{ and } y = 5 - \frac{6}{3}t = 5 - 2t.$$

How do we find a particular solution?

There is no unique answer to this question. There are a few approaches that work relatively well.

1 Trial and error, as in Example 2.

2 Using linear congruence (which you will study later in more detail).

The equation $ax + by = c$ can be rewritten as $ax - c = -by$, which implies that $ax \equiv c \pmod{b}$, which is simpler to solve.

For example:

$6x + 9y = 21 \Rightarrow 6x \equiv 21 \pmod{9}$

$\Rightarrow 2x \equiv 7 \pmod{3}$ [Theorem 6]

$\Rightarrow 2x \equiv (6 + 1) \pmod{3} \Rightarrow 2x \equiv 1 \pmod{3}$

Here we can find $x_0 = 2$ (or any number in its residual class!). Hence,

$y_0 = 1$, and our general solution is

$x = 2 + 3t$ and $y = 1 - 2t$.

When $t = -2$, we get $x = -4$ and $y = 5$, which is the solution found in Example 2.

3 Using ‘reverse’ Euclidean algorithm.

We know that $\text{gdc}(6, 9) = 3$, but to find a linear combination of 3 in terms of 6 and 9, we have to perform the algorithm first so that we can reverse it afterwards (as we did in the previous chapter). Otherwise, finding the linear combination will again be guesswork.

$9 = 1 \cdot 6 + 3$ and $6 = 2 \cdot 3 + 0$, so

$3 = 1 \cdot 9 - 6$, and now we multiply both sides by 7 to get $21 = 7 \cdot 9 - 7 \cdot 6$; so we choose $x_0 = -7$ and $y_0 = 7$ to be a particular solution.

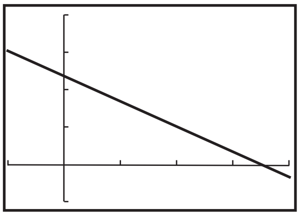
Hence, the general solution is

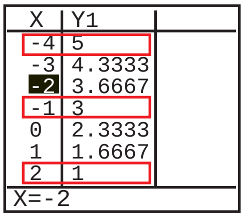
$x = -7 + 3t$ and $y = 7 - 2t$.

(Notice that if we substitute $t = 1$, we get the solution in **1** (Example 2) and if we substitute $t = 3$, we get the solution in **2**.)

Notice that the three solutions can be consolidated, and eventually they yield the same set of numbers.

Note: Since the solution for the equation, *if it exists*, is always an integer, and since this type deals with two variables, but gives only one equation, it is natural to expect an infinite number of solutions. One way to look at the solutions is to get an idea of the solution through a graph of the equation. As you know, $ax + by = c$ is the equation of a straight line. The line consists of all ordered pairs (x, y) that satisfy the equation. Not all of them are integers of course. By graphing and producing a table, you may be able to find a particular solution, after which the general solution is very simple.





```
 X  | Y1
 -4 | 5
 -3 | 4.3333
 -2 | 3.6667
 -1 | 3
 0  | 2.3333
 1  | 1.6667
 2  | 1
X=-2
```

Notice how you can find three particular solutions: $(-1, 3)$, $(-4, 5)$, and $(2, 1)$.

Example 3

Solve $12x + 25y = 331$.

Solution

We will use two methods to demonstrate their application and leave the trial and error for you to investigate. You might find the task easier if you set up a spreadsheet.

a) Euclidean algorithm:

We notice that 12 and 25 are relatively prime.

$25 = 2 \cdot 12 + 1$, and so $1 = 1 \cdot 25 - 2 \cdot 12$

$331 = 331 \cdot 25 - 662 \cdot 12$ Multiply both sides by 331.

$x_0 = -662$ and $y_0 = 331$ A particular solution.

$x = -662 + 25t$ and $y = 331 - 12t$ The general solution to this equation.

b) Linear congruence:

$$12x + 25y = 331 \Rightarrow 12x \equiv 331 \pmod{25}$$
$$\Rightarrow 12x \equiv (325 + 6) \pmod{25} \Rightarrow 12x \equiv 6 \pmod{25}$$
$$\Rightarrow 2x \equiv 1 \pmod{25} \qquad \text{[Corollary 1]}$$

Here we find $x_0 = 13$ and therefore $y_0 = 7$ to be a particular solution.

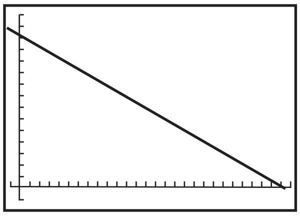
The general solution would be

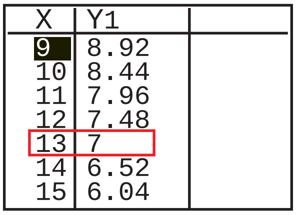
$x = 13 + 25t$ and $y = 7 - 12t$.

Notice that if we substitute $t = -27$, we get

$x = -662$ and $y = 331$.

Using a GDC here too helps you recognize $(13, 7)$ as a solution.





X	Y1
9	8.92
10	8.44
11	7.96
12	7.48
13	7
14	6.52
15	6.04

Note: Sometimes a constraint is added to the request of finding a solution. For instance, in Example 3, a condition is imposed that our solution must be positive. Luckily enough b) gave us a positive answer, but a) did not. However, to guarantee that it happens, we solve a system of two inequations.

$$-662 + 25t > 0 \text{ and } 331 - 12t > 0$$

$$\left.\begin{aligned} -662 + 25t > \Rightarrow t > \frac{662}{25} = 26\frac{12}{25} \\ 331 - 12t > 0 \Rightarrow t < \frac{331}{12} = 27\frac{7}{12} \end{aligned}\right\} \Rightarrow 26\frac{12}{25} < t < 27\frac{7}{12}$$

$t = 27$ is the only possibility, and hence $x = 13$ and $y = 7$.

Example 4

Solve the equation $6x + 51y = 22$.

Solution

Since $\gcd(6, 51) = 3 \nmid 22$, there is no solution.

Summary of the process of solving *ax* + *by* = *c*

Step 1: Calculate $g = \gcd(a, b)$.

Step 2: Check if $g \mid c$. If it is not true, then there are no solutions, so stop here. If $g \mid c$, then write $c = gk$.

Step 3: If $g \mid c$, then find integers u and v such that $au + bv = g$. Then $x_0 = uk$ and $y_0 = vk$ is a particular solution of $ax + by = c$. Use one of the three methods we discussed.

Step 4: Write the general solution $x = x_0 + \left(\frac{b}{g}\right)t$ and $y = y_0 - \left(\frac{a}{g}\right)t$ for all $t \in \mathbb{Z}$.

Example 5

Find the number of \$20 bills and the number of \$50 bills which will together make \$510.

Solution

The problem is equivalent to the Diophantine equation $20x + 50y = 510$, where x is the required number of \$20 bills and y is the required number of \$50 bills.

$\gcd(20, 50) = 10$, and $10 \mid 510$. So, $510 = 10 \cdot 51$.

$10 = 20 \cdot (-2) + 50 \cdot 1$ Using any of three methods discussed.

This implies that $10 \cdot 51 = 20 \cdot (-2 \cdot 51) + 50 \cdot 51$,
i.e. $510 = 20\,(-102) + 50 \cdot 51$

Thus, $x_0 = -102$ and $y_0 = 51$ is a particular solution.

The general solution of the Diophantine equation is

$$x = -102 + \left(\frac{50}{10}\right)t = -102 + 5t \text{ and } y_0 = 51 - \left(\frac{20}{10}\right)t = 51 - 2t.$$

We want to choose values of t so that x and y are positive.

Hence, we need $-102 + 5t \geqslant 0$ and $51 - 2t \geqslant 0$, which implies that

$$\frac{102}{5} = 20\frac{2}{5} \leqslant t \leqslant \frac{51}{2} = 25\frac{1}{2}.$$

Hence, only $t = 21, 22, 23, 24$, and 25 can be used.

Substituting these values of t into the expressions for x and y, we get the number of \$20 and \$50 bills which will make \$510 to be:
$(x, y) = (3, 9), (8, 7), (13, 5), (18, 3)$ and $(23, 1)$.

Example 6

a) Find the general solution of the linear Diophantine equation $172x + 20y = 1000$.

b) Find the positive integer solutions of this equation.

Solution

a) $\gcd(172, 20) = 4$ Use any method of your choice.

$172x \equiv 1000 \pmod{20} \Rightarrow 43x \equiv 250 \pmod 5 \Rightarrow (40 + 3)x \equiv 250 \pmod 5$

$\Rightarrow 3x \equiv 0 \pmod 5$; thus $x = 0$ (or any of its residue class mod 5)

A particular solution is $x_0 = 0$ and $y_0 = 50$. Substitute $x = 0$ into the equation.

A general solution is $x = 0 + 5t$ and $y = 50 - 43t$. $20 \div 4$ and $172 \div 4$.

If you choose to use the Euclidean algorithm (presented here for comparison purposes), then

$$\left.\begin{aligned} 172 &= 8(20)+12 \\ 20 &= 1(12)+8 \\ 12 &= 1(8)+4 \\ 8 &= 2(4)+0 \end{aligned}\right\}\gcd(172, 20) = 4$$

Now, we express $4 = 172u + 20v$.
From the calculations for finding gcd(172, 20), we have

$$4 = 12 - 8 = 12 - (20 - 12) = 2(12) - 20 = 2(172 - 8(20)) - 20 = 2(172) + (-17)20.$$

Hence, $u = 2$ and $v = -17$.

Since $\dfrac{1000}{4} = 250$, the particular solution (x_0, y_0) is given by

$x_0 = 2(250) = 5000$ and $y_0 = (-17)(250) = -4250$.

Hence, the general solution is given by

$$x = 500 + \left(\frac{20}{4}\right)t = 500 + 5t \text{ and } y = -4250 - \left(\frac{172}{4}\right)t = -4250 - 43t,\ t \in \mathbb{Z}.$$

Notice how there is an apparent difference in the solutions between the two methods. However, we leave it as an exercise for you to consolidate the two answers by the appropriate choice of values of t.

b) To find the positive integer solutions, t must be chosen so that $5t > 0$ and $50 - 43t > 0$.

This implies $0 < t < \dfrac{50}{43}$.

Thus, $t = 1$ is the only possible value, and so we have $x = 5$ and $y = 7$.

In the Euclidean method case:

$5t + 500 > 0$ and $-4250 - 43t > 0$.

This implies that $-100 < t < -98\dfrac{36}{43}$.

Hence, we take $t = -99$.

Thus, $x - 500 + 5(-99) = 5$ and $y = -4250 - 43(-99) = 7$, which is the same result as before.

Example 7

Solve the Diophantine equation $1492x + 1066y = -4$.

Solution

$\gcd(1492, 1066) = 2$. Since $2 \mid -4$, the Diophantine equation has infinitely many solutions.

Now, $2 = (-5)1492 + 7(1066)$. Since $-4 = 2 \cdot (-2)$, the particular solution (x_0, y_0) is given by $x_0 = (-5)(-2) = -10$ and $y_0 = 7(-2) = -14$.
Using (x_0, y_0), the general solution has the form

$$x = 10 + \left(\frac{1066}{2}\right)t = 10 + 533t \text{ and } y = -14 - \left(\frac{1492}{2}\right)t = 14 - 746t,\ t \in \mathbb{Z}.$$

Example 8

Find the smallest positive integer n such that the Diophantine equation $533x + 299y = 10\,000 + n$ has a solution, and for this value of n find the positive integer solutions.

Solution

gcd(533, 299) = 13. In order for the linear Diophantine equation $533x + 299 = 10\,000 + n$ to have a solution, $10\,000 + n$ must be divisible by 13. Thus, $1000 + n \equiv 0 \pmod{13} \Rightarrow 9997 + 3 + n \equiv 0 \pmod{13}$ $\Rightarrow 3 + n \equiv 0 \pmod{13} \Rightarrow n = 10$.

Hence, the equation to be solved is

$533x + 299y = 10\,010 \Rightarrow 533x \equiv 10\,010 \pmod{299} \Rightarrow 41x \equiv 770 \pmod{23}$(Why?)

$\Rightarrow 18x \equiv 11 \pmod{23} \Rightarrow x = 7$, since $18 \cdot 7 = 126 - 11 = 115 = 5 \cdot 23$.

By back substitution into the equation, we have $y = 21$.

Notice the difference if we were to use the Euclidean algorithm method.

Knowing gcd(533, 299) = 13, we need to find u and v such that $13 = 533u + 277v$.

We can find that $u = 9$ and $v = 16$.

A particular solution is given by $x_0 = \left(\frac{10\,010}{13}\right)9 = 6930$ and

$y_0 = \left(\frac{10\,010}{13}\right)(-16) = -12\,320$.

Hence, the general solution is given by $x = 6930 + 23t$ and $y = -12\,320 - 41t$.

For positive integer solutions both x and y are positive, so

$6930 + 23t > 0$ or $t > \frac{-6930}{23}$ and $-12\,320 - 41t > 0$ which implies

$-12\,320 > 41t$ or $41t < -12\,320$ or $t < \frac{-12\,320}{41}$.

Hence, $-301.304 < t < -300.975$.

On taking $t = -301$, $x = 6930 + 32(-301) = 7$ and
$y = -12\,320 - 41(-301) = 21$.

Exercise 2.2

1 Determine which of the following Diophantine equations have a solution.

a $51x + 6y = 22$

b $14x + 33y = 115$

c $35x + 14y = 93$

2 Determine the general solution of the following Diophantine equations.

a $13x - 7y = 21$

b $221x + 35y = 11$

c $1485x + 1745y = 15$

3 Determine the positive integer solutions of the linear Diophantine equations.

a $5x - 11y = 29$

b $32x + 55y = 71$

c $62x + 11y = 788$

4 A grocer orders apples and oranges for \$16.78. If apples cost him 25 cents each and oranges cost him 18 cents each, how many of each type of fruit did he order?

5 Kate spent €100.64 on posters. Some of the posters cost €4.98 each and some €5.98. How many did she buy?

6 A person has \$4.55 in change composed of dimes and quarters. Set up the linear Diophantine equation and find the maximum and the minimum number of coins that the person can have.

7 David collected \$75 at the market by selling chickens and geese. He got \$4 for each chicken and \$7 for each goose. How many of each did he sell?

8 A farmer purchased one hundred head of livestock for a total cost of \$4000. Calves, lambs, and piglets cost \$120, \$50, and \$25 each, respectively. If the farmer bought at least one animal of each type, how many of each type did he buy?

9 Roberto bought three dozen oranges and two dozen apples. He paid €8.04 in total. Each orange costs more than 10 cents, while an apple costs more than 15 cents. How much did he pay for the oranges?

10 Marco has a small grocery shop. He buys tomatoes from farmer Antonio in large boxes and then repackages them in smaller boxes. Marco bought 11 large boxes and sold 39 small boxes. A small box contains less than 12 tomatoes. At the end of the day, Marco was left with 19 tomatoes. How many tomatoes does each large box contain?

11 Farmer Josip owes farmer Tim €10. Neither of the two has any cash, but Josip has 14 sheep valued at €185 each. He suggests paying Tim in sheep with Tim paying the change in pigs, which are valued at €110 each. Is this possible? If yes, how; if not, why not?

In questions 12–34, either find all integral (integer) solutions to the given equation or show that it has none.

12 $3x + 2y = 1$

13 $3x - 2y = 1$

14 $17x + 14y = 4$

15 $33x - 12y = 9$

16 $91x + 221y = 15$

17 $361x + 2109y = 1000$

18 $401x + 503y = 20$

19 $26x + 14y = 2$

20 $27x + 15y = 3$

21 $217x + 341y = 62$

22 $117x + 247y = 39$

23 $2x + 3y = 50;\ x, y > 0$

24 $3x + 4y = 60;\ x, y > 0$

25 $4x + 6y = 60;\ x, y > 0$

26 $6x + 9y = 91;\ x, y > 0$

27 $4x + 6y = 25$

28 $3x + 5y = 50\,001$

29 $6x + 9y = 60\,001$

30 $21x - 14y = 10\,000$

31 $42y - 12x = 366$

32 $66x + 51y = 300$

33 $55x + 200y = -100$

34 $121x + 561y = 13\,200;\ x, y > 0$

35 $a, b \in \mathbb{Z}^+$, show that there exist $x, y \in \mathbb{Z}$ such that $\dfrac{1}{\text{lcm}(a, b)} = \dfrac{x}{a} + \dfrac{y}{b}$.

36 Show that if a and b are relatively prime, and $c \neq 0$, then $\gcd(ac, bc) \mid c$.

2.3 Linear congruences

A congruence of the form

$ax \equiv b \pmod{m}$, where x is an unknown integer,

is called a linear congruence in one variable. As you have seen in the previous section, the study of such congruences is similar to the work with linear Diophantine equations in two variables. In fact, we used linear congruences to solve some of these equations.

Example 9

Find a solution to linear congruence $3x \equiv 4 \pmod{7}$.

Solution

For now, let us try and find the solution by trial and error and some knowledge of congruence.

One way to approach this is to resort to the definition of congruence:

$3x \equiv 4 \pmod{7}$ implies that $7 \mid (3x - 4)$.

In other words, $3x - 4 = 7k$ for some integer k.

This means that $3x - 4$ should be equal to one of the multiples of 7 $\{0, \pm 7, \pm 14, \pm 21, \pm 28, \ldots\}$.

When $x = 6$, $3x - 4 = 14$ and we have a solution. If we let $x = -1$, $3x - 4 = -7$ and we have another solution. However, you know that $6 \equiv -1 \pmod{7}$. So, it appears that all members of the residue class of 6 will be solutions too.

If you recall some of the rules we learned earlier, you can solve the problem without guessing!

Multiply the equation by 5. This gives you $15x \equiv 20 \pmod 7$.

This, in turn, means $(14 + 1)x \equiv (14 + 6) \pmod 7$, which simplifies to $x \equiv 6 \pmod 7$.

From the previous discussion, you notice that if we have $x = x_0$ as a solution to the congruence $ax \equiv b \pmod m$, and if $x_1 \equiv x_0 \pmod m$, then $ax_1 \equiv ax_0 \equiv b \pmod m$, and hence x_1 is also a solution. Thus, if one member of a residue class modulo m is a solution, then the entire class is made up of solutions. The question remains: How many different '*incongruent*' solutions does the congruence have?

The following theorem tells you when to expect a solution and how many incongruent solutions modulo m the congruence has.

Theorem 11

Let $a, b, m \in \mathbb{Z}$, with $m > 0$ and $\gcd(a, m) = g$. If $g \nmid b$, then $ax \equiv b \pmod m$ has no solutions. If $g | b$, then $ax \equiv b \pmod m$ has exactly g '*incongruent*' solutions modulo m.

Proof

$ax \equiv b \pmod m$ can be written as $ax - b = my$, where y is an integer. (Definition of congruence.)

The last equation can be rewritten as $ax - my = b$. This is a Diophantine equation!

The Diophantine equation, by Theorem 9 and 10 combined, has no solution if $g \nmid b$, while it has infinitely many solutions if $g | b$. These solutions are given by

$$x = x_0 + \frac{m}{g}t \text{ and } y = y_0 - \frac{-a}{g}t = y_0 + \frac{a}{g}t,$$

where $x = x_0$ and $y = y_0$ is a particular solution of the equation. The values of x given above,

$$x = x_0 + \frac{m}{g}t,$$

are the solutions to the linear congruence. There are infinitely many of them, but they are congruent, as you notice from the equation.

To find out how many incongruent solutions there are, let us first look at the conditions under which two solutions like $x_1 = x_0 + \frac{m}{g}t$ and $x_2 = x_0 + \frac{m}{g}s$ are congruent modulo m.

Now, $x_1 \equiv x_2 \pmod{m} \Rightarrow x_0 + \frac{m}{g}t \equiv x_0 + \frac{m}{g}s \pmod{m} \Rightarrow \frac{m}{g}t \equiv \frac{m}{g}s \pmod{m}$. (1)

Now, $\gcd\left(m, \frac{m}{g}\right) = \frac{m}{g}$ since $\left(\frac{m}{g}\right) \mid m$, so by Theorem 6 we now have

$t \equiv s \pmod{g}$ [we divide (1) by (m/g)].

Therefore, to have a complete set of incongruent solutions $x = x_0 + \frac{m}{g}t$, we need to consider all residue classes modulo g. This proves the theorem.

Note: When $\gcd(a, m) = 1$, there is exactly one unique solution modulo m.

Example 10

Solve each of the following linear congruences.

a) $14x \equiv 13 \pmod{21}$

b) $9x \equiv 15 \pmod{21}$

c) $8x \equiv 7 \pmod{13}$

d) $9x \equiv 12 \pmod{15}$

e) $7x \equiv 1 \pmod{31}$

f) $7x \equiv 22 \pmod{31}$

g) $18x \equiv 30 \pmod{42}$

Solution

a) $\gcd(14, 21) = 7$, and $7 \nmid 13$, so the equation has no solution.

b) $\gcd(9, 21) = 3$, and $3 \mid 15$, so we have three incongruent solutions modulo 21.

Theorem 6 helps us rewrite the equation as

$3x \equiv 5 \pmod{7} \Rightarrow 3x \equiv (5 + 7) \pmod{7} \Rightarrow 3x \equiv 12 \pmod{7} \Rightarrow x \equiv 4 \pmod{7}$

This implies that the solutions to the equation are of the form

$x = x_0 + \frac{m}{g}t = 4 + 7t$, with $t = 0, 1$, and 2.

Thus, the solutions are: $x \equiv 4, 11, 18 \pmod{21}$.

c) $\gcd(8, 13) = 1$, so we have one solution modulo 13.

$8x \equiv 7 \pmod{13} \Rightarrow 8x \equiv (7 + 13) \pmod{13} \Rightarrow 8x \equiv 20 \pmod{13}$
$\Rightarrow 2x \equiv 5 \pmod{13}$, and again

$2x \equiv 5 \pmod{13} \Rightarrow 2x \equiv 18 \pmod{13} \Rightarrow x \equiv 9 \pmod{13}$, which is the solution.

d) $\gcd(9, 15) = 3$, and $3 \mid 12$, so we have exactly three incongruent solutions modulo 15.

Rewrite the equation: $3x \equiv 4 \pmod{5}$ Divided by 3.

$3x \equiv (4 + 20) \pmod 5 \Rightarrow 3x \equiv 24 \pmod 5 \Rightarrow x \equiv 8 \pmod 5$

$x = 8 + 5t$, with $t = 0, 1$, and 2.

Therefore, the solutions are given by

$x \equiv 8 \pmod{15}$, $x \equiv 13 \pmod{15}$, and $x \equiv 18 \equiv 3 \pmod{15}$.

e) $\gcd(7, 31) = 1$, so there is exactly one solution modulo 31.

$7x \equiv 1 \pmod{31}$ Multiply by 9.

$63x \equiv 9 \pmod{31} \Rightarrow (62x + x) \equiv 9 \pmod{31} \Rightarrow x \equiv 9 \pmod{31}$

Hint: $x \equiv 9 \pmod{31}$ is called an inverse of 7 modulo 31.

f) $\gcd(7, 31) = 1$, so there is exactly one solution.

$7x \equiv 22 \pmod{31}$ Multiply by 9.

$63x \equiv 198 \pmod{31} \Rightarrow x \equiv 12 \pmod{31}$ (Why?)

g) $\gcd(18, 42) = 6$, so we have six incongruent solutions modulo 42.

$18x \equiv 30 \pmod{42} \Rightarrow 3x \equiv 5 \pmod 7$

$\Rightarrow 3x \equiv 12 \pmod 7 \Rightarrow x \equiv 4 \pmod 7$

$x = 4 + 7t$, with $t = 0, 1, 2, 3, 4$, and 5.

Therefore, the solutions are given by

$x \equiv 4, 11, 18, 25, 32$, and $39 \pmod{42}$.

The Chinese remainder theorem

An old Chinese puzzle poses a question as follows:

Find a number that leaves a remainder of 1 when divided by 3, a remainder of 2 when divided by 5, and a remainder of 3 when divided by 7. Interpreting this puzzle using congruences, we get the following system:

$x \equiv 1 \pmod 3$

$x \equiv 2 \pmod 5$

$x \equiv 3 \pmod 7$

Even though systems with more than one variable can be solved, this section focuses on systems of simultaneous congruences with one variable but different moduli, like the one above.

The following theorem will provide us with a method for finding all solutions of simultaneous congruences similar to the given example.

Theorem 12: The Chinese remainder theorem

Let $m_1, m_2, \ldots, m_r$, be positive integers which are pairwise relatively prime,

i.e. $\gcd(m_i, m_j) = 1, \forall i \neq j, i, j = 1, 2 \ldots, r$.

The system of congruences

$$x \equiv a_1 \pmod{m_1}$$
$$x \equiv a_2 \pmod{m_2}$$
$$\vdots$$
$$x \equiv a_r \pmod{m_r}$$

has a unique solution modulo $M = m_1 m_2 \ldots m_r$.

Proof (Optional)

Let $M_k = \dfrac{M}{m_k} = m_1 m_2 \ldots m_{k-1} m_{k+1} \ldots m_r$.

In words, M_k is the product of all the moduli m_i, with the modulus m_k omitted.

By hypothesis, all the m_i are relatively prime in pairs, so the gcd$(M_k, m_k) = 1$. According to the previous section's theorems, it is possible to solve the congruence $M_k x \equiv 1 \pmod{m_k}$. Call that *unique* solution x_k. That is, $M_k x_k \equiv 1 \pmod{m_k}$.

Our aim now is to prove that the integer

$$x = a_1 M_1 x_1 + a_2 M_2 x_2 + \ldots + a_r M_r x_r$$

is a simultaneous solution of the given system.

To show this, we need to show that $x \equiv a_k \pmod{m_k}$ for $k = 1, 2, \ldots, r$.

Since $m_k \mid M_j$ whenever $j \neq k$, $M_j \equiv 0 \pmod{m_k}$. Thus, in the sum for x, all terms except the kth term are congruent to $0 \pmod{m_k}$.

Hence, $x \equiv a_k M_k x_k \pmod{m_k}$, with $M_k x_k \equiv 1 \pmod{m_k}$ implying that $x \equiv a_k \pmod{m_k}$.

This proves the existence of the solution.

Now, let y be another solution to the system.

Then for each k, $y \equiv x \equiv a_k \pmod{m_k}$, which means that $m_k \mid (x - y)$.

Then using Theorem 8, we see that $M = m_1 m_2 \ldots m_r \mid (x - y)$.

Therefore, $y \equiv x \pmod{M}$.

Example 11

Solve the system:

$$x \equiv 1 \pmod 3$$

$$x \equiv 2 \pmod 5$$

$$x \equiv 3 \pmod 7$$

Solution

$M = 3 \cdot 5 \cdot 7 = 105$

$M_1 = \frac{105}{3} = 35;\ M_2 = \frac{105}{5} = 21;\ M_3 = \frac{105}{7} = 15$

Now, to determine x_1, we solve $35x_1 \equiv 1 \pmod 3$, which simplifies to $x_1 \equiv 2 \pmod 3$.

For x_2, $21x_2 \equiv 1 \pmod 5$, we have $x_2 \equiv 1 \pmod 5$, and finally

$15x_3 \equiv 1 \pmod 7$, which gives $x_3 \equiv 1 \pmod 7$.

Therefore, our solution x is

$$x \equiv 1 \cdot 35 \cdot 2 + 2 \cdot 21 \cdot 1 + 3 \cdot 15 \cdot 1 \equiv 157 \equiv 52 \pmod{105}.$$

Checking back in the original system, you see that this solution satisfies the system:

$52 \equiv 1 \pmod 3$, since $51 = 3 \cdot 17$; $52 \equiv 2 \pmod 5$, since $50 = 10 \cdot 5$; and $52 \equiv 3 \pmod 7$, since $49 = 7 \cdot 7$.

Example 12

Solve the system:

$$x \equiv 2 \pmod 3$$

$$x \equiv 5 \pmod 4$$

$$x \equiv -3 \pmod 7$$

Solution

3, 4, and 7 are pairwise relatively prime.

$M = 3 \cdot 4 \cdot 7 = 84$

$M_1 = \frac{84}{3} = 28;\ M_2 = \frac{84}{4} = 21; M_3 = \frac{84}{7} = 12$

Now, to determine x_1, we solve $28x_1 \equiv 1 \pmod 3$, which simplifies to $x_1 \equiv 1 \pmod 3$.

For x_2, $21x_2 \equiv 1 \pmod 4$, we have $x_2 \equiv 1 \pmod 4$, and

$12x_3 \equiv 1 \pmod 7$, which gives $x_3 \equiv 3 \pmod 7$.

Therefore, our solution x is

$$x \equiv 1 \cdot 28 \cdot 2 + 1 \cdot 21 \cdot 5 + 3 \cdot 12 \cdot (-3) \equiv 53 \pmod{84}.$$

Again, checking back in the original system, you see that this solution satisfies the system:

$53 \equiv 2 \pmod 3$, since $51 = 17 \cdot 3$; $53 \equiv 5 \pmod 4$, since $48 = 12 \cdot 4$; and $53 \equiv -3 \pmod 7$, since $56 = 8 \cdot 7$.

The following example offers a slight variation on the same theme.

Example 13

Solve the linear congruence

$3x \equiv 11 \pmod{2275}$.

Solution

Since $\gcd(3, 2275) = 1$, the linear congruence has a unique solution modulo 2275.

We will approach the problem differently because of the size of the modulus.

Since $2275 = 5^2 \cdot 7 \cdot 13$, the original congruence may be replaced by the system:

$3x \equiv 11 \pmod{25}$

$3x \equiv 11 \pmod{7}$

$3x \equiv 11 \pmod{13}$

$M = 25 \cdot 7 \cdot 13 = 2275$

$M_1 = \dfrac{2275}{25} = 91$; $M_2 = \dfrac{2275}{7} = 325$; $M_3 = \dfrac{2275}{13} = 175$

Now, to determine x_1, we solve $91x_1 \equiv 16x_1 \equiv 1 \pmod{25}$, which simplifies to $x_1 \equiv 11 \pmod{25}$. Verify.

For x_2, $325x_2 \equiv 3x_2 \equiv 1 \pmod{7}$, we have $x_2 \equiv 5 \pmod{7}$, and

$175x_3 \equiv 6x_3 \equiv 1 \pmod{13}$, which gives $x_3 \equiv 11 \pmod{13}$.

We still need to determine the particular solutions, a_is, since the linear congruences are not in the standard $x \equiv a_i \pmod{m_i}$ form.

$3x \equiv 11 \pmod{25}$ will give $a_1 = 12$.

$3x \equiv 11 \pmod{7}$ will give $a_2 = 6$.

$3x \equiv 11 \pmod{13}$ will give $a_3 = 8$.

Thus, the solution to the original congruence is now given by

$$x \equiv 12 \cdot 91 \cdot 11 + 6 \cdot 325 \cdot 5 + 8 \cdot 175 \cdot 11 \equiv 37\,162 \equiv 762 \pmod{2275}.$$

What we observe here is that, even though we had to solve six congruences, the moduli of these congruences are relatively small as compared to 2275 and could mostly be solved by mere inspection. This method offers a way to perform computer arithmetic with large integers.

Alternative method of solution

There is also a method similar to solving systems of equations by substitution that you are familiar with from early years.

This is an iterative method where we find a general solution for the variable in one congruence and substitute that value into another congruence, until we finish. We will demonstrate this method with an example.

Example 14

Solve the system:

$x \equiv 1 \pmod 5$ ....(1)

$x \equiv 2 \pmod 6$ ....(2)

$x \equiv 3 \pmod 7$ ....(3)

Solution

Rewrite (1) using the definition of congruence, i.e. $x - 1 = 5t$ with $t \in \mathbb{Z}$, which leads to $x = 5t + 1$. Now, for this solution to serve as a solution to the system, it must satisfy the second congruence:

$5t + 1 \equiv 2 \pmod 6$, i.e. $5t \equiv 1 \pmod 6$.

This can be solved to give $t \equiv 5 \pmod 6$.

So, $t = 5 + 6k$, where $k \in \mathbb{Z}$, and hence $x = 5t + 1 = 5(5 + 6k) + 1 = 30k + 26$.

This x in turn must satisfy the third congruence, and hence

$30k + 26 \equiv 3 \pmod 7$, i.e. $2k + 5 \equiv 3 \pmod 7 \Rightarrow 2k \equiv -2 \pmod 7$
$\Rightarrow k \equiv - 1 \pmod 7$, and thus $k \equiv 6 \pmod 7$.

Hence, $k = 6 + 7u$, where $u \in \mathbb{Z}$. Finally,

$x = 30k + 26 = 30(6 + 7u) + 26 = 210u + 206$, which is equivalent to saying

$x \equiv 206 \pmod{210}$, which is the simultaneous solution.

This method demonstrates that a system of simultaneous congruences can be solved by successively solving linear congruences. This can be done even if the moduli are not pairwise relatively prime.

Example 15

Solve the linear congruence

$17x \equiv 9 \pmod{276}$.

Solution

Observe that $276 = 3 \cdot 4 \cdot 23$, and hence the congruence is equivalent to the following system:

$17x \equiv 9 \pmod 3 \Rightarrow x \equiv 0 \pmod 3$ ....(1)

$17x \equiv 9 \pmod 4 \Rightarrow x \equiv 1 \pmod 4$ ....(2)

$17x \equiv 9 \pmod{23} \Rightarrow 17x \equiv 9 \pmod{23}$ ....(3)

We will approach this problem using the iterative method.

From (1) we have $x = 3k$, where $k \in \mathbb{Z}$. Now, we substitute this into (2):

$3k \equiv 1 \pmod 4 \Rightarrow 9k \equiv 3 \pmod 4 \Rightarrow k \equiv 3 \pmod 4$

Thus, $k = 3 + 4i$, with $i \in \mathbb{Z}$, and hence $x = 3k = 3(3 + 4i) = 9 + 12i$.

From (3), we have
$17x \equiv 9 \pmod{23} \Rightarrow 17(9 + 12i) \equiv 9 \pmod{23} \Rightarrow 153 + 204i \equiv 9 \pmod{23}$
$\Rightarrow 204i \equiv -144 \pmod{23} \Rightarrow 3i \equiv 6 \pmod{23} \Rightarrow i \equiv 2 \pmod{23}$, and so
$i = 2 + 23t$.

Therefore, $x = 9 + 12i = 9 + 12(2 + 23t) = 33 + 276t$, and finally

$x \equiv 33 \pmod{276}$ is the solution to the system of congruences, and hence a solution to $17x \equiv 9 \pmod{276}$.

Systems of linear congruences

We will consider systems of two congruences involving two unknowns. The modulus will also be the same in both congruences. Of course, more congruences and more unknowns are possible, but they go beyond the scope of this publication.

The process we follow in trying to solve such systems is equivalent to what we do in solving systems of simultaneous equations in algebra. We will explain the method through the use of an example.

Example 16

Find the solution to:

$$3x + 4y \equiv 5 \pmod{13}$$
$$2x + 5y \equiv 7 \pmod{13}$$

Solution

Multiply the first congruence by 5 and the second by 4 to obtain

$$15x + 20y \equiv 25 \pmod{13}$$
$$8x + 20y \equiv 28 \pmod{13}$$

By subtraction, we have

$7x \equiv -3 \pmod{13}$, which will give us a solution for x.

$x \equiv 7 \pmod{13}$

We leave the verification as an exercise.

If we multiply the first congruence by 2 and the second by 3, we have

$$6x + 8y \equiv 10 \pmod{13}$$
$$6x + 15y \equiv 21 \pmod{13}$$

By subtraction, we have

$7y \equiv 11 \pmod{13}$, which in turn will yield

$y \equiv 9 \pmod{13}$.

The solution to the system is therefore

$(x \equiv 7 \pmod{13}, y \equiv 9 \pmod{13})$.

Theorem 13 (Optional)

Let $a, b, c, d, e, f, m \in \mathbb{Z}$ with $m > 0$. The system of congruences

$$ax + by \equiv e \pmod{m}$$

$$cx + dy \equiv f \pmod{m}$$

will have a unique solution if $\gcd(ad - bc, m) = 1$.

Exercise 2.3

In questions 1–13, find all solutions of each of the linear congruences.

1 $5x \equiv 2 \pmod{7}$

2 $6x \equiv 3 \pmod{9}$

3 $17x \equiv 30 \pmod{40}$

4 $5x \equiv 9 \pmod{49}$

5 $107x \equiv 333 \pmod{888}$

6 $490x \equiv 750 \pmod{800}$

7 $2x \equiv 3 \pmod{7}$

8 $12x \equiv 6 \pmod{18}$

9 $19x \equiv 16 \pmod{24}$

10 $15x \equiv 9 \pmod{25}$

11 $128x \equiv 833 \pmod{1001}$

12 $14x \equiv 5 \pmod{45}$

13 $3x \equiv 2 \pmod{78}$

14 For what integer values of k, where $k \in [0, 36[$, does the congruence $16x \equiv k \pmod{36}$ have solutions? When it has solutions, how many incongruent solutions are there?

In questions 15–19, attempt to use both methods, the Chinese remainder and the iterative methods, in solving each system.

15 Solve: $x \equiv 2 \pmod{3}$, $x \equiv 3 \pmod{4}$

16 Solve: $x \equiv 7 \pmod{9}$, $x \equiv 13 \pmod{23}$, $x \equiv 1 \pmod{2}$

17 Solve: $2x \equiv 3 \pmod{5}$, $4x \equiv 3 \pmod{7}$

18 Solve: $6x \equiv 8 \pmod{10}$, $15x \equiv 30 \pmod{55}$

19 Solve: $x \equiv 0 \pmod{2}$, $x \equiv 0 \pmod{3}$, $x \equiv 1 \pmod{5}$, $x \equiv 6 \pmod{7}$

20 Find an integer that leaves a remainder of 9 when divided by 10 or 11, but is divisible by 13.

21 Find the solution of

$x + 2y \equiv 1 \pmod 5$

$2x + y \equiv 1 \pmod 5$

22 Find the solution of

$x + 3y \equiv 1 \pmod 5$

$3x + 4y \equiv 2 \pmod 5$

23 Find the solution of

$4x + y \equiv 2 \pmod 5$

$2x + 3y \equiv 1 \pmod 5$

24 Find the solution of

$2x + 3y \equiv 5 \pmod 7$

$x + 5y \equiv 6 \pmod 7$

2.4 Integer representations and operations

We usually use the decimal notation to represent integers. It is a positional numeral system with base 10.

In this section, we shall show that any positive integer can be uniquely represented in a base b, where b is a positive integer. When $b = 2$, the representation is called a **binary representation**; when $b = 16$, the representation is called the **hexadecimal expansion**. We will describe a method of finding the base b representation of an integer, and describe a procedure to carry out integer arithmetic.

Use of bases other than ten is known from the history of mathematics (see Howard Eves, *An Introduction to the History of Mathematics*, 6th edition (Thomson Brooks/Cole, 1990) pages 19–27). Between 2000 to 500 BCE, the Babylonians evolved a sexagesimal system (base 60). The Mayan numerical system used base 20, but a positional system of its own. Some African tribes used base 5, and base 2 appears in Chinese mathematics. Some of the Egyptian calculations used base 7.

Before we discuss representation of an integer in an arbitrary base, we examine our familiar decimal system and build the rest of our work on that.

Decimal representation of integers

1765 in base 10 is written as

$$1765 = 1000 + 700 + 60 + 5 = 1 \cdot 10^3 + 7 \cdot 10^2 + 6 \cdot 10 + 5 \cdot 10^0.$$

In general, if n is a natural number whose decimal representation is

$a_r a_{r-1} \dots a_1 a_0$, where $0 \leqslant a_k \leqslant 9$, $k = 0, 1, \dots, r$, then

$$n = a_r \cdot 10^r + a_{r-1} \cdot 10^{r-1} + \dots + a_1 \cdot 10^1 + a_0 \cdot 10^0 = \sum_{k=0}^{r} a_k 10^k.$$

Each a_k is called a decimal digit of n.

For another example, when we write 54 273, we mean

$$5 \cdot 10^4 + 4 \cdot 10^3 + 2 \cdot 10^2 + 7 \cdot 10 + 3.$$

Theorem 14

Let b be a positive integer with $b > 1$. Then every positive integer n can be written uniquely in the form

$$n = a_r \cdot b^r + a_{r-1} \cdot b^{r-1} + \ldots + a_1 \cdot b^1 + a_0 \cdot b^0 = \sum_{k=0}^{r} a_k b^k$$

where r and a_r are non-negative integers, with $a_r \leqslant b - 1$ for $k = 0, 1, 2, \ldots, r$, and the initial coefficient $a_r \neq 0$.

Proof (Optional)

We obtain an expression of the desired type by applying the division algorithm in sequence in the following manner:

Divide n by b to get $n = bq_0 + a_0$, $0 \leqslant a_0 < b$. If $q_0 \neq 0$, continue dividing by b to get:

$$q_0 = bq_1 + a_1, 0 \leqslant a_1 < b$$

We continue this process to obtain:

$$q_1 = bq_2 + a_2, 0 \leqslant a_2 < b$$

$$q_2 = bq_3 + a_3, 0 \leqslant a_3 < b$$

$$\vdots$$

$$q_{r-2} = bq_{r-1} + a_{r-1}, 0 \leqslant a_{r-1} < b$$

$$q_{r-1} = b \cdot 0 + a_r, 0 \leqslant a_r < b$$

The last step of the process is achieved when a quotient of 0 is obtained.

Now, as you recall from the division algorithm,

$$n > q_0 > q_1 > \ldots \geqslant 0.$$

Since this sequence is a decreasing sequence of non-negative integers which continues as long as its terms are positive, the last term is 0.

Now, combining what we obtained above, we get

$$n = bq_0 + a_0 = b(bq_1 + a_1) + a_0 = b(b(bq_2 + a_2) + a_1) + a_0$$
$$= b(b(b(bq_3 + a_3) + a_2) + a_1) + a_0, \text{ and so on.}$$

$$n = a_r \cdot b^r + a_{r-1} \cdot b^{r-1} + \ldots + a_1 \cdot b^1 + a_0 \cdot b^0$$

The uniqueness can also be proved, but will not be included here.

Note: When a number is expressed in a base different from decimal, it is a convention to write it as

$$(a_r a_{r-1} \ldots a_1 a_0)_b.$$

b is usually called the **base** or **radix** of the system or expansion. Recall that our system, with base 10, is called the decimal system. Base 2 is the binary system, base 8 is the octal system, and base 16 is the hexadecimal system (or *hex* for short).

Example 17

Follow the outlined process in Theorem 13 to find an expression for 1948 in base 2 and in base 5.

Solution

Base 2: $1948 = 2 \cdot 974 + 0$

$974 = 2 \cdot 487 + 0$

$487 = 2 \cdot 243 + 1$

$243 = 2 \cdot 121 + 1$

$121 = 2 \cdot 60 + 1$

$60 = 2 \cdot 30 + 0$

$30 = 2 \cdot 15 + 0$

$15 = 2 \cdot 7 + 1$

$7 = 2 \cdot 3 + 1$

$3 = 2 \cdot 1 + 1$

$1 = 2 \cdot 0 +1$

Therefore, the number in base 2 is $(11110011100)_2$.

Base 5: $1948 = 5 \cdot 389 + 3$

$389 = 5 \cdot 77 + 4$

$77 = 5 \cdot 15 + 2$

$15 = 5 \cdot 3 + 0$

$3 = 5 \cdot 0 + 3$

Therefore, the number in base 5 is $(30243)_5$.

To verify, we can change these numbers back into decimal by writing their base expansion:

$(11110011100)_2 = 1 \cdot 2^{10} + 1 \cdot 2^9 + 1 \cdot 2^8 + 1 \cdot 2^7 + 0 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 0 = 1948$

$(30243)_5 = 3 \cdot 5^4 + 0 \cdot 5^3 + 2 \cdot 5^2 + 4 \cdot 5^1 + 3 = 1948$

If systems use more digits than the decimal system, then they need more digits. No-one so far has invented new digits. Number theorists have been using letters to represent the extensions. For example, in base 16, the digits used are: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, and F. The letters A, B, C, D, E, and F are used to represent the digits that correspond to 10, 11, 12, 13, 14, and 15 (written in decimal notation). Next is an example to demonstrate the conversion between the two systems.

Example 18

a) Convert $(\text{A35B0F})_{16}$ to decimal notation.

b) Convert 38609905 to hex.

Solution

a) $$(\text{A35B0F})_{16} = \text{A} \cdot 16^5 + 3 \cdot 16^4 + 5 \cdot 16^3 + \text{B} \cdot 16^2 + 0 \cdot 16^1 + \text{F}$$
$$= 10 \cdot 16^5 + 3 \cdot 16^4 + 5 \cdot 16^3 + 11 \cdot 16^2 + 0 \cdot 16^1 + 15$$
$$= 10705679_{10}$$

b) $38609905 = 16 \cdot 2413119 + 1$

$2413119 = 16 \cdot 150819 + 15$ (F)

$150819 = 16 \cdot 9426 + 3$

$9426 = 16 \cdot 589 + 2$

$589 = 16 \cdot 36 + 13$ (D)

$36 = 16 \cdot 2 + 4$

$2 = 16 \cdot 0 + 2$

Therefore, $38609905_{10} = (\text{24D23F1})_{16}$.

Note: A simple conversion is possible between binary and hexadecimal notations. Each hex digit is written as a block of four binary digits according to the following table.

Hex digit	Binary	Hex digit	Binary	Hex digit	Binary
0	0000	6	0110	C	1100
1	0001	7	0111	D	1101
2	0010	8	1000	E	1110
3	0011	9	1001	F	1111
4	0100	A	1010		
5	0101	B	1011		

Example 19

a) Convert from hex to binary: $(\text{3FCB9})_{16}$

b) Convert from binary to hex: $(11011111011010100111 00)_2$

Solution

a) We simply replace each digit with its binary equivalent. However, for the first digit to the left, if it starts with zeros, then they should be omitted (similar to decimal representation when we are talking about 0213, we mean 213).

$$(\text{3FCB9})_{16} = (\cancel{00}111111111001011 1001)_2 = (111111110010111001)_2$$

b) We break the number into blocks of four, starting from the right. If the last block is missing digits, we add the initial zeros.

$(110111101101010011100)_2$

$= (\underline{000}110111101101010011100)_2$

$= (1BDA9C)_{16}$

Operations in different systems

The operations of addition, subtraction, and multiplication can be performed using similar methods to those you learned in the decimal system. We will explain a few operations using examples.

Example 20: Addition in base 4

Add: $(32032)_4 + (10203)_4$

Solution

Before you perform any operation, it is advisable that you set up a table for that operation. So, for addition in base 4, here is the addition table.

	1	2	3
1	2	3	10
2	3	10	11
3	10	11	12

1			1	1	
	3	2	0	3	2
	1	0	2	0	3
1	0	2	3	0	1

Starting at the right: $2 + 3 = 11$ Write 1, and retain 1.

$1 + 3 + 0 = 10$ Write 0, and retain 1; and so on.

Therefore, $(32032)_4 + (10203)_4 = (102301)_4$.

Example 21: Multiplication in base 6

Find the product $(352)_6 \times (524)_6$.

Solution

We set up a multiplication table to make our task simple.

	1	2	3	4	5
1	1	2	3	4	5
2	2	4	10	12	14
3	3	10	13	20	23
4	4	12	20	24	32
5	5	14	23	32	41

We arrange the numbers in a similar manner to decimal multiplication.

$$\begin{array}{rrrrrr}
 & & & 5 & 2 & 4 \\
 & & & 3 & 5 & 2 \\
\hline
 & & 1 & 4 & 5 & 2 \\
 & 4 & 3 & 1 & 2 & 0 \\
2 & 4 & 2 & 0 & 0 & 0 \\
\hline
3 & 3 & 1 & 0 & 1 & 2
\end{array}$$

Start at the right.

$2 \times 4 = 12$ Write 2, and retain 1 to the next step.

$2 \times 2 = 4, 4 + 1 = 5$ Write 5.

$2 \times 5 = 14$ Write 14 as it is the last product on this line.

Next, you shift left one digit and do the multiplication by 5.

Finally, you add, in base 6, all the products you found.

Therefore, $(352)_6 \times (524)_6 = (331012)_6$.

Some divisibility rules

Rule 1: divisibility by 10^n

Consider an integer a written in decimal notation.

$$a = a_n a_{n-1} a_{n-2} \ldots a_1 a_0$$

This number, as discussed earlier, is a notation for the following decimal expansion:

$$a = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + a_{n-2} \cdot 10^{n-2} + \ldots + a_1 \cdot 10 + a_0$$

We can split this number into two parts as follows:

$$\begin{aligned}
a &= a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + a_{n-2} \cdot 10^{n-2} + \ldots + a_1 \cdot 10 + a_0 \\
&= a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + a_{n-2} \ldots a_1 a_0 \\
&= 10^{n-1}(10a_n + a_{n-1}) + a_{n-2} \ldots a_1 a_0 \\
&= 10^{n-1} \cdot k + \underbrace{a_{n-2} \ldots a_1 a_0}_{(n-1)\text{ digits}}
\end{aligned}$$

$k = 10a_n + a_{n-1}$ is an integer because it is the sum of two integers.

Now, if we let $m = n - 1$, a can now be written as

$$a = 10^m \cdot k + \underbrace{a_{m-1} \ldots a_1 a_0}_{m\text{ digits}} = 10^m \cdot k + p.$$

Therefore, a can be written as the sum of a multiple of mth power of 10 and a number p represented by the last m digits of a.

Now, we know that

$10^m \cdot k \equiv 0 \pmod{10^m}$, and hence

$10^m \cdot k + p \equiv p \pmod{10^m}$, and thus

$a \equiv p \pmod{10^m}$,

and this means that a and p have the same remainder when divided by 10^m.

We can conclude that the remainder when dividing any integer by 10^m is the number formed by its last m digits from the right.

For instance, the remainder of dividing 34 527 by 1000 is 527.

As a direct consequence, a number is divisible by 10^m if its last m digits are zeros.

Rule 2: divisibility by 2 and 5

As a consequence of the previous result, we can claim that every integer a can be written as

$a = 10 \cdot k + p$, and hence p represents the last digit!

Now,

$10 \equiv 0$ (mod 2 or 5)

$\Rightarrow 10 \cdot k + p \equiv p$ (mod 2 or 5), and so

$a \equiv p$ (mod 2 or 5).

Therefore, any integer has the same remainder when divided by 2 or 5 as its last digit. Consequently, a number is divisible by 2 or 5 if the last digit is divisible by 2 or 5.

The remainder of dividing 23 456 789 by 2 is 1 since the remainder of dividing 9 by 2 is 1.

The number 123 455 is divisible by 5 because the last digit is divisible by 5.

Rule 3: divisibility by 4 and 25

$a = 100k + p$, where p represents the last two digits.

Similarly to previous discussions,

$a \equiv p$ (mod 4 or 25), which leads to the rule:

The remainder of dividing any integer by 4 or 25 is the same as the remainder of the number representing the last two digits. Similarly the case with divisibility.

The number 123 432 is divisible by 4 since 32 is divisible by 4.

The number 123 432 leaves a remainder of 7 when divided by 25 because 32 does!

8 and 125 have similar rules, but with the last three digits!

Rule 4: divisibility by 3 and 9

Since a can be written as

$a = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + a_{n-2} \cdot 10^{n-2} + \ldots + a_1 \cdot 10 + a_0$, and since

$10 \equiv 1 \pmod{3 \text{ or } 9}$, which also implies that $10^k \equiv 1^k \pmod{3 \text{ or } 9}$, then

$$\begin{aligned} a_n \cdot 10^n &\equiv a_n \pmod{3 \text{ or } 9} \\ a_{n-1} \cdot 10^{n-1} &\equiv a_{n-1} \pmod{3 \text{ or } 9} \\ a_{n-2} \cdot 10^{n-2} &\equiv a_{n-2} \pmod{3 \text{ or } 9} \\ &\vdots \\ a_1 \cdot 10 &\equiv a_1 \pmod{3 \text{ or } 9} \\ a_0 &\equiv a_0 \pmod{3 \text{ or } 9} \end{aligned}$$

Hence, $a \equiv a_n + a_{n-1} + a_{n-2} + \ldots + a_1 + a_0 \pmod{3 \text{ or } 9}$.

Therefore, the remainder of dividing a number by 3 or 9 is the same as the remainder of dividing the sum of its digits by 3 or 9.

Similarly, we can say that a number is divisible by 3 or 9 iff the sum of its digits is divisible by 3 or 9.

Rule 5: divisibility by 11

Since $10 \equiv -1 \pmod{11}$,

$10^2 \equiv 1 \pmod{11}$, and hence

$10^{2k} \equiv 1 \pmod{11}$, and $10^{2k+1} \equiv -1 \pmod{11}$, and thus

$a \equiv (a_0 + a_2 + \ldots + a_{2k} + \ldots) - (a_1 + a_3 + \ldots + a_{2k+1} + \ldots) \pmod{11}$.

This means that the remainder of dividing a number by 11 is equal to the remainder when the difference between the sum of its digits with even position and the sum of its digits with odd position is divided by 11. Similarly, the number is divisible by 11 if the difference between these sums is divisible by 11.

For example, 6 570 289 is divisible by 11 because
$(9 + 2 + 7 + 6) - (8 + 0 + 5) = 11$.

Exercise 2.4

1 Convert $(2009)_{10}$ to base 7 notation.

2 Convert $(3060)_7$ to decimal notation.

3 Convert $(452091)_{10}$ to base 8 notation.

4 Convert $(713060)_8$ to decimal notation.

5 Convert $(1001110011010)_2$ to base 10 notation.

6 Convert $(2010)_{10}$ to binary notation.

7 Convert $(2012452091)_{10}$ to hex notation.

8 Convert $(7B1CE3060)_{16}$ to decimal notation.

9 Convert $(10001111001)_2$ to hex notation.

10 Convert $(11101001110)_2$ to hex notation.

11 Convert $(FECDB)_{16}$ to binary notation.

12 Convert $(7DEFACED89)_{16}$ to binary notation.

13 A number N in base 10 consists of the same digit a repeated n times. For example, 4444444.

a When does $11|N$? **b** When does $3|N$? **c** When does $2|N$?

2.5 Fermat's little theorem

When working with congruences relating to exponents, the next theorem is of great value.

Theorem 15: Fermat's little theorem

If p is prime and a is a positive integer with $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$.

For example, $6^{7-1} \equiv 1 \pmod{7}$, i.e. $6^6 - 1$ is a multiple of 7.

Proof (Optional)

We begin by considering the first $p - 1$ positive multiples of a:

$a, 2a, 3a, 4a, \ldots, (p-1)a$

None of these numbers is congruent to any other modulo p, nor is any congruent to zero. Since if that were the case, then with $1 \leqslant r \leqslant s \leqslant p - 1$, $ra \equiv sa \pmod{p}$. Then using the cancellation law as $\gcd(a, p) = 1$, we will have

$r \equiv s \pmod{p}$, which cannot happen as both s and r are smaller than p.

Hence, the set of integers $a, 2a, 3a, 4a, \ldots, (p-1)a$ would each leave a remainder when divided by p, and the set of these remainders constitute the $p - 1$ residue classes $1, 2, 3, \ldots, p - 1$. Thus,

$$a \cdot 2a \cdot 3a \cdot 4a \cdot \ldots \cdot (p-1)a \equiv 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-1) \pmod{p}$$

$$a \cdot a \cdot a \cdot a \cdot \ldots \cdot a\,(1 \cdot 2 \cdot 3 \cdot 4 \cdot \ldots \cdot (p-1)) \equiv 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-1) \pmod{p},$$

$$\Rightarrow a^{p-1} \cdot (p-1)! \equiv (p-1)! \pmod{p}$$

Another version of this theorem is also used:

If p is prime and a is a positive integer with $p \nmid a$, then $a^p \equiv a \pmod{p}$.

and since $\gcd(p, (p-1)!) = 1$, we can cancel $(p-1)!$, and therefore

$$a^{p-1} \equiv 1 \pmod{p}.$$

Example

This example demonstrates the proof of Fermat's little theorem.

Let $p = 7$ and $a = 5$.

We will consider the first six multiples of 5:

$$1 \cdot 5 \equiv 5 \pmod{7}, \quad 2 \cdot 5 \equiv 3 \pmod{7}, \quad 3 \cdot 5 \equiv 1 \pmod{7},$$
$$4 \cdot 5 \equiv 6 \pmod{7}, \quad 5 \cdot 5 \equiv 4 \pmod{7}, \quad 6 \cdot 5 \equiv 2 \pmod{7}$$

Hence,

$$\begin{aligned}&(1 \cdot 5)(2 \cdot 5)(3 \cdot 5)(4 \cdot 5)(5 \cdot 5)(6 \cdot 5) \equiv 5 \cdot 3 \cdot 1 \cdot 6 \cdot 4 \cdot 2 \pmod{7}\\ &\Rightarrow (1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6) \cdot 5^6 \equiv 6! \pmod{7}\\ &\Rightarrow (6!) \cdot 5^6 \equiv 6! \pmod{7}\end{aligned}$$

Since $\gcd(6!, 7) = 1$, cancel 6!, and therefore

$$5^6 \equiv 1 \pmod{7}.$$

Example 22

Show that $5^{38} \equiv 4 \pmod{11}$.

Solution

We know that $5^{10} \equiv 1 \pmod{11}$, and so $5^{30} \equiv 1 \pmod{11}$.

Also $5^2 \equiv 3 \pmod{11}$, giving us $5^8 \equiv 3^4 \equiv 4 \pmod{11}$.

Therefore, $5^{38} \equiv 1 \cdot 4 \equiv 4 \pmod{11}$.

Example 23

Find the least positive residue of 3^{201} mod 11.

Solution

We know that $3^{10} \equiv 1 \pmod{11}$, and hence $3^{201} = (3^{10})^{20} \cdot 3 \equiv 3 \pmod{11}$.

Example 24

Solve $7x \equiv 3 \pmod{13}$ for x.

Solution

Since $7^{12} \equiv 1 \pmod{13}$, then $7^{11} \cdot 7x \equiv 7^{11} \cdot 3 \pmod{13}$, and hence $x \equiv 7^{11} \cdot 3 \pmod{13}$. Therefore, $x \equiv 2 \cdot 3 \pmod{13} \equiv 6 \pmod{13}$.

Note: Example 24 can be generalized to solve linear congruencies of the form $ax \equiv b \pmod{p}$ when p is prime in the following manner:

If $ax \equiv b \pmod{p}$, then $a^{p-2} \cdot ax \equiv a^{p-2} \cdot b \pmod{p}$, which implies that $a^{p-1}x \equiv a^{p-2} \cdot b \pmod{p}$, and knowing that $a^{p-1} \equiv 1 \pmod{p}$, we will have $x \equiv a^{p-2} \cdot b \pmod{p}$.

Exercise 2.5

1 Find x such that $3^{12} \equiv x \pmod{11}$.

2 Find x such that $3^{21} \equiv x \pmod{11}$.

3 Find the value of $5^{173} \pmod{13}$.

4 Find the value of $6^{47} \pmod{17}$.

5 Find the value of $10^{321} \pmod{11}$.

6 Solve $8x \equiv 7 \pmod{17}$.

7 Solve $3x \equiv 10 \pmod{17}$.

8 Solve $7x \equiv 12 \pmod{17}$.

9 Solve $3x \equiv 4 \pmod{11}$.

10 Solve $3^{14} \equiv x \pmod{13}$.

11 Solve $3^{45} \equiv x \pmod{13}$.

12 **a** Use Fermat's little theorem to calculate: $7^{2009} \pmod{11}$, $7^{2009} \pmod{13}$, and $7^{2009} \pmod{17}$.

b Hence, calculate: $7^{2009} \pmod{2431}$.

13 Find the remainder upon dividing 512^{372} by 13.

14 Find the remainder upon dividing 3444^{3233} by 17.

15 Find the remainder upon dividing 314^{159} by 31.

16 **a** Show that if p is a prime number then $(a + 1)^p \equiv a^p + 1 \pmod{p}$, where a is an integer.

b Hence, derive Fermat's little theorem.

17 Show that $11^{104} + 1$ is a multiple of 17.

18 Let x and 35 be relatively prime numbers. Show that $x^{12} \equiv 1 \pmod{35}$.

19 Let x and 42 be relatively prime numbers. Show that $x^6 \equiv 1 \pmod{168}$.

20 Show that each of the following is true:

a $b^{21} \equiv b \pmod{15}$, for all integers b.

b $b^7 \equiv b \pmod{42}$, for all integers b.

c $b^9 \equiv b \pmod{30}$, for all integers b.

2.6 Recurrence relations

Sometimes it is difficult to define a function, a sequence, or a set explicitly. However, it may be easier to define it in terms of itself! This process is called recursion.

For instance, you recall from Chapter 4 in your book that we can use recursion to define sequences. For example, the arithmetic sequence is defined, recursively by stating the first term a_1 and by writing down the rule for finding any term of the sequence from the previous one. In the case of the arithmetic sequence, this rule is:

$a_n = a_{n-1} + d$, where d is the common difference.

Similarly, you know that a geometric sequence is defined by stating the first term g_1 and the rule for finding each term from previous ones, namely:

$g_n = g_1 r^{n-1}$, where r is the common ratio.

The arithmetic and geometric sequences have their explicit forms of course. Moving between an explicit form and a recursive form is a necessity in many cases. Specifically, the explicit form of these types is easier to work with in cases where the value of a large term is required. Imagine that you need to find the value of the 100th term in an arithmetic sequence. Using the recursive definition means that you have to know the 99th term in order to get to your target, while knowing the explicit form enables you to find the requested term by simply substituting 100 for n in the explicit formula

$$a_{\text{n}} = a_1 + (n - 1)d.$$

Consider the following situation. You are given a sequence with $a_0 = 1$ and $a_n = 3a_{n-1}$ for $n > 0$. By looking at a few terms you can easily recognize this sequence as that of the powers of 3 i.e. $a_n = 3^n$ for $n \geqslant 0$. Of course, it is simpler to work with the latter form.

Example 25

Find an explicit formula for the following sequence.

$$a_0 = 1$$
$$a_n = na_{n-1} \text{ for } n > 0$$

Solution

The first few terms will give you an idea of what the explicit form of the definition is

$$a_0 = 1,\ a_1 = 1 \times 1,\ a_2 = 2 \times 1 = 2,\ a_3 = 3 \times 2 = 6,\ a_4 = 4 \times 6 = 24,\ \ldots$$

This in fact is nothing but n!

To prove this, we can use mathematical induction.

Basis step

$a_1 = 1 \times 1!$

Inductive step

Assume that the statement is true for $n = k$, i.e. $a_k = k!$. We prove that it is true for $n = k + 1$.

By definition, $a_{k+1} = (k + 1)a_k = (k + 1) \cdot k! = (k + 1)!$

Recurrence relations

As we discussed above, you notice that a recursive definition of a sequence identifies one or more early terms and a law for defining later terms from those preceding them. Such rules are called **recurrence relations**. So, when the problem is to discover an explicit formula for a recursively defined sequence, the recursive formula is called a **recurrence relation**. Remember that to define a sequence well, a recursive formula must be supplemented by information about some earlier terms of the sequence. This information is called the **initial condition(s)** for the sequence.

Definition 1

A **recurrence relation** for a sequence $\{a_n\}$ is a formula that expresses a_n in terms of one or more of the previous terms of the sequence: a_{n-1}, a_{n-2}, etc.…

Definition 2

A sequence is called a **solution** of a recurrence relation, if its terms satisfy the recurrence relation.

Definition 3

Initial conditions are explicitly given values for a certain number of the terms of the sequence.

Examples

i The recurrence relation $a_n = 2a_{n-1} + 3$ for $n > 1$ with $a_1 = 5$ defines the sequence 13, 29, 61, 125, ….

ii The recurrence relation $F_n = F_{n-1} + F_{n-2}$ for $n > 2$ with **initial conditions** $F_1 = 1$ and $F_2 = 1$ describes the well-known Fibonacci sequence 1, 1, 2, 3, 5, 8, …

iii In Example 25 above, the recurrence relation is $a_n = na_{n-1}$ for $n > 0$ and the initial condition $a_0 = 1$ describes the sequence $a_n = n!$

Example 26

Consider the recurrence relation $u_{n+1} = 2u_n - u_{n-1}$ for $n \geqslant 1$. Which of the following is a solution of this relation?

a) $u_n = 3n$ b) $u_n = 2^n$ c) $u_n = 5$

Solution

a) For $u_n = 3n$ with $n \geqslant 1$, we see that according to the recurrence relation, if u_n is a solution, then

$$u_n = 2u_{n-1} - u_{n-2} = 2(3(n-1)) - 3(n-2) = 6n - 6 - 3n + 6 = 3n$$

$\therefore\ u_n = 3n$ is a solution.

b) For $u_n = 2^n$ with $n \geqslant 1$, we see that according to the recurrence relation, if u_n is a solution, then

$$u_n = 2u_{n-1} - u_{n-2} = 2(2^{(n-1)}) - 2^{(n-2)} = 2^n - 2^{n-2} \neq 2^n$$

$\therefore\ u_n = 2^n$ is not a solution.

c) For $u_n = 5$ with $n \geqslant 1$, we see that according to the recurrence relation, if u_n is a solution, then

$$u_n = 2u_{n-1} - u_{n-2} = 2(5) - 5 = 5$$

$\therefore\ u_n = 5$ is a solution.

Consider this sequence: $u_n = 3n + 5$.

If u_n is a solution, then

$$\begin{aligned} u_n &= 2u_{n-1} - u_{n-2} \\ &= 2(3(n-1)+5) - 3(n-2) - 5 \\ &= 6n - 6 + 10 - 3n + 6 - 5 \\ &= 3n + 5 \end{aligned}$$

$\therefore\ u_n = 3\mathrm{n} + 5$ is a solution.

This demonstrates a theorem which we will prove in Section 2.8 that if u and v are solutions of a linear recurrence relation, then $au + bv$ where a and b are arbitrary constants, is also a solution.

Example 27

Consider the recurrence relation $a_{n+1} = 2a_n + 1$ with the initial condition $a_1 = 7$.

a) Find a_2, a_3, and a_4.

b) Show that $a_n = 2^{n+2} - 1$ is a solution to this recurrence relation.

Solution

a) $a_2 = a_{1+1} = 2a_1 + 1 = 15$

$a_3 = a_{2+1} = 2a_2 + 1 = 31$

$a_4 = a_{3+1} = 2a_3 + 1 = 63$

b) Notice that $a_2 = 15 = 2^{2+2} - 1$, $a_3 = 31 = 2^{3+2} - 1$, and $a_4 = 63 = 2^{4+2} - 1$

Now, substituting $a_n = 2^{n+2} - 1$ into a_{n+1} will give us:

$$a_{n+1} = 2^{n+1+2} - 1 = 2^{n+3} - 1 = 2\,(2^{n+2}) - 1 = 2(2^{n+2} - 1) + 1 = 2a_n + 1$$

2.7 Modelling with recurrence relations

We can use recurrence relations to model a diverse range of situations. Such situations include counting bit strings with specific properties, compound interest, counting growth of populations under specific constraints, and some counting related to recreational mathematics! Here are some examples.

Compound interest

Consider that a person makes a one-off deposit of an amount of P_0 in a savings account that pays r in annual interest. (r is in decimal notation. For example for 5% , $r = 0.05$)

How much money will be in the account after n years?

Solution

Let P_n denote the amount in the account after n years. Then P_n is equal to the amount that has accumulated over the last $n - 1$ years, P_{n-1}, plus the interest earned during the nth year, $r\,P_{n-1.}$

Therefore, $P_n = P_{n-1} + r\,P_{n-1} = (1 + r)\,P_{n-1}$.

To find an explicit formula for the amount of money, we can use an iterative approach for that purpose. (It is also called **backtracking**.)

$$P_1 = (1 + r)\,P_0$$

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

$$P_3 = (1 + r)^3\,P_0$$

$$\vdots$$

$$P_n = (1 + r)^n\,P_0$$

You have seen this formula in Chapter 4 of the textbook too. We can use mathematical induction to establish its validity.

Basis step
For $n = 0$, $P_0 = (1 + r)^0 P_0 = P_0$

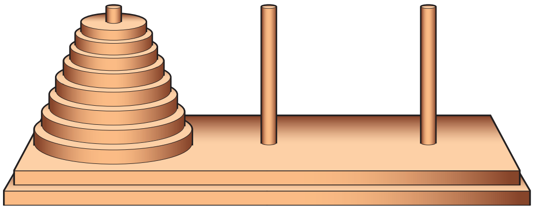
Inductive step
Assume this to be true for $n = k$, i.e. $P_k = (1 + r)^k\,P_0$

For $n = k + 1$

$$P_{k+1} = (1 + r)\,P_k = (1 + r)\,(1 + r)^k\,P_0 = (1 + r)^{k+1}\,P_0$$

$\therefore\ P_n = (1 + r)^n\,P_0$ for all possible values of n.

Tower of Hanoi



The Tower of Hanoi puzzle involves moving a pile of different-sized disks from one peg to another, using an intermediate peg. Only one disk at a time can be moved, a disk can only be moved if it is the top disk on a pile, and a larger disk can never be placed on a smaller one. Our task is to find the number of moves needed to move all the n disks from peg 1 to peg 3 for example.

Solution

Let d_n represent the number of moves required to move the disks from peg 1 to 3, using peg 2 as an auxiliary 'stop'.

We can move the top $n - 1$ disks, following the rules of the game, from peg 1 to peg 2, leaving the largest disk at peg 1. This can be done in d_{n-1} ways.

Now we move the largest disk, in one move from peg 1 to peg 3. The next step is then to move the $n - 1$ disks from peg 2 to peg 3, which can be done in d_{n-1} ways again. Hence, the total number of moves is now

$$d_n = d_{n-1} + d_{n-1} + 1 = 2d_{n-1} + 1$$

This is the recurrence relation leading us to the solution. The initial condition here is $d_1 = 1$, because one disk requires one move only to be transferred from peg 1 to peg 3.

We can use an iterative method (backtracking) to solve this recurrence relation

$$\begin{aligned} d_n &= 2d_{n-1} + 1 \\ &= 2(2d_{n-2} + 1) + 1 = 2^2d_{n-2} + 2 + 1 \\ &= 2^2(2d_{n-3} + 1) + 2 + 1 = 2^3d_{n-3} + 2^2 + 2 + 1 \\ &= 2^3(2d_{n-4} + 1) + 2^2 + 2 + 1 = 2^4d_{n-4} + 2^3 + 2^2 + 2 + 1 \\ &\vdots \\ &= 2^{n-1}d_1 + 2^{n-2} + 2^{n-3} + \cdots + 2 + 1 \end{aligned}$$

However, $d_1 = 1$, and so

$$d_n = 2^{n-1} + 2^{n-2} + 2^{n-3} + \cdots + 2 + 1$$

The right-hand side of this equation is a geometric series, with first term 1 and common ratio 2, and hence

$$d_n = 2^{n-1} + 2^{n-2} + 2^{n-3} + \cdots + 2 + 1 = 1 \cdot \frac{1 - 2^n}{1 - 2} = 2^n - 1$$

The formula can be proved using mathematical induction:

Basis step
For $n = 1$, $d_1 = 2^1 - 1 = 1$, which is true.

Inductive step
Assume this to be true for $n = k$, i.e. $d_k = 2^k - 1$.

For $n = k + 1$,

$d_{k+1} = 2d_k + 1$. according to the recurrence relation, and thus
$d_{k+1} = 2(2^k - 1) + 1 = 2^{k+1} - 2 + 1 = 2^{k+1} - 1$ as required.

Fibonacci's Rabbits

The imaginative problem that Fibonacci probed (in the year 1202) was about how fast rabbits could breed in *ideal* settings.

Presume that a newborn pair of rabbits, one male and one female, are put in a field. Rabbits are able to mate at the age of 1 month so that at the end of its second month a female can produce another pair of rabbits. Assume that our rabbits **never die** and that the female always produces **one new pair (one male, one female) every month** from the second month on. The puzzle that Fibonacci posed was this:

How many pairs will there be in 1 year?

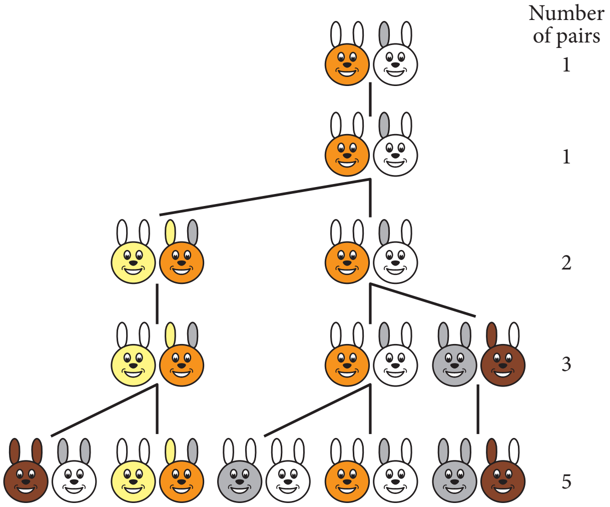
Our task here is to find a recurrence relation for the number of pairs of rabbits after *n* months.

Solution

Consider the situation according to the way it is set up.

- By the end of the first month, there is only one pair, the original.
- At the end of the second month, they mate, but there is still one only 1 pair.
- At the end of the third month the female produces a new pair, so now there are 2 pairs of rabbits in the field.
- At the end of the fourth month, the original female produces a second pair, making 3 pairs in all in the field, the newborn pair mate but no new children yet.
- At the end of the fifth month, the original female has produced yet another new pair, the female born two months ago produces her first pair also, making 5 pairs.

Now let r_n be the number of pairs of rabbits at the end of n months. So, at the end of the first month there is only one pair, i.e., $r_1 = 1$. At the end of the second month, still one pair, i.e., $r_2 = 1$. At the end of the third month there are two pairs, i.e., $r_3 = 2$ and so on.



To find the number after n months, we add the number in the field in the previous month, r_{n-1}, and the number of the newborn pairs, which will be

r_{n-2}, since each newborn pair comes from a pair at least 2 months old.

Consequently:

$$r_n = r_{n-1} + r_{n-2}$$

This, along with the initial conditions $r_1 = 1$and $r_2 = 1$ describes the Fibonacci sequence which you know already.

2.8 Solving linear recurrence relations

As you have seen earlier, some of the recurrence relations can be solved using iteration (backtracking), others can be solved by some other improvised techniques, and a specific type known as **linear homogeneous recurrence relations with constant coefficients** can be solved explicitly in a systematic manner.

Definition 1

A **linear homogeneous recurrence relations of degree *k* with constant coefficients** is a recurrence relation of the form

$$a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k},$$

where $c_1, c_2, \ldots, c_k$ are real numbers with $c_k \neq 0$.

This relation is linear because the right-hand side is a linear combination of the earlier terms of the sequence, and homogeneous because *all* terms are multiples of the a_is. The coefficients of the terms are all constants rather than functions of n. The degree of the relation is k because a_n is expressed in terms of the previous k terms of the sequence.

Note: In this book, we will limit our discussion to linear recurrence relations of at most second degree.

Example 28

Which of the following recurrence relations are linear homogeneous?

a) $s_n = 3s_{n-1}$

b) $f_{n+1} = f_n + f_{n-1}$

c) $b_n = 2b_{n-1}\,b_{n-2}$

d) $a_n = 2a_{n-1} + 5n$

e) $A_n = 1.09A_{n-1}$

f) $c_n = 2c_{n-1} - c^2{}_{n-2}$

Solution

a) This is linear homogeneous since the nth term is a constant multiple of the previous term.

b) This is linear homogeneous since the nth term is the sum of the previous two terms.

c) This is not linear homogeneous since the *n*th term is the product of the previous two terms and not a constant multiple of one of them.

d) This is not linear homogeneous since the right-hand side contains a function of *n* rather than a constant.

e) This is linear homogeneous since the *n*th term is a constant multiple of the previous term.

f) This is not linear homogeneous since the right-hand side contains a power of one term that is higher than 1.

The basic approach for solving linear homogeneous recurrence relations is to look for solutions of the form $a_n = x^n$, where x is a constant.

Obviously $a_n = x^n$ is a solution of the recurrence relation

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

if and only if

$$x^n = c_1 x^{n-1} + c_2 x^{n-2} + \ldots + c_k x^{n-k}$$

Multiplying both sides by x^{k-n} and simplifying will yield the equation

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

Note
$x^k - c_1 x^{k-1} - c_2 x^{k-2} - \ldots - c_k$ is known as the **characteristic polynomial**.

This is called the **characteristic equation** of the recurrence relation. Obviously too, the sequence $\{a_n\}$ with $a_n = r^n$ is a solution to the recurrence relation if and only if r is a solution of the characteristic equation.

We will demonstrate the general method of solving linear homogeneous relations with constant coefficients by finding an explicit solution to a second order relation first.

Example 29

Solve the recurrence relation $a_n = 2a_{n-1} + 8a_{n-2}$ with initial conditions $a_0 = 4$, $a_1 = 10$.

Solution

The associated characteristic equation is:

$$x^2 - 2x - 8 = 0$$

Solving this equation, we have two solutions.

$$r_1 = 4 \text{ or } r_2 = -2.$$

At this point, we have two solutions of the recursive relation.

$$s_n = 4^n \text{ or } t_n = (-2)^n$$

In Example 27 of Section 2.6, we verified a theorem that if *u* and *v* are solutions, then a linear combination of *u* and *v* will also be a solution. Thus

$$a_n = b(4^n) + d(-2)^n$$

is a solution to the relation.

To satisfy the initial conditions, we must have

$$a_0 = 4 \Rightarrow b(4^0) + d(-2)^0 = 4 \Rightarrow b + d = 4$$

$$a_1 = 10 \Rightarrow b(4^1) + d(-2)^1 = 10 \Rightarrow 4b - 2d = 10$$

Solving this system we find that $b = 3$ and $d = 1$, and thus

$$a_n = 3(4^n) + (-2)^n$$

is the solution to the recurrence relation.

Notice that

$$\begin{aligned} 2a_{n-1} + 8a_{n-2} &= 2(3(4^{n-1}) + (-2)^{n-1}) + 8(3(4^{n-2}) + (-2)^{n-2}) \\ &= 6(4^{n-1}) + 2\,(-2)^{n-1} + 24\,(4^{n-2}) + 8\,(-2)^{n-2} \\ &= 6(4^{n-1}) + 2\,(-2)^{n-1} + 6\,(4^{n-1}) - 4\,(-2)^{n-1} \\ &= 12(4^{n-1}) - 2\,(-2)^{n-1} \\ &= 3(4^n) + (-2)^n \\ &= \mathrm{a}_n \end{aligned}$$

This verifies that a_n is a solution to the recurrence relation.

Theorem 1

If u_n and v_n are solutions to the second order linear homogeneous recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2}$, then $t_n = bu_n + dv_n$ is also a solution.

Proof

Since u_n and v_n are solutions then

$$u_n = c_1u_{n-1} + c_2u_{n-2} \text{ then } v_n = c_1v_{n-1} + c_2v_{n-2}$$

Thus

$$\begin{aligned} t_n &= bu_n + dv_n = b(c_1u_{n-1} + c_2u_{n-2}) + d(c_1v_{n-1} + c_2v_{n-2}) \\ &= c_1(bu_{n-1} + dv_{n-1}) + c_2(bu_{n-2} + dv_{n-2}) \\ &= c_1t_{n-1} + c_2t_{n-2} \end{aligned}$$

Therefore t_n is a solution to $a_n = c_1a_{n-1} + c_2a_{n-2}$.

Theorem 2

1 If the characteristic polynomial $x^2 - c_1x - c_2$ of the recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2}$ has two distinct zeros r_1 and r_2, then $a_n = br_1{}^n + dr_2{}^n$ where b and d depend on the initial conditions, is the explicit formula for the solution sequence.

2 If the characteristic polynomial $x^2 - c_1x - c_2$ of the recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2}$ has a single zero r, then $a_n = br^n + dnr^n$ where b and d depend on the initial conditions, is the explicit formula for the solution sequence.

3 If the characteristic polynomial $x^2 - c_1x - c_2$ of the recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2}$ has two conjugate complex zeros z_1 and z_2, then we express these zeros in polar form where $z_1 = (r, \theta)$ and $z_2 = (r, -\theta)$ and the solution will be of the form $a_n = r^n(b\cos(n\theta) + d\sin(n\theta))$ where b and d depend on the initial conditions.

Proof

1 Suppose that r_1 and r_2 are zeros of $x^2 - c_1x - c_2$, so $r_1^2 - c_1r_1 - c_2 = 0, r_2^2 - c_1r_2 - c_2 = 0$, and $a_n = br_1{}^n + dr_2{}^n$, for $n \geqslant 1$. We show that this definition of a_n defines the same sequence as $a_n = c_1a_{n-1} + c_2a_{n-2}$.

First we note that b and d are chosen so that the initial conditions are satisfied. That is

$a_1 = br_1 + dr_2$ and $a_2 = br_1^2 + dr_2^2$.

Thus

$$\begin{aligned} a_n &= br_1^n + dr_2^n \\ &= br_1^{n-2}r_1^2 + dr_2^{n-2}r_2^2 \end{aligned}$$

Now, using the fact that r_1 and r_2 are zeros of $x^2 - c_1x - c_2$, we have

$$r_1^2 - c_1r_1 - c_2 = 0 \Rightarrow r_1^2 = c_1r_1 + c_2$$

and $$r_2^2 - c_1r_2 - c_2 = 0 \Rightarrow r_2^2 = c_1r_2 + c_2$$

Thus

$$\begin{aligned} a_n &= br_1^{n-2}r_1^2 + dr_2^{n-2}r_2^2 \\ &= br_1^{n-2}(c_1r_1 + c_2) + dr_2^{n-2}(c_1r_2 + c_2) \\ &= c_1(br_1^{n-1} + dr_2^{n-1}) + c_2(br_1^{n-2} + dr_2^{n-2}) \\ &= c_1a_{n-1} + c_2a_{n-2} \end{aligned}$$

2 This part may be proved in a similar manner and is left as an exercise.

3 The proof of this part is beyond the scope of this book.

Example 30

Find the solution to the recurrence relation $a_n = 3a_{n-1} - 2a_{n-2}$, where $a_1 = 5$, $a_2 = 3$.

Solution

The characteristic equation associated with this relation is

$$x^2 - 3x + 2 = 0$$

The characteristic roots are 1 and 2.

Thus, the solution to the relation is of the form

$$a_n = br_1{}^n + dr_2{}^n = b(1^n) + d(2^n),$$

With the initial conditions, we have

$$\left.\begin{aligned} b + 2d &= 5 \\ b + 4d &= 3 \end{aligned}\right\} \Rightarrow b = 7, d = -1$$

Therefore the solution is

$$a_n = 7 - 2^n$$

Note: Notice here that using $a_n = 7 - 2^n$, we find that the first 5 terms are: 5, 3, –1, –9, and –25 and using the recurrence relation, we have 5, 3, –1, –9, and –25.

Example 31

Solve the recurrence relation $u_n = 4u_{n-1} - 4u_{n-2}$, where $u_0 = 1$, $u_1 = 1$.

Solution

The associated characteristic equation is

$$x^2 - 4x + 4 = 0$$

This has one solution, $x = 2$.

According to theorem 2, the solution to this equation has the form

$$u_n = bx^n + dnx^n$$

Thus, the solution for this relation is

$$u_n = b2^n + dn2^n$$

The initial conditions yield

$$\left.\begin{aligned} 1 &= b \\ 1 &= 2b + 2d \end{aligned}\right\} \Rightarrow b = 1, d = -\frac{1}{2}$$

Therefore, the solution is

$$u_n = 2^n - \frac{1}{2}n2^n = 2^n - n2^{n-1}$$

Example 32

Solve the recurrence relation $v_n = 2v_{n-1} - 2v_{n-2}$, where $v_0 = 1$, $v_1 = 3$.

Solution

The characteristic equation for the recurrence relation is

$$t^2 - 2t + 2 = 0$$

The characteristic roots are then

$$z_1 = 1 + i, z_2 = 1 - i$$

When written in polar form, the roots are

$$z_1 = \sqrt{2}cis\frac{\pi}{4},\ z_2 = \sqrt{2}cis\frac{-\pi}{4}$$

Thus any solution of the relation is of the form

$$v_n = \left(\sqrt{2}\right)^n\left(b\cos\left(n\frac{\pi}{4}\right) + d\sin\left(n\frac{\pi}{4}\right)\right)$$

With the initial conditions we have

$$\left.\begin{array}{l} v_0 = 1 = \left(\sqrt{2}\right)^0\left(b\cos\left(0\cdot\frac{\pi}{4}\right) + d\sin\left(0\cdot\frac{\pi}{4}\right)\right) = b \\ v_1 = 3 = \left(\sqrt{2}\right)^1\left(b\cos\left(\frac{\pi}{4}\right) + d\sin\left(\frac{\pi}{4}\right)\right) = \sqrt{2}\left(b\frac{1}{\sqrt{2}} + d\frac{1}{\sqrt{2}}\right) \end{array}\right\} \Rightarrow \left.\begin{array}{l} b = 1 \\ b + d = 3 \end{array}\right\} \Rightarrow b = 1;\ d = 2$$

The solution of the recurrence equation is then

$$v_n = \left(\sqrt{2}\right)^n\left(\cos\left(n\frac{\pi}{4}\right) + 2\sin\left(n\frac{\pi}{4}\right)\right)$$

Example 33

Consider the Fibonacci sequence $F_n = F_{n-1} + F_{n-2}$ for $n > 2$ with initial conditions $F_1 = 1$ and $F_2 = 1$.

Find an explicit expression for F_n.

Solution

The characteristic equation associated with this is

$$x^2 - x - 1 = 0$$

The characteristic roots are then

$$r_1 = \frac{1+\sqrt{5}}{2} \text{ and } r_2 = \frac{1-\sqrt{5}}{2}.$$

Thus, any solution to Fibonacci's sequence is of the form

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

Now using the initial conditions we have

$$\left.\begin{array}{l} F_1 = b\left(\frac{1+\sqrt{5}}{2}\right)^1 + d\left(\frac{1-\sqrt{5}}{2}\right)^1 = 1 \\ F_2 = b\left(\frac{1+\sqrt{5}}{2}\right)^2 + d\left(\frac{1-\sqrt{5}}{2}\right)^2 = 1 \end{array}\right\} \Rightarrow b = \frac{1}{\sqrt{5}}, d = \frac{-1}{\sqrt{5}}$$

Hence, Fibonacci's nth term is

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

Linear non-homogeneous recurrence relations with constant coefficients

We have seen how to solve linear homogeneous recurrence relations by using characteristic polynomials and some other relations by using iteration. This section explores techniques that can be used to solve non-homogeneous relations.

For example, $a_n = 2a_{n-1} + 3n$ is a recurrence relation but not homogeneous.

Definition
A recurrence relation of the form

$a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k} + f(n)$

where c_i for $i = 1, 2, \ldots, k$ are real numbers and $f(n)$ is a function of n not identically zero is a **linear non-homogeneous recurrence relation with constant coefficients**.

The recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k}$ is called the associated homogeneous recurrence relation. It plays a very important role in the solution of the non-homogeneous recurrence relation.

Examples

Each of the following recurrence relations are linear non-homogeneous.

i $a_n = 2a_{n-1} + 3n$

ii $b_n = b_{n-1} - 3b_{n-2} + n^2 + 2n$

iii $u_n = 2u_{n-1} + u_{n-2} + 2n5^n$

Each of the following is the associated linear homogeneous recurrence relation.

i $a_n = 2a_{n-1}$

ii $b_n = b_{n-1} - 3b_{n-2}$

iii $u_n = 2u_{n-1} + u_{n-2}$

The importance of the associated homogeneous relations in the solution of the non-homogeneous relations is shown by the following theorem.

Theorem 3 (Optional)

If p_n is a particular solution of the linear non-homogeneous recurrence relation with constant coefficients $a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k} + f(n)$ and if h_n is a solution of the associated homogeneous relation $a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k}$, then every solution of the non-homogeneous relation is of the form $p_n + h_n$.

Proof

If p_n is a solution of the non-homogeneous relation, then

$$p_n = c_1 p_{n-1} + c_2 p_{n-2} + \dots + c_k p_{n-k} + f(n).$$

Suppose that q_n is another solution of the non-homogeneous equation, then

$$q_n = c_1 q_{n-1} + c_2 q_{n-2} + \dots + c_k q_{n-k} + f(n)$$

Subtracting the two equations gives

$$q_n - p_n = c_1 (q_{n-1} - p_{n-1}) + c_2 (q_{n-2} - p_{n-2}) + \dots + c_k (q_{n-k} - p_{n-k})$$

This shows that $q_n - p_n$ is a solution of the associated homogeneous relation

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

Call this solution h_n and so: $h_n = a_n = q_n - p_n \Rightarrow q_n = h_n + p_n$

Example 34 (Optional)

Find a solution of the recurrence relation $a_n = 2a_{n-1} + 3 \cdot 2^n$.

Solution

The associated homogeneous relation is

$$a_n = 2a_{n-1}$$

This is easily spotted to be a geometric sequence and hence has a solution

$$h_n = b2^n$$

To find a particular solution, we can attempt $p_n = dn \cdot 2^n$

To find d we substitute p_n back into the original non-homogeneous relation

$$dn \cdot 2^n = 2d\,((n-1)2^{n-1}) + 3 \cdot 2^n$$

Simplify the equation by dividing through by 2^{n-1} to get

$$2dn = 2d\,(n-1) + 3 \cdot 2 \Rightarrow d = 3$$

Thus, the particular solution we seek is $p_n = 3n \cdot 2^n$ and hence the general solution of the non-homogeneous relation is the sum of the solution to the homogeneous relation and this one: $h_n = b2^n + 3n2^n = (b + 3n)2^n$.

Exercise 2.6–2.8

In questions 1–4, give the first five terms and identify the recurrence relation as linear homogeneous or not. If the relation is linear homogeneous, then what is its degree?

1 $b_n = \frac{5}{2} b_{n-1}; b_1 = 6$

2 $a_n = -3a_{n-1} - 2a_{n-2};\ a_1 = -2,\ a_2 = 4$

3 $a_n = 2^{n-1} a_{n-1}; a_1 = 5$

4 $b_n = 5b_{n-1} + 3; b_1 = 1$

In questions 5–10, solve each of the recurrence relations.

5 $b_n = \frac{5}{2}b_{n-1}; b_1 = 4$

6 $a_n = 5a_{n-1} + 3; a_1 = 3$

7 $a_n = a_{n-1} + n; a_1 = 4$

8 $b_n = -\frac{11}{10}b_{n-1}; b_1 = 10$

9 $a_n = a_{n-1} - 2; a_1 = 0$

10 $b_n = nb_{n-1}; b_1 = 8$

In questions 11–13, solve each of the recurrence relations.

11 $b_n = 4b_{n-1} + 5b_{n-2}; b_1 = 6,\ b_2 = 6$

12 $a_n = -3a_{n-1} - 2a_{n-2}; a_1 = -2,\ a_2 = 4$

13 $a_n = 2a_{n-1} - 2a_{n-2}; a_1 = 1,\ a_2 = 4$

14 Develop a general explicit formula for a recurrence relation of the form $u_n = au_{n-1} + b$ where a and b are real numbers.

Apply the result to the situations above that fit that model.

Practice questions 2

- Practice questions 1–10 cover work from Chapters 1–2 inclusive.

1 For any positive integers a and b, let gcd(a, b) and lcm (a, b) denote the greatest common divisor and the least common multiple of a and b, respectively.
Prove that

$a \times b = (\gcd(a, b)) \times (\text{lcm}(a, b))$.

2 **a** Using Euclid's algorithm, find integers x and y such that $17x + 31y = 1$.

b Given that $17p + 31q = 1$, where $p, q \in \mathbb{Z}$, show that

$|p| \geqslant 11$ and $|q| \geqslant 6$.

3 Find the remainder when 67^{101} is divided by 65.

4 **a** Convert the number 95 from base 10 to base 6.

b Working in base 6, square your answer to part **a**.

c Convert your answer to part **b** to a base 10 number.

5 The function $f: \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$ is defined by $f(x) = \gcd(x, 6)$.

a Find the range of the function f.

b Show that the function f is periodic and find its period.

c Find the set of positive integers satisfying $f(x) = 2$.

6 **a** Use the Euclidean algorithm to find the greatest common divisor of 43 and 73.

Consider the equation $43x + 73y = 7$, where $x, y \in \mathbb{Z}$.

b **i** Find the general solution of this equation.

ii Find the solution which minimizes $|x| + |y|$.

7 **a** Use the Euclidean algorithm to show that 275 and 378 are relatively prime.

b Find the general solution to the Diophantine equation $275x + 378y = 1$.

8 **a** Define what is meant by the statement $x \equiv y \pmod{n}$, where $x, y, n \in \mathbb{Z}^+$.

b Hence, prove that if $x \equiv y \pmod{n}$ then $x^2 \equiv y^2 \pmod{n}$.

c Determine whether or not $x^2 \equiv y^2 \pmod{n}$ implies that $x \equiv y \pmod{n}$.

9 **a** **i** Given that $a \equiv d \pmod{n}$ and $b \equiv c \pmod{n}$, prove that

$(a + b) \equiv (c + d) \pmod{n}$.

ii Hence, solve the system:

$$\begin{cases} 2x + 5y \equiv 1 \pmod{6} \\ x + y \equiv 5 \pmod{6} \end{cases}$$

b Show that $x^{97} - x + 1 \equiv 0 \pmod{97}$ has no solution.

10 **a** Given that $ax \equiv b \pmod{p}$, where $a, b, p, x \in \mathbb{Z}^+$, p is prime and a is not a multiple of p, use Fermat's little theorem to show that

$x \equiv a^{p-2}b \pmod{p}$.

b Hence, solve the simultaneous linear congruences

$3x \equiv 4 \pmod{5}$

$5x \equiv 6 \pmod{7}$

giving your answer in the form $x \equiv c \pmod{d}$.

Questions 1–10 © International Baccalaureate Organization

3 Graphs

Terminology

You should be aware that many different terminologies exist in graph theory and that different textbooks may employ different combinations of these.

In IB examination questions, the terminology used will be as it appears in the syllabus. A summary of the terminology is provided below.

Graph Consists of a set of vertices and a set of edges; an edge joins its endpoints (vertices).

Subgraph A graph within a graph.

Weighted graph A graph in which each edge is allocated a number or weight.

Loop An edge whose endpoints are joined to the same vertex.

Multiple edges Multiple edges occur if more than one edge joins the same pair of vertices.

Walk A sequence of linked edges.

Trail A walk in which no edge appears more than once.

Path A walk with no repeated vertices.

Circuit A walk that begins and ends at the same vertex, and has no repeated edges.

Cycle A walk that begins and ends at the same vertex, and has no other repeated vertices.

Hamiltonian path A path that contains all the vertices of the graph.

Hamiltonian cycle A cycle that contains all the vertices of the graph.

Eulerian trail A trail that contains every edge of a graph.

Eulerian circuit A circuit that contains every edge of a graph.

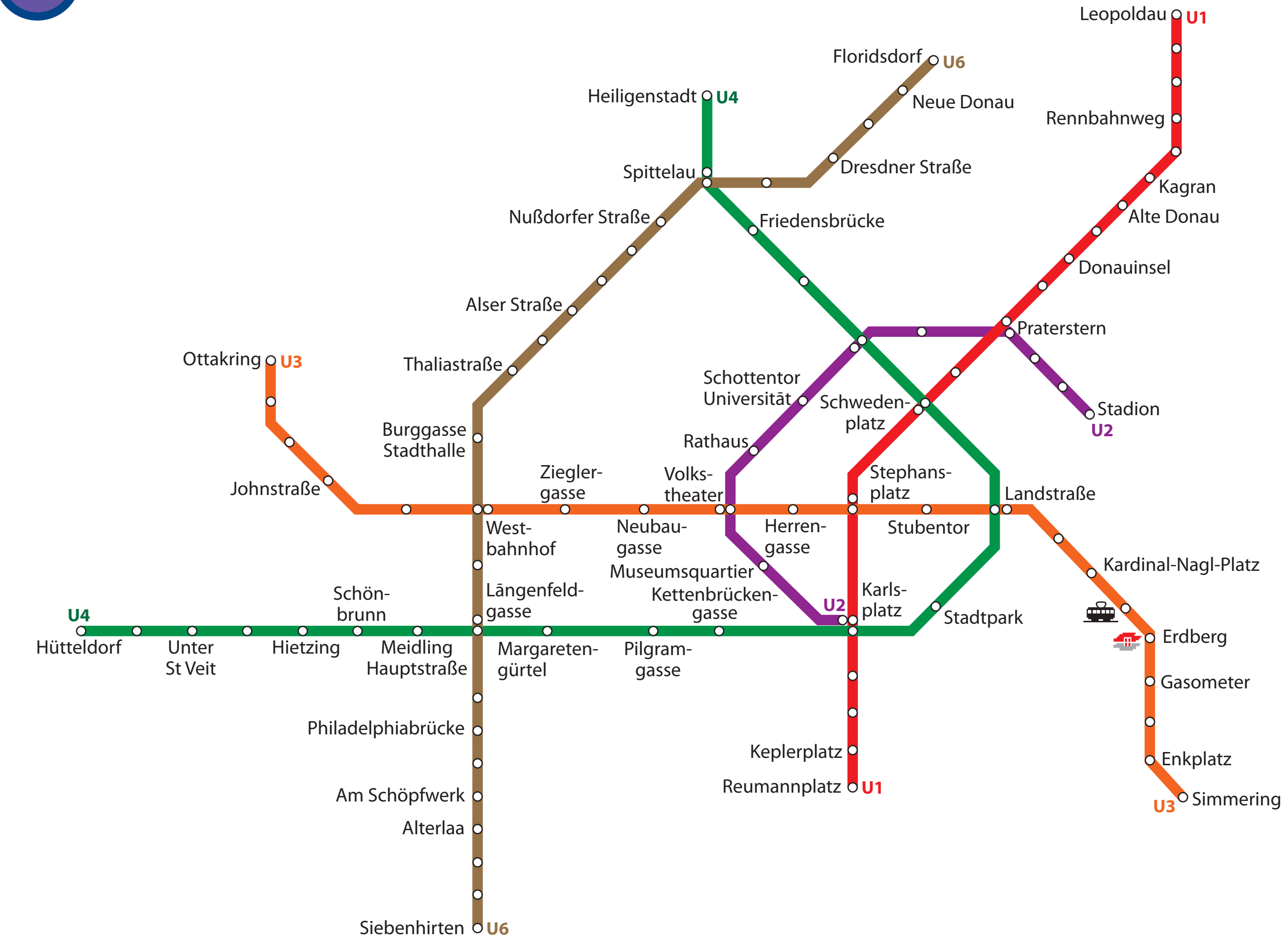
Degree of a vertex The number of edges joined to the vertex; a loop contributes two, one for each of its endpoints.

Simple graph A graph without loops or multiple edges.

Complete graph A simple graph where every vertex is joined to every other vertex.

Connected graph	A graph that has a path joining every pair of vertices.
Disconnected graph	A graph that has at least one pair of vertices not joined by a path.
Tree	A connected graph that contains no cycles.
Weighted tree	A tree in which each edge is allocated a number or weight.
Spanning tree of a graph	A subgraph containing every vertex of the graph, which is also a tree.
Minimum spanning tree	A spanning tree of a weighted graph that has the minimum total weight.
Complement of a graph G	A graph with the same vertices as G but which has an edge between any two vertices if and only if G does not.
Graph isomorphism between two simple graphs G and H	A one-to-one correspondence between vertices of G and H such that a pair of vertices in G is adjacent if and only if the corresponding pair in H is adjacent.
Planar graph	A graph that can be drawn in the plane without any edge crossing another.
Bipartite graph	A graph whose vertices can be divided into two sets and in which edges always join a vertex from one set to a vertex from the other set.
Complete bipartite graph	A bipartite graph in which every vertex in one set is joined to every vertex in the other set.
Adjacency matrix of G, denoted by A_G	The adjacency matrix, A_G, of a graph G with n vertices, is the $n \times n$ matrix in which the entry in row i and column j is the number of edges joining the vertices i and j. Hence, the adjacency matrix will be symmetric about the diagonal.
Cost adjacency matrix of G, denoted by C_G	The cost adjacency matrix, C_G, of a graph G with n vertices is the $n \times n$ matrix in which the entry in row i and column j is the weight of the edges joining the vertices i and j.

3.1 Introduction



The diagram above is a map of Vienna's underground. Maps like this one do not generally correspond to the real geographic sites in the city but rather the way in which the different stations are organized. This way, a passenger using the underground can plan a route from one station to another. The map as presented is simply a diagrammatic means of representing how the stations are interconnected.

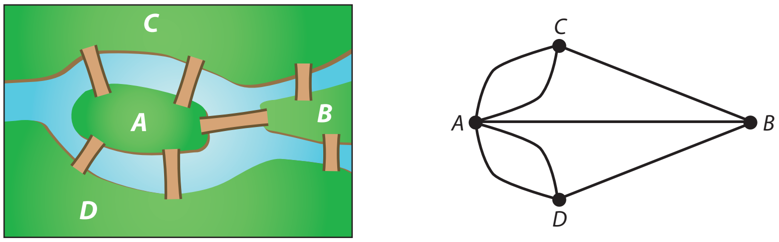
The above situation is one simple application of graph theory. The theory has many applications, including chemical molecules, floor plans, electrical and computer networks, and many others. We will begin with some basic definitions.

3.2 Graphs: definitions

When we are using a map, we are more concerned with seeing how to get from one point to another using the routes available. Consequently, we are dealing with two sets of objects: locations and routes. Such situations involving two sets give rise to relations between the elements of the sets.

If V denotes the set of vertices (also called **nodes** or **points**) and E denotes the set of edges (routes, lines), graph G is the non-empty set consisting of vertices and edges, as shown below.

Figure 3.1



a)

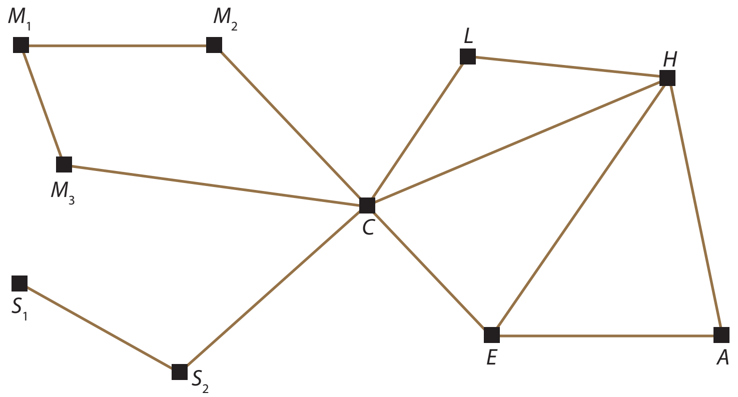
b) Graph G

Related to the discussion above is the Königsberg bridge problem (Figure 3.1a). The Pregel river passes through the Prussian city of Königsberg and divides it into two banks and two islands in the middle. Seven bridges connect the four land areas of the city. Residents of the city had a problem – namely to determine whether it was possible to walk through the city using each of the bridges exactly once.

The Königsberg problem inspired Euler to find a solution which appeared in his paper *Solutio problematis ad geometriam situs pertinentis*, published in 1736. Euler realized that the physical layout of land, water and bridges could be modelled by the graph shown in Figure 3.1b). The land parts are represented by points A, B, C, and D, and the bridges by lines (edges) which could be curved. By means of such a graph, the real problem is transformed into a mathematical one: given the graph in Figure 3.1b), is it possible to choose a vertex, traverse the edges one after the other, and return to the starting vertex using every edge only once? Euler showed that it was impossible. This is a problem we will visit later in the chapter.

Consider Figure 3.2 below, representing a school network. Each computer is connected to the network by one cable. In this network, there is at most one cable between any two computers and there is no cable that connects a computer to itself. This network can be modelled by a **simple graph**, which consists of vertices that represent the computers and undirected edges that represent the cables. Each edge connects two different vertices and no two edges connect the same pair of vertices.

Figure 3.2



Definition 1

A **graph *G* = (*V*, *E*)** consists of two sets: *V*, a non-empty set of **vertices** (**nodes** or **points**), and *E*, a set of unordered pairs of different elements of *V* called **edges** (**arcs** or **sides**).

1 In this publication, all graphs are assumed to be **finite graphs**, which means that they consist of a finite number of vertices and edges.

2 Edges in a graph are allowed to cross each other without intersecting at a vertex. See Figure 3.3 right.

3 A graph with no direction assigned to its edges is **undirected**.

4 **Notation:** Vertices are denoted by single letters or by numbers, so we can say vertex *A* or *a*, or 1, and edges connecting two vertices *u* and *v* by either (*u*, *v*), *u*-*v*, *uv*, or by a single variable such as e_1. See Figure 3.3.

5 A graph where all pairs of adjacent vertices are connected by only one edge are **simple graphs**. The graph in Figure 3.3 is simple.

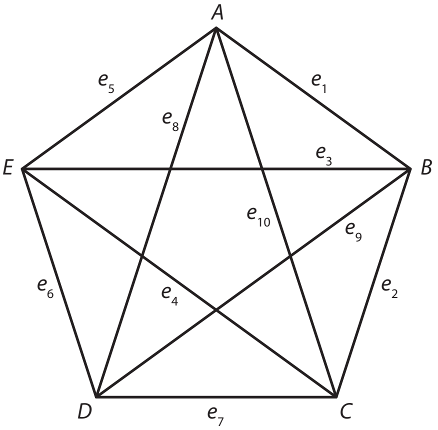


Figure 3.3

Note: In graph theory we do not concern ourselves with the shape of edges or position of the vertices. What is important is which vertices are connected by which edges. The same graph in Figure 3.4 (below) can be represented in different ways, two of which are shown. We consider those two graphs as equivalent.

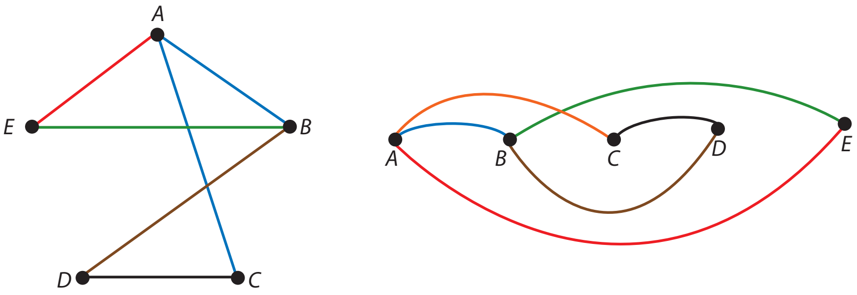


Figure 3.4

Definition 2

1 Two vertices *A* and *B* in an undirected graph *G* are called **adjacent** (or neighbours) if *u* = {*A*, *B*} is an edge in *G*. The edge *u* is said to be **incident** with vertices *A* and *B*. The edge *u* is also said to **connect** *A* and *B*. *A* and *B* are also called the **endpoints** of {*A*, *B*}. Two edges are said to be **adjacent** if they have a vertex in common.

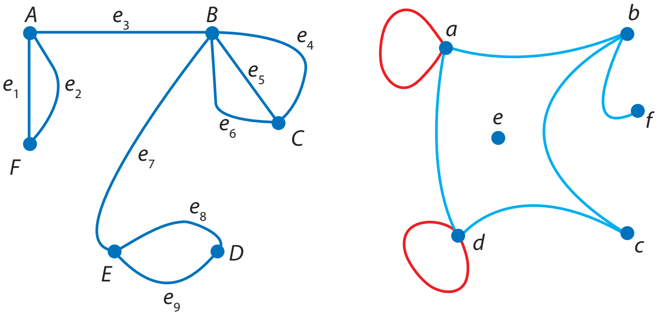
2 If an edge has only one endpoint, then the edge joins the vertex to itself and is called a **loop**.

3 If two edges have the same endpoints, they are called **multiple edges** or **parallel edges**.

4 The degree of a vertex in an undirected graph is the number of edges incident with it. The loop, however, contributes two degrees to the vertex it is incident with. The degree of a vertex *a* is denoted by deg(*a*). A vertex with degree 0, is said to be **isolated** and a vertex with degree 1 is **pendant**. Vertices with odd degrees are called **odd vertices** and those with even degrees are **even vertices**.

Example 1

Identify the elements of the two graphs below.



a) b)

Solution

a) *A* is adjacent to *B* and *F*, while *F* and *B* are not adjacent. *B* is adjacent to *C* and *E* but not to *D*.

e_1 is incident with *F* and *A*, and so is e_2. e_1 and e_2 are multiple (parallel) edges. Also, e_4, e_5, and e_6 are multiple (parallel) edges, as are e_8 and e_9. There are no loops. Deg(A) = 3, deg(B) = 5, and deg(E) = 3. *A*, *B*, *C*, and *E* are odd, while *F* and *D* are even. e_1 and e_3 are adjacent since they have *A* as a common vertex. e_6 and e_7 are also adjacent.

b) *a* and *d* have loops incident with them. Deg(a) = 4, with 2 degrees from the loop! Edges *cd* and *cb* are adjacent since they have vertex *c* in common. Vertex *e* with deg(e) = 0 is isolated while vertex *f* with deg(f) = 1 is pendant.

Now we give a formal definition of a simple graph.

Definition 3

A **simple graph** $G = (V, E)$ is a graph that contains no loops or parallel edges. If there is more than one edge adjacent to two vertices, the graph is called a **multiple graph** or a **multigraph**.

For instance, the graphs in Example 1 above are multigraphs while the graphs in Figures 3.2, 3.3, or 3.4 are simple.

Theorem 1 (The handshaking theorem)

Let $G = (V, E)$ be a graph with *e* edges, i.e. $|E| = e$. Then the sum of all degrees of the vertices in *V* is twice the number of edges. That is,

$$\sum_{v\in V}\deg(v) = 2e.$$

Note: This applies even if the graph is a multigraph.

Proof

Every edge contributes 2 to the sum of the degrees of the vertices, since every edge is incident with exactly two vertices (they may be equal!). So by adding all the vertex degrees we count each edge twice.

For instance, in Example 1, graph a) has 9 edges and $3 + 5 + 3 + 2 + 3 + 2 = 18$ degrees. Graph b) has 7 edges and $4 + 3 + 2 + 4 + 0 + 1 = 14$ degrees.

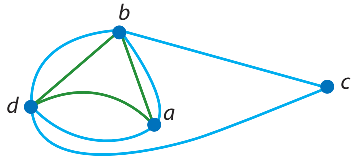
This is called the **Handshaking theorem**, because of the resemblance between an edge having two endpoints and a handshake involving two hands!

Example 2

In a graph with four vertices *a*, *b*, *c*, and *d*, the degrees are as follows: $\deg(a) = 4$, $\deg(b) = \deg(d) = 5$, and $\deg(c) = 2$. Is this graph possible? If yes, draw a representation, and if not, justify why not.

Solution

Since the sum of the degrees is 16, there is a possible graph with $16/2 = 8$ edges. On the right is a demonstration of such a graph.



Theorem 1 gives rise to another important theorem.

Theorem 2

An undirected graph $G = (V, E)$ can only have an even number of odd vertices.

Proof

The degree of a vertex is either odd or even. Let V_O consist of all odd vertices in V, and V_E consist of all even vertices in V.

Since $V = V_O \cup V_E$ and $V_O \cap V_E = \varnothing$, then

$$2e = \sum_{v\in V} \deg(v) = \sum_{v\in V_O} \deg(v) + \sum_{v\in V_E} \deg(v).$$

Since $2e$ is even, the right-hand side of the equation must be even. Also, the even vertices will have an even sum! Thus, the odd vertices can only have an even sum since the sum of odd numbers cannot be even, and since all the terms in this sum are odd, there must be an even number of them. Thus, there is an even number of odd vertices.

Example

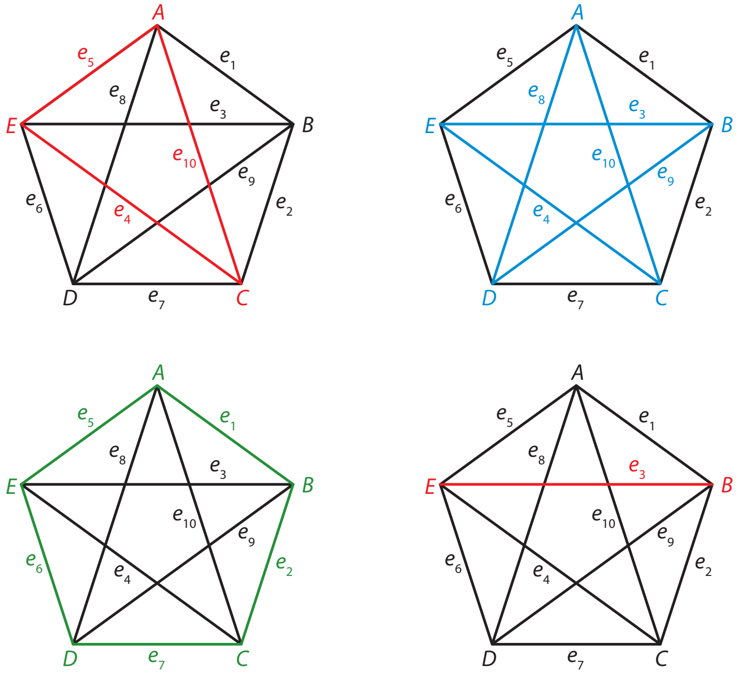
In Figure 3.2, the graph has two odd vertices, S_1 and E; in Figure 3.4, A and B are odd vertices; in Example 1, A, B, C, and E are the odd vertices in graph a), while b and f are odd in graph b); and finally, in Example 2, b and d are the odd vertices.

Definition 4: Subgraphs

Given that $G = (V, E)$ is a **graph**, then, $G_1 = (V_1, E_1)$ is called a **subgraph** of G if $V_1 \subseteq V$, $E_1 \subseteq E$, and $V_1 \neq \varnothing$.

Example

The following are subgraphs of the graph in Figure 3.3. The subgraphs are coloured to distinguish them from the parent one.

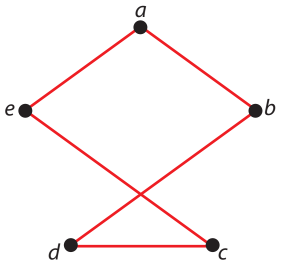


Definition 5: Union (optional)

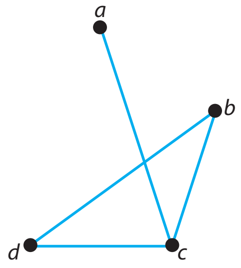
The union of two simple graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ is the simple graph with vertex set $V_1 \cup V_2$ and edge set $E_1 \cup E_2$. The union of G_1 and G_2 is denoted by $G_1 \cup G_2$.

Example 3

Find the union of the graphs G_1 and G_2 shown below.



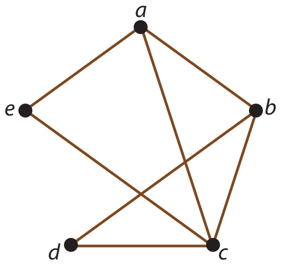
G_1



G_2

Solution

The vertex set of the union $G_1 \cup G_2$ is the union of the two vertex sets. So, $E = E_1 \cup E_2 = \{a, b, c, d, e\}$. The edge set is the union of the two edge sets, i.e. $V = V_1 \cup V_2 = \{ae, ab, ac, bc, bd, cd, ce\}$. The union is displayed on the right.



$G_1 \cup G_2$

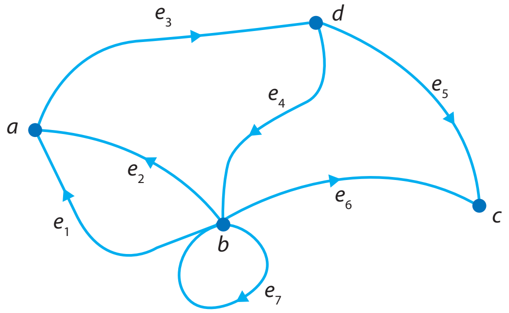
Some special graphs

So far we have only considered **undirected** graphs. Adding direction to edges gives us a new look at a slightly different graph, the **directed graph** or simply **digraph**. The difference from the previous discussion is that edges in a directed graph have directions. That is, for example, the edge *ab* is not the same as the edge *ba*.

Definition 6: Digraphs

A **directed graph** or **digraph** $G = (V, E)$ consists of two sets: V, a non-empty set of **vertices** (**nodes** or **points**) and E, a set of **ordered** pairs of different elements of V called **edges** (arcs or sides).

Here is a representation of a digraph. Notice that the difference from a graph is that each edge e_i is represented by an arrow rather than simply an arc.



G consists of four vertices a, b, c, and d; and seven arcs: $e_1 = (b, a)$, $e_2 = (b, a)$, $e_3 = (a, d)$, $e_4 = (d, b)$, $e_5 = (d, c)$, $e_6 = (c, b)$, and $e_7 = (b, b)$. Each directed arc has an **initial vertex** and a **terminal vertex.** So, e_3 has a as its initial point and d as its terminal point. e_7 is a loop with the same initial and terminal vertex b. e_1 and e_2 are called **parallel** edges since they have the same initial vertex b and terminal vertex a.

Definition 7: Degrees in digraphs

In a digraph, the **in-degree** of a vertex v, $\deg^-(v)$, is the number of edges with v as their terminal vertex. The **out-degree** of v, $\deg^+(v)$, is the number of edges with v as their initial vertex.

Note: According to the definition, a loop contributes one in-degree and one out-degree for the vertex.

In the graph for a digraph on the previous page, for example, $\deg^-(a) = 2$ and $\deg^+(a) = 1$. Also, $\deg^-(b) = 2$ [one degree from e_4 and one from e_7], while $\deg^+(b) = 4$. Moreover, $\deg^-(c) = 2$ and $\deg^+(c) = 0$.

Theorem 3

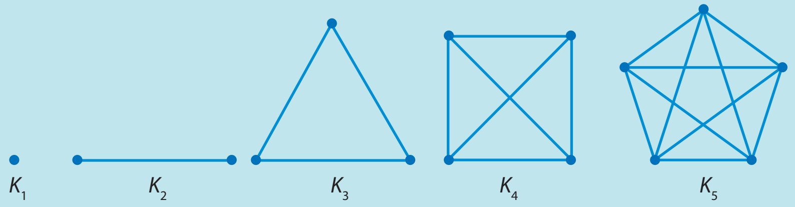
In a digraph $G = (V, E)$, $|E| = \sum_{v\in V}\deg^+(v) = \sum_{v\in V}\deg^-(v)$.

Proof

Since each edge has an initial vertex and a terminal vertex, the sum of the in-degrees is the same as the number of edges. The same is true for the out-degrees.

Definition 8: Complete graphs

A **simple graph** $G = (V, E)$ is called a **complete graph** if for all $a, b \in V$ there is an edge $\{a, b\}$. A complete graph with n vertices is denoted by K_n.
Here are the graphs of K_n, where $n = 1, 2, \ldots, 5$.



Theorem 4

The number of edges in a **complete graph** K_n is given by $|K_n| = \frac{n(n-1)}{2}$.

Proof

The number of vertices is n and each edge connects two vertices; therefore, there are $\binom{n}{2} = \frac{n(n-1)}{2}$ edges.

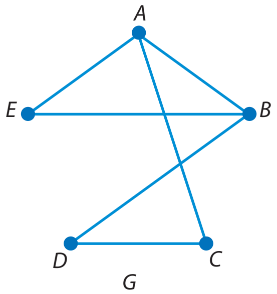
Definition 9: Complement

Let $G = (V, E)$ be a **simple graph**. Then the **complement** of G, denoted by G', is a graph that contains the same set of vertices as the graph G and contains all the edges that are not in G.

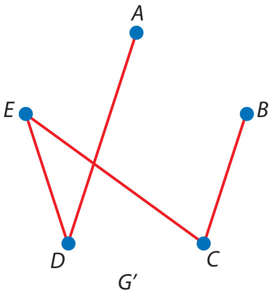
When dealing with sets, the complement of a set A is the set containing the elements of the universal set U that are not in the given set itself. The complete graphs here play a similar role to the universal set. The complement of G which has n vertices is the subgraph of K_n consisting of the n vertices in G and all the edges that *are not* in G. So, two vertices are adjacent in G' if and only if they are not adjacent in G.

Example

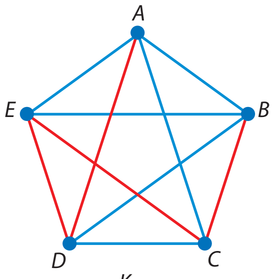
The graph G presented in the figure below is coloured in blue, while G' is coloured in red.



G



G'



K_5

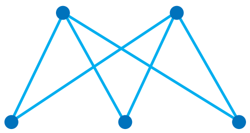
We notice that the graphs G and G' together form a K_5. In some books it is said that those two graphs complement each other to a complete graph.

Another similarity with the complement of a set can be seen here when we look for the complement of K_n. K_n's complement consists of all the vertices and no edges and it is called a **null graph**. This is similar to the case when we look for the complement of U. It is the empty set.

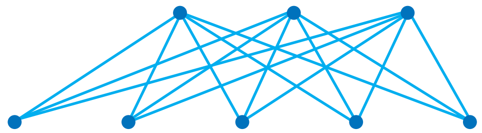
Definition 10: Bipartite graphs

A simple graph $G = (V, E)$ is said to be a **bipartite graph** if the vertex set V can be separated into two subsets V_1 and V_2 such that $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \varnothing$, often called a **partition**, and all the edges for the set E are of the form $\{X, Y\}$ such that $X \in V_1$ and $Y \in V_2$ (no edge in G connects either two vertices in V_1 nor two vertices in V_2).
A bipartite graph is said to be a **complete bipartite graph** if every vertex from V_1 is adjacent to every vertex from V_2. The most common notation of a complete bipartite graph is $K_{m,n}$, where $|V_1| = m$ and $|V_2| = n$.

Here are some examples of complete bipartite graphs.



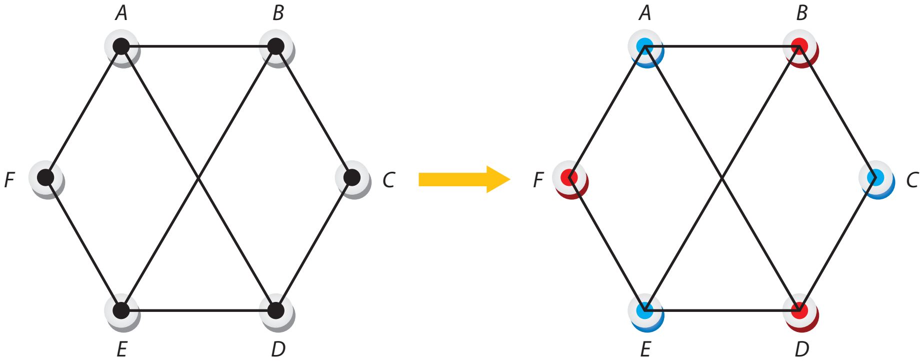
$K_{2,3}$



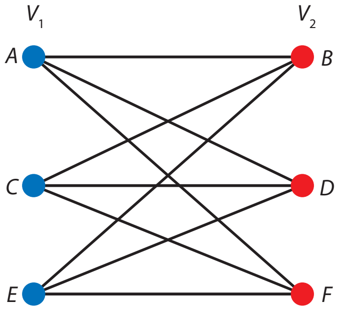
$K_{3,5}$

Example

The graph on the following figure (page 1588) is bipartite. As we carefully investigate it we notice that the vertices can be split into two disjoint sets and no edge connects two vertices from the same set. If we simply colour vertices with different colours (red and blue), we observe that no blue vertex is adjacent to a red vertex; therefore, two possible partitions are $V_1 = \{A, C, E\}$ and $V_2 = \{B, D, F\}$.

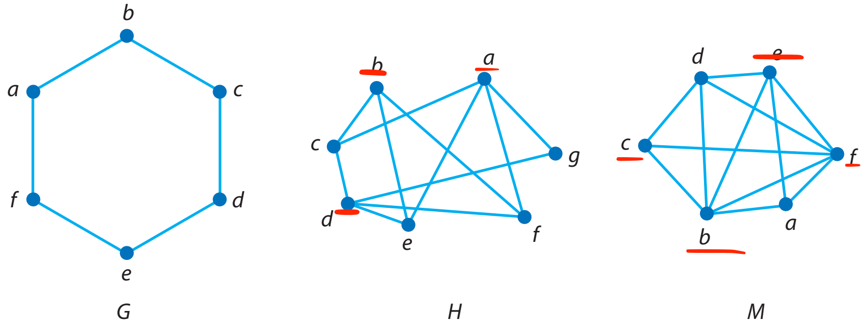


This can be made clearer by rearranging the graph without changing the way the vertices are connected. With this, it becomes obvious that we have a bipartite graph.



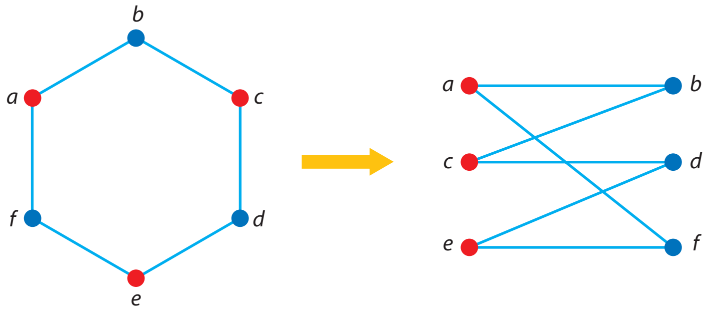
Example 4

Which of the following graphs are bipartite?

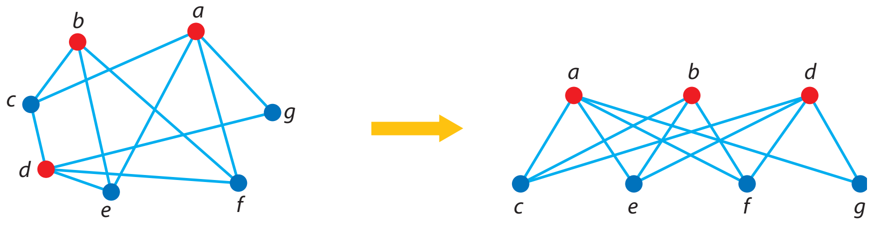


Solution

G: If we colour the vertices with two different colours, we notice that we can do that without any two adjacent vertices sharing a colour. By rearranging the vertices, you can clearly see that we are able to separate them into two sets. So, *G* is bipartite.



H: Doing the same thing here will also yield a bipartite graph.



M: *M* cannot be bipartite. If you consider vertex *b* and vertex *f*, they cannot be in the same subset as they are adjacent. So, they should be in different subsets. Now, *a* can either be in the subset containing *b*, but it cannot since the two are adjacent; or *a* could be in the subset containing *f*, but that cannot happen either.

Exercise 3.1 and 3.2

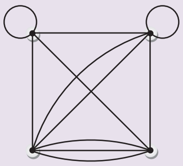
1 For each graph write down:

i the number of vertices

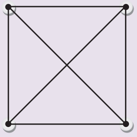
ii the number of edges

iii the degree of each vertex.

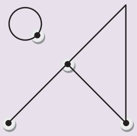
a



b



c



2 Consider a group of 5 people at a party. Is it possible for each of them to chat with:

a 3 other people from the group

b 4 other people from the group?

If possible, represent the solution in the form of a graph.

3 What is the minimum number of edges a simple connected graph with *n* vertices can have?

4 A graph has *n* vertices. What is the number of edges if the graph is complete?

5 Find the number of vertices and edges for the following graphs:

a $K_{3,4}$ **b** $K_{13,17}$ **c** $K_{m,n}$

6 A complete bipartite graph $K_{m,n}$ has altogether 24 vertices and 128 edges. Find the number of vertices in each partition.

7 A graph is called ***r*-regular** if all the vertices have the same degree *r*.

a How many vertices does a 3-regular graph have if it has 12 edges?

b Is it possible to have a regular simple graph with 14 edges? Explain your solution.

c How many regular simple graphs are there with *p* edges, where *p* is a prime number?

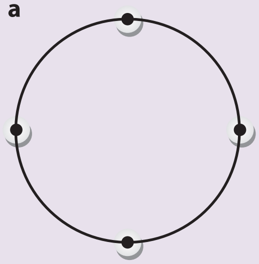
d If the number of edges in a graph is *e* and vertices *v*, show that, if the graph is simple and connected, then $v-1 \leqslant e \leqslant \frac{v(v-1)}{2}$.

8 Show that in a simple connected graph there are at least two vertices of the same degree.

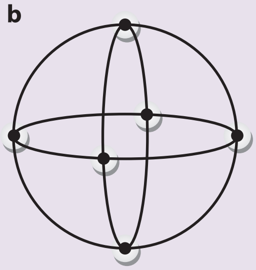
9 Prove that any subgraph of a bipartite graph must be bipartite.

10 Explain which of the following graphs are bipartite:

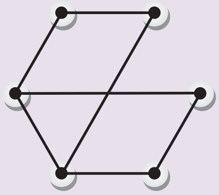
a



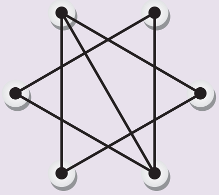
b



c



d



11 A graph with $v = 7$ has the following vertex degrees: 2, 3, 3, 3, 4, 4, 5. What is the number of edges of this graph?

12 In each of the following, determine whether it is possible to have a simple graph. If yes, draw it. If not, explain why not.

a Number of vertices $v = 5$, vertex degrees: 1, 3, 3, 4, 4

b Number of vertices $v = 6$, vertex degrees: 1, 3, 3, 4, 4, 5

c Number of vertices $v = 6$, vertex degrees: 1, 2, 2, 3, 3, 3

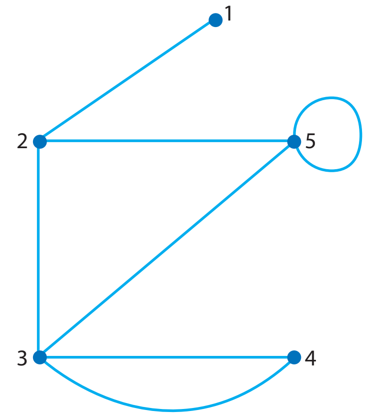
3.3 Graph representation

Diagrams are very helpful and useful in representing graphs and sometimes they are the best way to understand them. However, there are other methods used to represent graphs and a few of these may at times be more convenient. In this section we will see how we can represent graphs in different ways.

Adjacency matrices

For any graph, we can store information about the number of edges connecting each pair of vertices in matrix form. Consider the graph given below with the matrix at the right.

The syllabus does not mention matrices. The term used is 'adjacency tables', which is an equivalent but is not universally used. In this publication, we will continue to use the adjacency matrix notation. In some cases, the table (matrix) may use 'T' for 1 and 'F' for 0.



$$\begin{array}{cc}
 & \begin{matrix} 1 & 2 & 3 & 4 & 5 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{matrix} \\
\begin{matrix} 1 \rightarrow \\ 2 \rightarrow \\ 3 \rightarrow \\ 4 \rightarrow \\ 5 \rightarrow \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}
\end{array}$$

Every row corresponds to a vertex and every column corresponds to a vertex too. The entries in each row correspond to the number of edges connecting that vertex to the vertices represented by the columns. For example, row 1 has only 1 in the second entry. This is because there is one edge connecting vertex 1 to vertex 2. Row 3, for another example, has 0 in the first entry because vertex 3 has no edges with vertex 1 (i.e. they are not adjacent), has 1 in entry 2 because there is 1 edge connecting vertex 3 to vertex 2, and has 2 in entry 4 because there are 2 edges connecting vertices 3 and 4. Notice that row 5 has an entry corresponding to column 5, because there is a loop at vertex 5.

The following definition formalizes the idea and introduces some notation.

Definition 11

The **adjacency matrix A_G** of a **simple graph** $G = (V, E)$ with n vertices is an $n \times n$ matrix containing 1 or 0 in such a way that any entry of the matrix

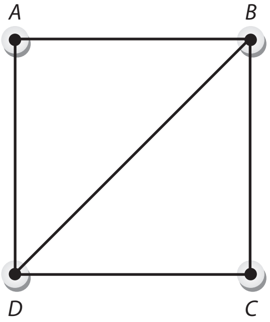
$$a_{i,j} = \begin{cases} 1 & \text{if } \{v_i, v_j\} \text{ is an edge from } E \\ 0 & \text{otherwise} \end{cases}$$

Note: For a multigraph, the definition can be adjusted to reflect the fact that there could be more than one edge between two vertices. So, for a multigraph, we can say that the adjacency matrix has the property

$$a_{i,j} = \begin{cases} k(i, j) & k = \text{number of edges between } v_i \text{ and } v_j \\ 0 & \text{otherwise} \end{cases}$$

Example 5

a) Use an adjacency matrix to represent the given graph.



b) Draw a graph represented by the given adjacency matrix.

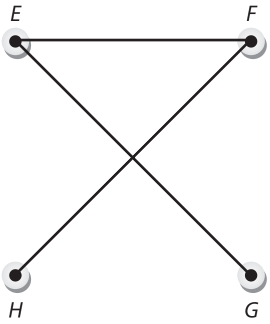
$$B_G = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Solution

a)

$$A_G = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

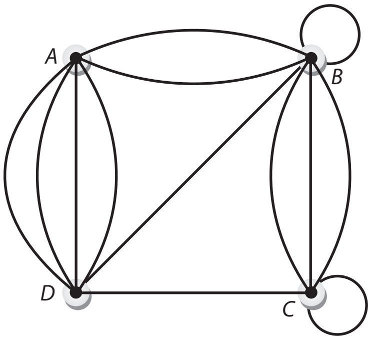
b)



We notice that in a simple graph all the entries on the main diagonal of its adjacency matrix are 0. This is the case since there are no loops in a simple graph. The matrix is also symmetric about its main diagonal since the simple graph is not a digraph, and thus when there is an edge between v_i and v_j this contributes 1 to the (i, j) entry. Similarly, when the edge is between v_j and v_i, this contributes 1 to the (j, i) entry. In the case of a multigraph that contains loops and multiple edges, the entries on the leading diagonal will be 1 if there is a loop at that vertex, whilst multiple edges will contribute correspondingly to a non-diagonal, and hence the matrix may not be symmetric.

Example 6

Use an adjacency matrix to represent the following multigraph.



Solution

$$A_G = \begin{pmatrix} 0 & 2 & 0 & 4 \\ 2 & 1 & 3 & 1 \\ 0 & 3 & 1 & 1 \\ 4 & 1 & 1 & 0 \end{pmatrix}$$

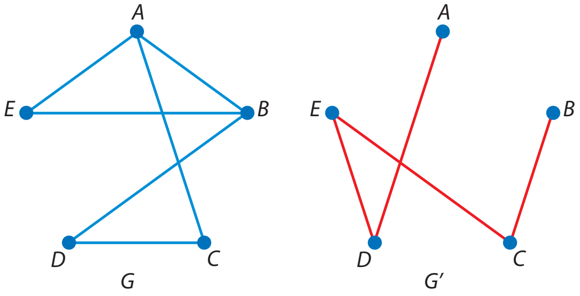
Note: We notice that adjacency matrices of complete graphs have all entries equal to 1 except on the main diagonal where they are all 0. For example,

the adjacency matrix of K_3 is $A_{K_3} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

Note: The adjacency matrices of complementary graphs each have the main diagonal as 0, but all the other entries are complementary 1 and 0. That means whenever there is a 1 in one matrix it is 0 in the other matrix and vice versa, apart from the main diagonal, of course. When we add them we obtain an adjacency matrix of a complete graph.

Example 7

Consider the graphs G and G' below and write their adjacency matrices.



Solution

$$G \Rightarrow \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}, \text{and } G' \Rightarrow \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

Incidence matrices (Optional)

Another way that can be helpful in comparing different graphs to check if they have similar structures is the **incidence matrix**. The incidence matrix consists of n rows corresponding to the vertices that a graph has, and k columns corresponding to the edges that this graph has. The matrix will have a 1 in the entry (i, j) if the edge e_j is incident with the vertex v_i.

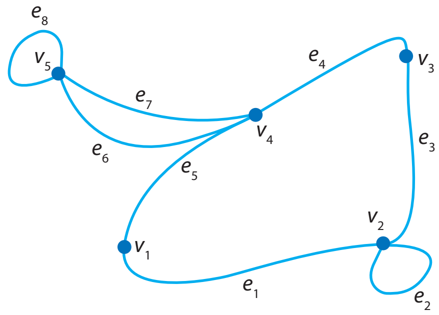
Definition 12

The **incidence matrix** $\boldsymbol{I}_G$ of a **simple** graph $G = (V, E)$ with n vertices and k edges is an $n \times k$ matrix containing 1 or 0 in such a way that any entry of the matrix

$$a_{i,j} = \begin{cases} 1 & \text{if } e_j \text{ is incident with } v_i \\ 0 & \text{otherwise} \end{cases}$$

Example 8

Represent the graph shown below with an incidence matrix.



Solution

$$\begin{array}{c|ccccc|c|c|c|} & e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 & e_8 \\ v_1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ v_2 & 1 & \boxed{1} & 1 & 0 & 0 & 0 & 0 & 0 \\ v_3 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ v_4 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ v_5 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & \boxed{1} \end{array}$$

Notice how multiple edges are represented by columns with identical entries while loops are the only columns with exactly one entry equal to 1.

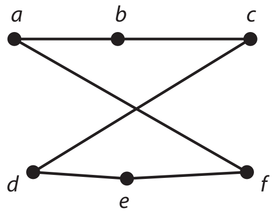
In the case of simple graphs, the row totals give the degree of each vertex of the graph. In multigraphs, however, the entries corresponding to loops should be multiplied by 2 to give the degree of the vertex involved.

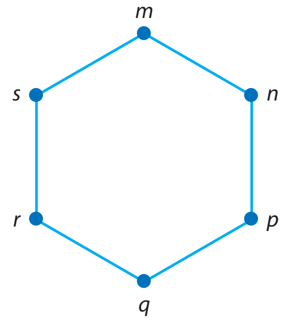
Isomorphic graphs

Isomorphic comes from the Greek words ***iso*** (the same as) and ***morphe*** (form).

Try the following experiment with two of your classmates:
Give one of them the following instructions: 'Draw and label the six vertices *a*, *b*, *c*, *d*, *e*, and *f* of a graph *G*. Now connect *a* to *b*, *c* to *b*, *c* to *d*, *d* to *e*, *f* to *e*, and *a* to *f*.' Now give the other the following instructions: 'Draw and label the six vertices *m*, *n*, *p*, *q*, *r*, and *s* of a graph *H*. Now connect *m* to *n*, *n* to *p*, *p* to *q*, *q* to *r*, *r* to *s*, and *s* to *m*.'

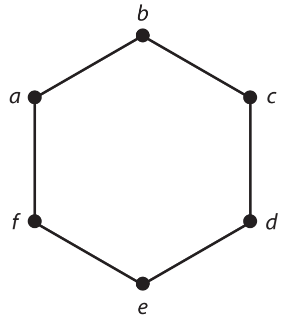
An experiment that was performed in one class produced the following two graphs.

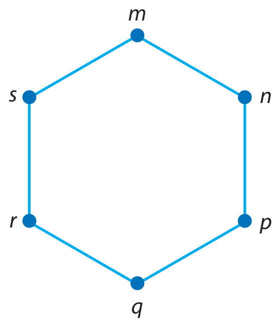




Although the syllabus does not include isomorphic graphs, we will still use them here because they help to make some operations more efficient. Obviously this will not jeopardize your chances of earning marks. All sound mathematical methods are acceptable in exams. Moreover, 'isomorphism' is still on the list of terms in the syllabus.

You may have noticed already that these two graphs define the same situation. However, they appear to be different. If we rearrange the way we graphed them, you will see that they are equivalent. Here is a rearrangement.





Such graphs are said to be **isomorphic**. You can set up a one-to-one correspondence between the vertices of the two graphs, keeping the adjacent vertices in one graph and the images of the adjacent vertices in the other. For example, here we can match *a* with *m*, *b* with *n*, and so on. This way, any two vertices that are adjacent in one graph have their images adjacent in the same way. We say that the two graphs have the same structure.

One-to-one correspondence means that the function f is **surjective** and **injective**, i.e. it is a **bijection**.

Definition 13

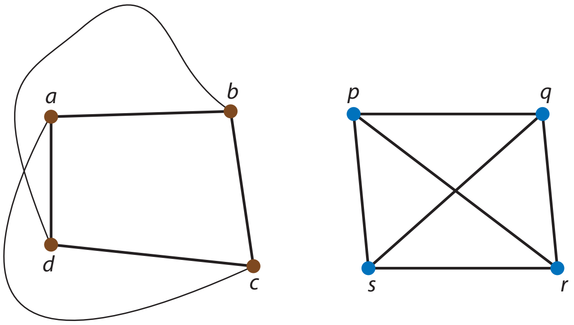
Let $G = (V, E)$ and $G' = (V', E')$ be two **simple graphs**. If there is a one-to-one correspondence $f: V \to V'$, such that for every pair of vertices v_i and v_j that are adjacent in the graph G vertices $f(v_i)$ and $f(v_j)$ are adjacent in G', then the graphs G and G' are said to be **isomorphic**. The function f is called a **graph isomorphism**.

Stated differently, when two graphs are isomorphic, there is a bijection between the vertices of the two graphs that *preserves* the *adjacency* association. In the previous example, the bijection could be defined by

$$g(a) = m, g(b) = n, g(c) = p, g(d) = q, g(e) = r, \text{ and } g(f) = s.$$

Example 9

Consider the graphs G and H given below. Examine whether the two graphs are isomorphic.



Solution

We set up the following function: $f(a) = p, f(b) = q, f(c) = r, f(d) = s$.

This function preserves adjacency as is easily verified, and hence it is an isomorphism. Take the adjacent vertices a and b, for example, $f(a) = p$ is adjacent to $f(b) = q$. The rest can clearly be seen.

Hence, the two graphs can be considered the same, as far as graph structure is concerned.

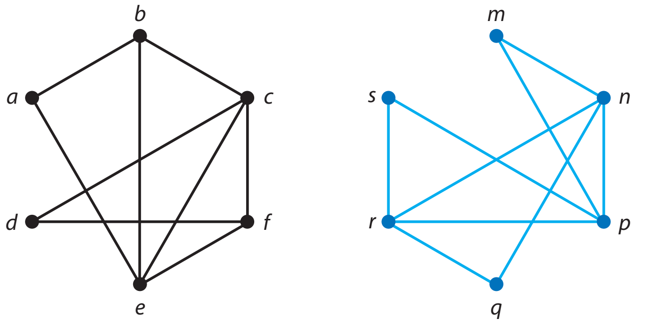
Note: If we set up the adjacency matrices for the two graphs above, we get:

$$\begin{array}{c|cccc} & a & b & c & d \\ \hline a & 0 & 1 & 1 & 1 \\ b & 1 & 0 & 1 & 1 \\ c & 1 & 1 & 0 & 1 \\ d & 1 & 1 & 1 & 0 \end{array} \Leftrightarrow \begin{array}{c|cccc} & p & q & r & s \\ \hline p & 0 & 1 & 1 & 1 \\ q & 1 & 0 & 1 & 1 \\ r & 1 & 1 & 0 & 1 \\ s & 1 & 1 & 1 & 0 \end{array}$$

It is important to note that when you arrange the matrices of two isomorphic graphs in such a way that the corresponding vertices occupy the same rows and columns, the adjacency matrices of both are identical, as you see above.

Example 10

Consider the following two graphs and examine whether they are isomorphic.



Solution

If we consider the adjacency matrices for both, we get:

$$\begin{array}{cc} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \text{ and } \begin{array}{cc} & \begin{matrix} m & n & p & q & r & s \end{matrix} \\ \begin{matrix} m \\ n \\ p \\ q \\ r \\ s \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Since these graphs are simple, then the column/row totals are the degrees of each vertex. We can clearly see that the degree sequence of the first graph is 2, 2, 3, 3, 4, 4, while the second graph is 2, 2, 2, 4, 4, 4. This means that we cannot set up a correspondence to preserve adjacency, and hence the two graphs are not isomorphic.

The **degree sequence** of a graph is the list of degrees of the vertices of the graph, listed from smallest (largest) degree to largest (smallest).

Example 10 leads us to the following theorem.

Theorem 5

Let $G = (V, E)$ and $G' = (V', E')$ be two **isomorphic graphs** and $f{:}V \to V'$ a **graph isomorphism**. If a is any vertex from set V, then $\deg(a) = \deg(f(a))$.

Stated differently, corresponding vertices in an isomorphism must have the same degree.

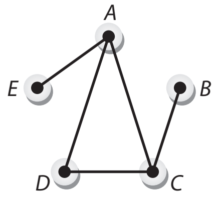
Proof

Assume that $\deg(a) \neq \deg(f(a))$, then we have two cases to consider.

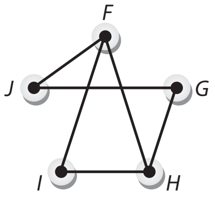
The first case is $\deg(a) > \deg(f(a))$, and if this is true, then there is a vertex b such that b is adjacent to a in G, but $f(b)$ is not adjacent to $f(a)$, which is a contradiction to the definition of **graph isomorphism** f. A similar argument is true for the case when $\deg(a) < \deg(f(a))$. Therefore, $\deg(a) = \deg(f(a))$.

Example 11

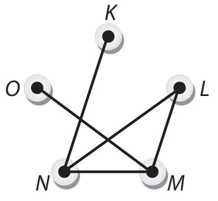
Determine which pairs of graphs are isomorphic.



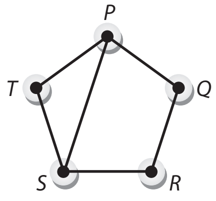
1



2



3



4

Solution

Looking at a table showing the degrees of the corresponding vertices of the graphs, we can try to construct a graph isomorphism.

Graph	1					2					3					4				
Vertex	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
Degree	3	1	3	2	1	3	2	3	2	2	1	2	3	3	1	3	2	2	3	2

Obviously graphs **1** and **3** have the same degree sequence: 1, 1, 2, 3, 3. Therefore, we would proceed in trying to find an isomorphism between them.

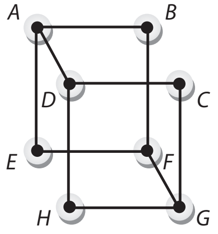
One possible isomorphism between **1** and **3** is $f(A) = M$, $f(B) = K$, $f(C) = N$, $f(D) = L$, $f(E) = O$.

Note that we have to be careful with respect to the vertices with degree one because if we assign $f(A)$ to M then we must assign $f(E)$ to O, since A and E are adjacent in **1**. Another alternative is to assign $f(E)$ to K which would give us a contradiction to Theorem 5, since M and K are not adjacent in graph **3**.

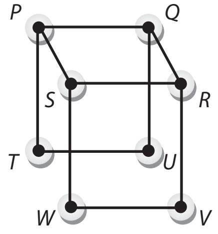
Similarly, **2** and **4** have the same degree sequence: 2, 2, 2, 3, 3. An isomorphism between graphs **2** and **4** could be $g(F) = P$, $g(G) = Q$, $g(H) = S$, $g(I) = T$, $g(J) = R$. Again, here we need to be careful not to assign two adjacent vertices of degree 2 in graph **4** to vertices I and J in **2** which are not adjacent. If we do, we will be violating Theorem 5's conclusion.

Example 12

Determine whether the following pair of graphs are isomorphic.



Graph G



Graph H

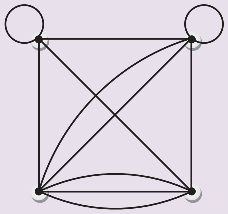
Solution

These two graphs are not isomorphic even though they have an equal number of vertices of degree 2, as well as degree 3. The problem arises with the fact that in graph *G* all the vertices of degree 2 are adjacent only to vertices of degree 3, while in graph *H* all the vertices of degree 2 are connected to one vertex of degree 3 and one of degree 2. Let's take one such pair, for example, *B* and *U*. Both have a degree of 2. *B* is adjacent to vertices *A* and *F* both of which are of degree 3, while *U* in graph *H* is adjacent to *Q* with degree 3 and *T* with degree 2. A function that matches vertex *B*, for example, to vertex *U* will have to match *A* and *F* to *T* and *Q*. Since *A* and *F* have degree 3, one of them will be matched with *T* which is of degree 2. This will contradict Theorem 4. Any attempt to set up a correspondence will meet the same obstacle, and therefore there is no isomorphism between graphs *G* and *H*.

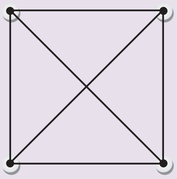
Exercise 3.3

1 For each graph, write down its adjacency matrix.

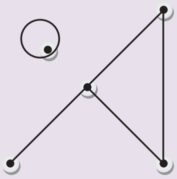
a



b



c

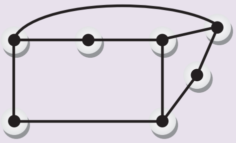


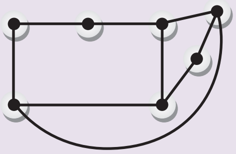
2 Draw the graph for each adjacency matrix and determine pairs of isomorphic graphs.

a $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ **b** $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ **c** $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$

d $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$ **e** $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ **f** $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$

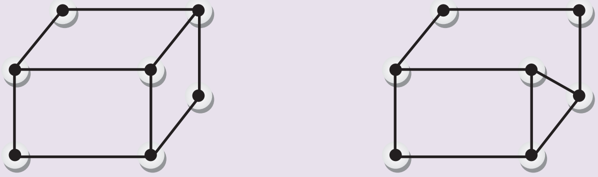
3 Determine whether the following graphs are isomorphic. Explain your answer.



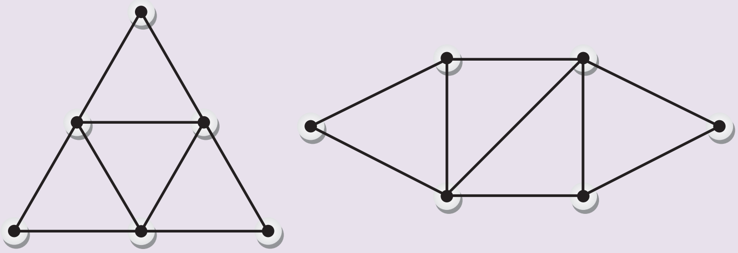


4 Determine whether the following pairs of graphs are isomorphic.

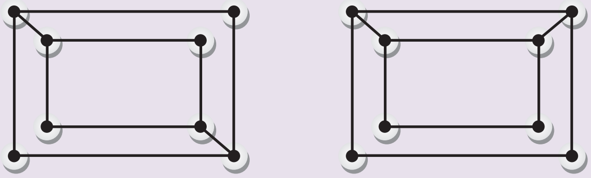
a



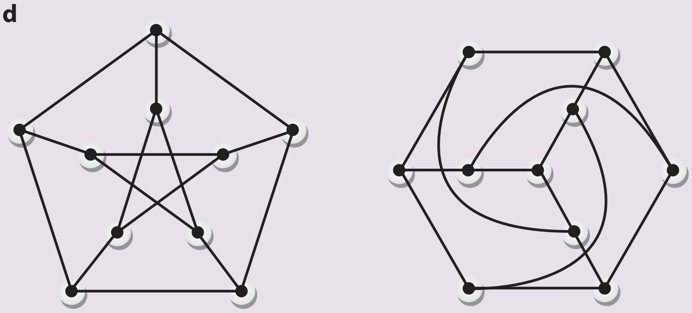
b



c



d



5 Draw two non-isomorphic graphs with three vertices and two edges. How many such non-isomorphic graphs are possible?

6 Draw two non-isomorphic graphs with four vertices and three edges. How many such non-isomorphic graphs are possible?

7 Draw all possible non-isomorphic simple regular graphs with four vertices.

3.4 Paths, walks and trails

Many of the applications of graph theory have to do with paths formed by travelling along the edges of graphs. The example of the Königsberg bridges (page 1580) is one of the oldest. Some current applications include network links, how messages travel between different nodes, postal routes, refuse collection, etc.

We will start this section by stating a few additional necessary definitions.

Definition 14: Walks

Let $G = (V, E)$ be a **graph**. A **walk** is a sequence of alternating vertices and edges that starts and ends with a vertex and where each edge is adjacent to its neighbouring vertices. Stated slightly differently, a $v_0 - v_n$ walk in G is a finite alternating sequence

$$v_0, e_1, v_1, e_2, v_2, \ldots, e_{n-1}, v_{n-1}, e_n, v_n$$

of vertices and edges starting at vertex v_0 and ending at vertex v_n and involving the n edges

$$e_i = \{v_{i-1}, v_i\}, \text{ where } 1 \leqslant i \leqslant n.$$

v_0 and v_n do not have to be different.

The **length of a walk**, n, is the number of edges used in the sequence.

Note: A walk may repeat both edges and vertices.

Note: Like several things in graph theory, unfortunately there is still no unique way of labelling walks. For example, if a graph G has the set of vertices $V = \{a, b, c, \ldots\}$, then a walk can be described as

$$a, \{a, b\}, b, \{b, c\}, \ldots$$

or simply as

$$\{a, b\}, \{b, c\}, \ldots$$

or as

$$a, b, c, \ldots$$

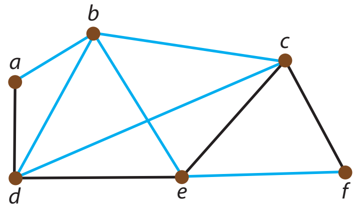
or as

$$abc\ldots$$

We will use the following example to introduce slight variations to the above definition.

Example

Consider the graph below.



The blue coloured walk is the **walk** *abdcbef*. Notice here that vertex b has been visited twice. The length of this walk is 6. No edge has been visited more than once.

The walk *abdcedb* has a length of 6 and uses the edge *bd* twice and the vertices b and d are used twice.

A walk like the first one is known as a **trail**.

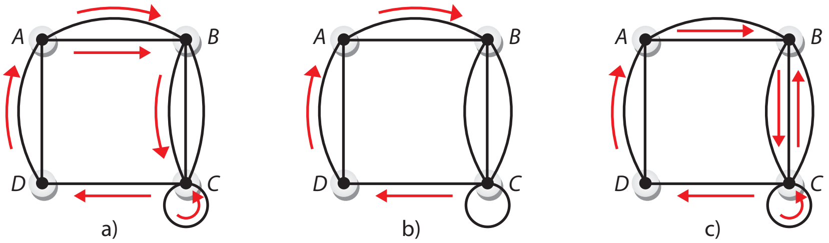
Definition 15

1 A **trail** is a walk in which no edge appears more than once. A trail (like *abcebda*) which begins and ends at the same vertex is called a **circuit**.

2 A walk (like *abef*) where no vertex is visited more than once is called a **path**. A path (like *abceda*) which begins and ends at the same vertex is called a **cycle**.

Example 13

Determine whether each sequence shown is a walk, a path or a trail.



a) $A, \{A, B\}_{\text{lower}}, B, \{B, C\}, C, \{C, C\}_{\text{loop}}, C, \{C, D\}, D, \{D, A\}, A, \{A, B\}_{\text{upper}}, B$

b) $C, \{C, D\}, D, \{D, A\}, A, \{A, B\}_{\text{upper}}, B$

c) $C, \{C, C\}_{\text{loop}}, C, \{C, D\}, D, \{D, A\}, A, \{A, B\}_{\text{upper}}, B, \{B, C\}_{\text{middle}}, C, \{C, B\}_{\text{middle}}, B$

Solution

a) The sequence is a trail since no edge has been repeated. Starting at vertex A to vertex B we used the lower edge, while at the end of the sequence again from vertex A to vertex B we used the upper edge. This sequence cannot be a path since vertices C, A, and B have been repeated.

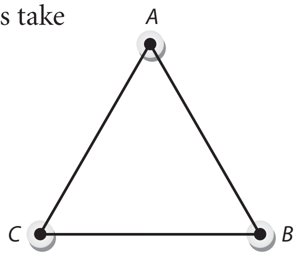
b) The sequence is a path since no vertex has been repeated.

c) The sequence is a walk, since it cannot be a trail as the middle edge from B to C has been repeated twice.

Note: Every path is a trail, while a trail can be a path only in a simple graph.

Adjacency matrices and walks

Adjacency matrices can be very useful in determining the number of possible walks in a graph. Let's take a K_3 and its adjacency matrix for example.



$$A_{K_3} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

The adjacency matrix also represents walks of length 1.

How many different walks of length 2 can we have in K_3?

We observe that since this graph is regular, all the vertices will be treated equally. Start walking from A and note where we can arrive after travelling through two edges:

A **regular graph** is a graph where all vertices have the same degree.

$A, \{A, B\}, B, \{B, C\}, C \qquad A, \{A, B\}, B, \{B, A\}, A$

$A, \{A, C\}, C, \{C, B\}, B \qquad A, \{A, C\}, C, \{C, A\}, A$

We notice that two walks of length 2 will end up back at A, while only one walk of length 2 will end up at B or C.

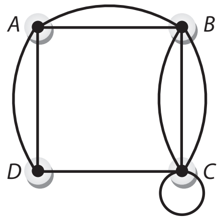
Now, look at the square of the adjacency matrix:

$$A_{K_3}^2 = A_{K_3} \cdot A_{K_3} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

We notice that the entries are the number of walks of length 2 in K_3. Two walks from each vertex back to the same vertex and one walk from each vertex to each of the other two.

Example

Consider the multigraph given right.



Its adjacency matrix is $A_G = \begin{pmatrix} 0 & 2 & 0 & 2 \\ 2 & 0 & 3 & 0 \\ 0 & 3 & 1 & 1 \\ 2 & 0 & 1 & 0 \end{pmatrix}$ and the square of the matrix is $A_G^2 = \begin{pmatrix} 8 & 0 & 8 & 0 \\ 0 & 13 & 3 & 7 \\ 8 & 3 & 11 & 1 \\ 0 & 7 & 1 & 5 \end{pmatrix}$.

Here, for example, the matrix suggests that there are eight walks of length 2 from A to C. We will not list them, we will just explain how to find them. There are 2 edges from A to B and then 3 edges to get from B to C. Therefore, by the counting principle, there are $3 \times 2 = 6$ walks from A to C through B. On the other hand, there are 2 edges from A to D and only 1 edge from D to C. Therefore, there are 2 ways from A to C through D. Now, the total number of walks from A to C is then $6 + 2 = 8$, which is suggested by the matrix. On the other hand, it looks like there are so many walks of length 2 from C back to itself. There are 3 edges to B and 3 edges back, and therefore nine walks through B altogether. There is only one walk to D and back. At the end there is a loop at C; therefore, if we go through it twice that is the last possible walk, which sums up to 11.

To summarize both generalizations we will state the following theorem.

Theorem 6

Let G be a graph containing v vertices and A_G be its adjacency matrix. The number of walks of length n from vertex v_i to v_j is given by the (i, j)th entry of A_G^n, $n \in \mathbb{Z}^+$.

Proof

We will conduct the proof by using mathematical induction on n.

Basis step: Every entry in the adjacency matrix is the number of edges from A_i to A_j; therefore, walks of length 1. The statement is true for $n = 1$.

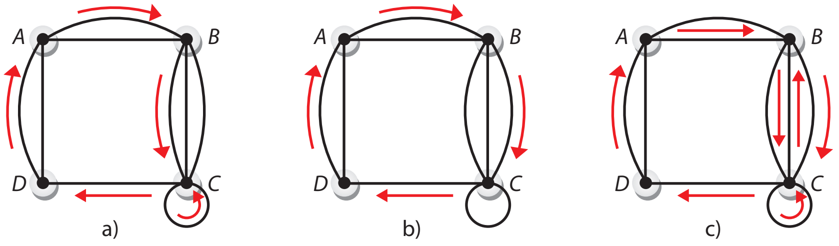
Inductive step: We will assume that every entry of matrix A_G^k is the number of walks of length k between two vertices. Since $A_G^{k+1} = A_G^k \cdot A_G$ then the (i, j)th entry of the matrix A_G^{k+1} is calculated in the following way:

$c_{ij} = b_{i1} \times a_{1j} + b_{i2} \times a_{2j} + \ldots + b_{in} \times a_{nj}$, where b_{ik} is the number of walks of length k from vertex v_i to v_k, and a_{ki} is the number of walks of length 1 from vertex v_k to v_j, giving the total number of walks of length $k + 1$ from vertex v_i to v_j through the vertex v_k. When we add up all the walks from vertex v_i to v_j through different vertices v_k, we get the total sum of all possible walks of length $k + 1$ from vertex v_i to v_j.

Conclusion: Since the statement is true for $n = 1$ and $S(k) \Rightarrow S(k + 1)$, by the principle of mathematical induction, we can conclude that the statement is true for all $n \in \mathbb{Z}^+$.

Example 14

Determine whether each sequence shown below is a closed walk, a cycle or a circuit.



a) $C, \{C, C\}_{\text{loop}}, C, \{C, D\}, D, \{D, A\}, A, \{A, B\}_{\text{upper}}, B, \{B, C\}_{\text{left}}, C$

b) $D, \{D, A\}, A, \{A, B\}_{\text{upper}}, B, \{B, C\}_{\text{right}}, C, \{C, D\}, D$

c) $A, \{A, B\}_{\text{lower}}, B, \{B, C\}_{\text{middle}}, C, \{C, B\}_{\text{right}}, B, \{B, C\}_{\text{middle}},$
$C, \{C, C\}_{\text{loop}}, C, \{C, D\}, D, \{D, A\}, A$

Solution

a) The sequence is a circuit since it is closed and no edge has been repeated. This sequence cannot be a cycle because of the loop at C.

b) The sequence is a cycle since it is closed and no vertex has been repeated.

c) The sequence is a closed walk, since it cannot be a circuit as the middle edge from B to C has been repeated twice.

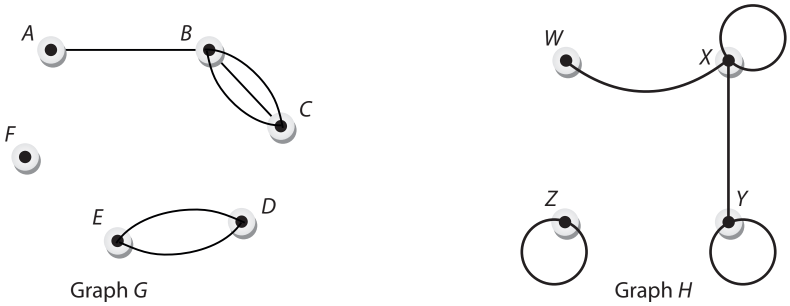
Every cycle is always a circuit, while a circuit can be a cycle only in a simple graph.

Definition 16

Let V be a non-empty set of **vertices** and E be a non empty set of **edges**. The graph $G = (V, E)$ is called a **connected graph** if there is a **path** between any two vertices from the set V.

Example

The graphs presented by all the figures so far are connected. The following graphs G and H are not connected since they contain vertices or even subgraphs that are isolated. Note that in the case of the vertex Z, even though it is isolated, the degree is not equal to zero unlike the vertex F.



Graph G

Graph H

The graphs G and H have the following adjacency matrices:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We notice that the adjacency matrix of a graph that is not connected contains only zeroes in a row and a column of the isolated vertex, or contains only one 1 at the diagonal position in that row or column. On the other hand, disconnected subgraphs can be shown as diagonal matrices where all the other entries are zeroes.

Properties of connected graphs

We will state some properties of connected graphs that will be helpful in later discussions. However, they are not required for examination purposes and their proofs are not supplied in this publication.

Property 1

Let $G = (V, E)$ be a **simple connected graph**, and let a and b be two vertices in G that are not adjacent. If a graph G_1 is formed by *adding* the edge ab to G, then G_1 has a cycle that contains the edge ab.

Property 2

When an edge is removed from a cycle in a connected graph, the result is a graph that is still connected.

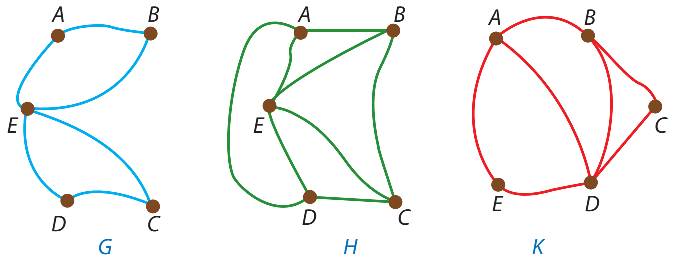
Eulerian graphs

Definition 17

Let $G = (V, E)$ be a **connected graph**. A **trail** where every edge of G appears only once is called an **Eulerian trail**. A **circuit** where every edge of G appears only once is called an **Eulerian circuit**. A connected graph with an Eulerian circuit is called an **Eulerian graph**.

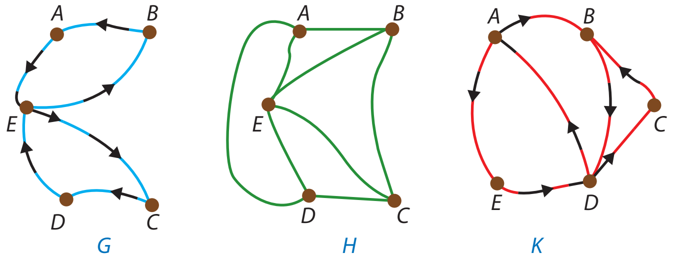
Example 15

Which of the undirected graphs below have an Eulerian circuit? Which have an Eulerian trail only?



Solution

Graph G has an Eulerian circuit. Look at $AECDEBA$, for example.



You can verify that H has neither an Eulerian circuit nor trail. You will be able to confirm this later in the chapter.

Graph K does not have an Eulerian circuit, but it has an Eulerian trail, *AEDCBDAB*.

Theorem 7

Let $G = (V, E)$ be a **connected graph**. G has an **Eulerian circuit** if and only if every vertex has an *even* degree.

Proof

(⇒) Suppose G has an Eulerian circuit. This means the circuit starts at a vertex v_0 (say) and continues with an edge v_0v_1 incident to it, and carries on with the rest of the vertices until it gets back to v_0, i.e. $v_0, v_1, v_2, v_3, \ldots, v_{n-1}, v_0$. Now, v_0v_1 contributes one degree to v_0 and one degree to v_1, but v_1v_2 contributes another degree to v_1, which implies that the circuit contributes two degrees to every vertex it visits. Also, $v_{n-1}v_0$ contributes another degree to v_0, making the total for v_0 at least 2 degrees. Thus, the degree of every vertex, including v_0, is an even integer.

(⇐) Conversely (a short argument that can be expanded), if we assume that each vertex has an even degree, then the circuit can visit each vertex through one edge and leave it using another unused edge. Thus, we can form an Eulerian circuit since the graph is connected.

Example

Refer to Example 15. Graph G has $\deg(A) = \deg(B) = \deg(C) = \deg(D) = 2$, and $\deg(E) = 4$. That is why G is Eulerian.

Graph H has $\deg(A) = \deg(B) = \deg(C) = \deg(D) = 3$, and $\deg(E) = 4$. Only one of the vertices is even while the rest are all odd; thus it cannot be Eulerian.

Graph K has $\deg(C) = \deg(E) = 2$, and $\deg(D) = 4$, while $\deg(A) = \deg(B) = 3$. This is why it does not have an Eulerian circuit. We know however that it has an Eulerian trail. This can be confirmed using the following theorem.

Theorem 8

Let $G = (V, E)$ be a **connected graph**. G has an **Eulerian trail** but not an Eulerian circuit, if and only if it has *exactly two* vertices of *odd* degree.

Proof

(⇒) Suppose G has an Eulerian trail. This means the trail starts at a vertex v_0 (say) and continues with an edge v_0v_1 incident to it, and carries on with the rest of the vertices until it gets to v_n, i.e. $v_0, v_1, v_2, v_3, \ldots, v_n$. Now, v_0v_1 contributes one degree to v_0 and one degree to v_1, but v_1v_2 contributes another degree to v_1, which implies that the trail contributes two degrees to every (internal) vertex it visits. However, since it stops at v_n, then it only contributes one degree to v_n. Thus, the degree of exactly two vertices is odd.

($\Leftarrow$) Conversely, suppose G contains exactly two vertices of odd degree, say v_0v_n. Now, add a new (auxiliary) edge v_0v_n to the graph and the result will be a new graph G_1 with all even degrees. Hence, G_1 has an Eulerian circuit. Removing the auxiliary edge from the circuit leaves you with a trail.

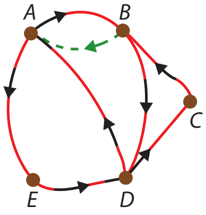
Note: An Eulerian trail must begin and end with a vertex of odd degree!

This is an informal approach to Theorems 7 and 8.

Consider an Eulerian walk W as a sequence of edges $e_1e_2e_3, \ldots, e_n$. Consider a vertex v. Each edge incident with v is used exactly once in the walk. Say v is not the first or last vertex of the walk. Let's walk along W. Each time we arrive at v, say along edge e_i, we must leave along edge e_{i+1}. Thus, each time we visit v we use two edges. Say the number of times we visit v is k. Then v has degree $2k$, an even number. What if v is the first or last vertex? Then the same reasoning applies except for the first or last edge in the walk. If the walk is closed (circuit), then the first and last edge both visit v and we still have an even number. If the walk is open (trail), then either the first or last edge visits v, but not both and we see that v has an odd degree. Thus, the first and last vertices of W have odd degree and we have two vertices of odd degree.

Example

Consider the graph K in Example 15.

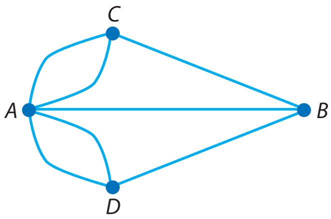


K

By adding an edge BA, we are able to have the circuit $AEDCBDABA$. By removing the edge BA, we get the trail $AEDCBDAB$.

Example 16

Consider the Königsberg bridge problem again (page 1580).
Can we solve it?



Solution

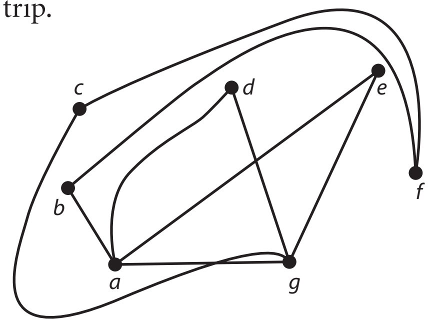
Notice here that deg(B) = deg(C) = deg$(D) = 3$, and deg$(A) = 5$.

Thus, by Theorems 7 and 8, no Eulerian circuit is possible in such a graph, nor an Eulerian trail.

The next example will offer a way in which an Eulerian circuit can be constructed in an Eulerian graph.

Example 17

The vertices in the following graph are the roads connecting several cities that you want to visit on a short holiday. You don't want to use the same road twice and you want to return home to city *a*. Find a route for your trip.

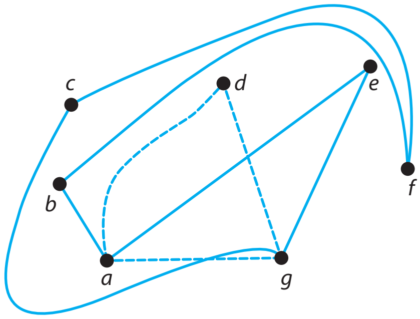


Solution

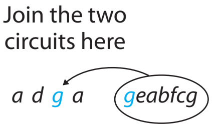
This is asking you to find an Eulerian circuit for the given graph.

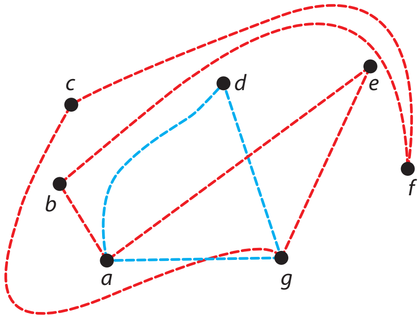
This is an Eulerian graph since all vertex degrees are even.

First construct a circuit *C* beginning with *a* (say); *adga* is such a circuit. Since it does not include all edges, it is not Eulerian. Next, look for a vertex in *C* that is adjacent to a non-used edge; *a* and *g* are such vertices. Beginning with *g*, for example, construct a circuit using unused edges; *geabfcg* is such a circuit. Use a broken line as before.



Since no more solid edges remain, the procedure stops here. To combine the two circuits, join them at vertex *g* where the second circuit started.



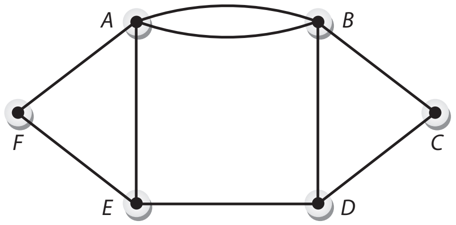


Thus, the Eulerian circuit for the graph is

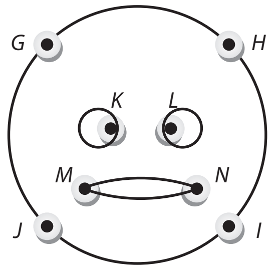
adgeabfcga.

Example 18

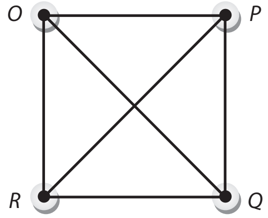
In which of the following graphs is it possible to find an Eulerian trail or an Eulerian circuit? When possible, find an example of the trail or circuit. When not possible, explain the reasons for the absence of an Eulerian trail or circuit.



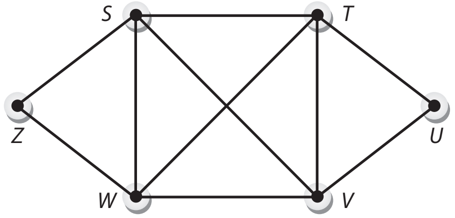
a)



b)



c)



d)

Solution

a) Looking at vertex degrees, we have:

$\deg(A) = \deg(B) = 4$, $\deg(C) = \deg(F) = 2$ and $\deg(D) = \deg(E) = 3$

Since two vertices have odd degrees, it is possible to find a trail.

We need to start from a vertex of an odd degree, so one possible Eulerian trail would be:

$D, \{D, C\}, C, \{C, B\}, B, \{B, D\}, D, \{D, E\}, E, \{E, F\}, F, \{F, A\}, A, \{A, B\}_{\text{upper}}, B, \{B, A\}_{\text{lower}}, A, \{A, E\}, E.$

b) Even though all vertices are of an even degree (2) the graph is not connected; therefore, it is not possible to find either an Eulerian trail or an Eulerian circuit.

c) All the vertices are of the same degree (3), so it is not possible to find either an Eulerian trail or an Eulerian circuit.

d) Looking at vertex degrees we have:

$\deg(S) = \deg(T) = \deg(V) = \deg(W) = 4$ and $\deg(U) = \deg(Z) = 2$

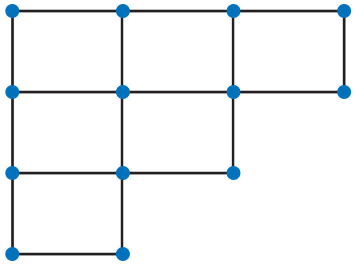
Thus, it is possible to find a circuit. We can start from any vertex, so one of the possible Eulerian circuits would be:

STVWSVUTWZS.

If we apply the algorithm presented in Example 17 above, we can start with a circuit *SZWS*, for example. Then *WTVW*, and lastly *VUTSV*. Now we join the first two at *W*, getting a new circuit *SZWTVWS*. Lastly, we join this circuit with the third one at *V*, thus getting *SZWTVUTSVWS* as our Eulerian circuit.

Hamiltonian graphs

Below is a graph where the vertices represent locations of postal boxes where mail has to be picked up every day. Postal services must find a route so that mail can be picked up from each of these boxes. Would an Eulerian circuit suffice for this job?



The answer is No! An Eulerian circuit would not provide a good solution since the primary goal is simply visiting each vertex rather than travelling each edge. In this problem, it would be very inefficient to require each edge to be travelled since this would force multiple visits to the same vertex.

In general, Eulerian circuits/paths are not the appropriate tool for analyzing problems where it is only important to visit each vertex. For problems of this type, whether an edge is travelled is not important.

Remember, Eulerian circuits/paths deal with situations where it is important to travel every edge.

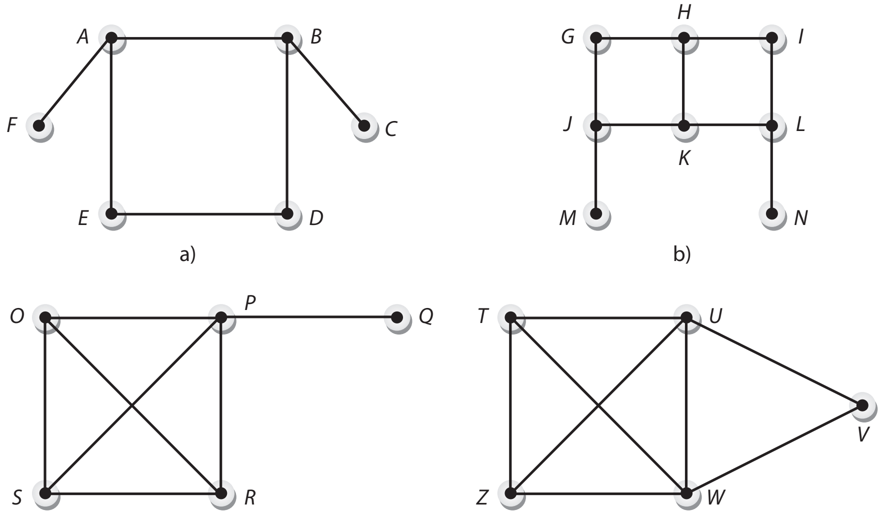
We have found some conditions for the existence of trails and circuits containing all the edges of a graph only once. Can we do a similar task with vertices? Is it possible to find a path or a cycle that contains all the vertices in a given graph?

Definition 18

Let $G = (V, E)$ be a **connected graph**. A **path** that contains all vertices of G is called a **Hamiltonian path**. A **cycle** that contains all vertices of G is called a **Hamiltonian cycle**. A connected graph that contains a Hamiltonian cycle is called a **Hamiltonian graph**.

Example 19

In which of the following graphs is it possible to find a **Hamiltonian path** or a **Hamiltonian cycle**? When possible, find an example of the path or cycle; and when not, explain the reasons for the absence of a Hamiltonian path or cycle.



a) b) c) d)

Solution

a) Two vertices have a degree of 1, so if we leave either of these two vertices we cannot come back to them; therefore, it is not possible to find a cycle. A possible Hamiltonian path would be:

$F, \{F, A\}, A, \{A, E\}, E, \{E, D\}, D, \{D, B\}, B, \{B, C\}, C.$

b) It is not possible to find a Hamiltonian cycle because there are two vertices of degree 1. Neither is it possible to find a Hamiltonian path since at the end there are two non-adjacent vertices that we need to connect.

c) There is only one vertex of a degree of 1; therefore, it is not possible to find a cycle. A possible Hamiltonian path would be *QPOSR*.

d) It is possible to find a Hamiltonian cycle. We can start from any vertex, so one such possible cycle would be *VUZTWV*.

Unlike the situation with Eulerian trails and circuits, there is no well-known test, or listing of requisites, that can be employed to establish whether a graph contains a Hamiltonian path or cycle. In its place, there are some negative tests, which can explain that a certain graph cannot contain such a cycle or path. There are some theorems that establish either necessary conditions or sufficient conditions for a graph to have a Hamiltonian path or cycle. We will examine some of these in the following pages. When faced with certain graphs, however, we will time and again resort to trial and error.

Theorem 9 (Optional but extremely helpful)

Let $G = (V, E)$ be a **simple connected graph**. If $|V| = n$, $n \geqslant 3$ and, for each vertex $A \in V$, $\deg(A) \geqslant \frac{n}{2}$, then the graph G has a Hamiltonian cycle. This fact is known as **Dirac's theorem**.

The proof is beyond the scope of this publication.

Note: We can easily see that this is not a necessary condition. The dodecahedron graph corresponding to Hamilton's original game has $n = 20$ and $\deg(v) = 3$ for every vertex v, yet the graph is Hamiltonian.

Theorem 10 (Optional)

Let $G = (V, E)$ be a **simple connected graph**. If $|V| = n$, $n \geqslant 3$ and, for each pair of **non-adjacent** vertices $A, B \in V$, $\deg(A) + \deg(B) \geqslant n$, then the graph G has a Hamiltonian cycle. This fact is known as **Ore's theorem**. This is a generalization of Dirac's theorem.

Proof

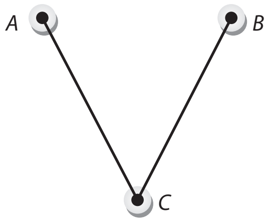
It can be proved by Dirac's theorem. Since for any two vertices A and B on graph G

$\deg(A) \geqslant \frac{n}{2}$, $\deg(B) \geqslant \frac{n}{2} \Rightarrow \deg(A) + \deg(B) \geqslant \frac{n}{2} + \frac{n}{2} = n$, so this must be true for two non-adjacent vertices too.

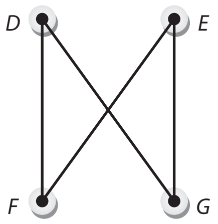
Unfortunately these two theorems give us only **sufficient** conditions, not **necessary** conditions for the statement. Also, once we know of the existence of a Hamiltonian cycle, there is no guidance for finding that cycle or how to find a Hamiltonian path.

Example 20

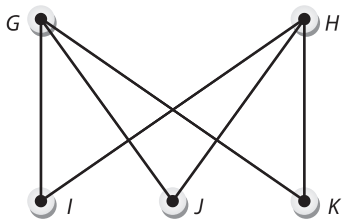
In which of the following bipartite graphs is it possible to find a Hamiltonian path or a Hamiltonian cycle? If possible, find an example of it and if not possible, give a reason why not.



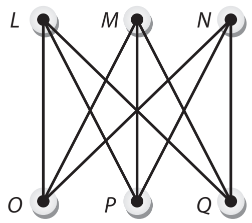
a)



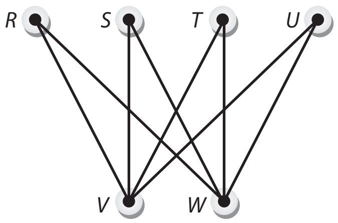
b)



c)



d)



e)

Solution

a) There is a Hamiltonian path A, $\{A, C\}$, C, $\{C, B\}$, B, but no cycle. We can see that the vertices don't satisfy the conditions of Theorems 9 or 10.

b) There is a Hamiltonian cycle. One such possible cycle would be:

D, {*D*, *F*}, *F*, {*F*, *E*}, *E*, {*E*, *G*}, *G*, {*G*, *D*}, *D*.

We can observe that all four vertices have a degree of 2 and they satisfy the conditions of Theorems 9 and 10.

c) There is a Hamiltonian path but no cycle. To find one such path we need to start from a vertex of a degree 2 and not repeat a vertex before we travel through all of them. One possible path is *J*, {*J*, *G*}, *G*, {*G*, *I*}, *I*, {*I*, *H*}, *H*, {*H*, *K*}, *K*. We notice that the vertices don't satisfy the conditions of the theorems since vertices *I*, *J*, and *K* have a degree of 2, which is less than 2.5. Also, taken two at a time, the sum of their degrees is 4, which is less than 5.

d) There is a Hamiltonian cycle. One such possible cycle would be:

L, {*L*, *O*}, *O*, {*O*, *M*}, *M*, {*M*, *P*}, *P*, {*P*, *N*}, *N*, {*N*, *Q*}, *Q*, {*Q*, *L*}, *L*.

We notice that all four vertices have a degree of 3 and they satisfy the conditions of Theorems 9 and 10.

e) There is no Hamiltonian path nor cycle. The problem is that every time we visit a 2-degree vertex, we need to leave it, revisiting a 4-degree vertex. And hence there is no Hamiltonian cycle.

The above example points to two possible negative tests.

Bipartite graphs – negative tests

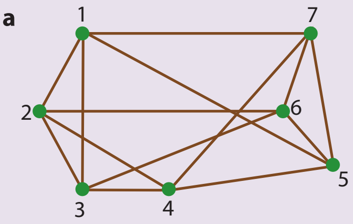
G is a bipartite graph with V_1 and V_2 subsets of vertices. Let subset 1 have *m* vertices and subset 2, *n vertices.*

- If $m \neq n$, *G* cannot have a Hamiltonian cycle. The case with Example 20 a), c), and e).
- If *m* and *n* differ by 2 or more, there is no Hamiltonian path. The case with Example 20 e).

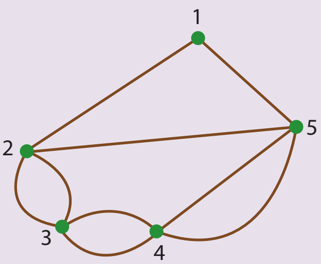
Exercise 3.4

1 Explain why each of the following graphs is Eulerian and find an Eulerian circuit for each.

a

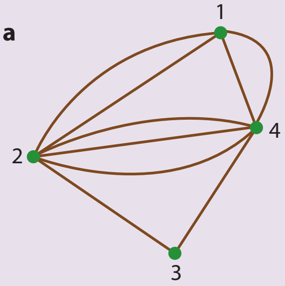


b

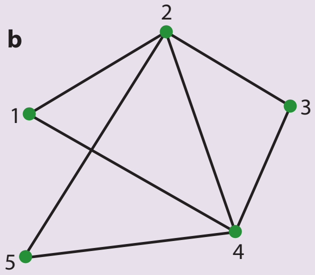


2 In each of the graphs below, find an Eulerian circuit or explain why no Eulerian circuit exists.

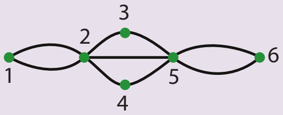
a



b



c

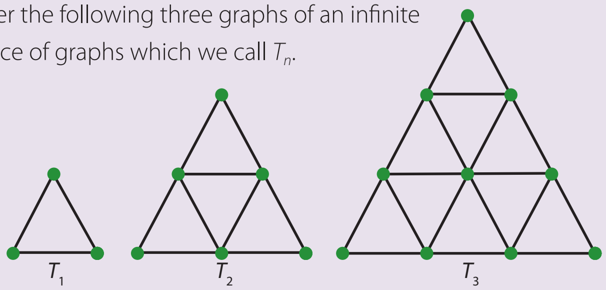


3 Under what conditions would each of the following be Eulerian? Justify your answer.

a K_n **b** $K_{m,n}$

4 Are the graphs in questions 1 and 2 Hamiltonian? If one is not Hamiltonian but has a Hamiltonian path, find it.

5 Consider the following three graphs of an infinite sequence of graphs which we call T_n.

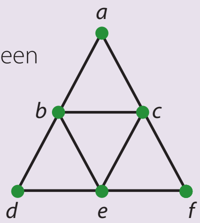


a Find an Eulerian circuit when possible, or justify why not when one does not exist.

b Find a Hamiltonian cycle when possible, or justify why not when one does not exist.

c When is T_n Eulerian? Hamiltonian?

6 How many walks of length 1, 2, 3, or 4 are there between *a* and *e* in the simple graph right?



7 Find the number of walks of length x between the vertices in K_5 when x is

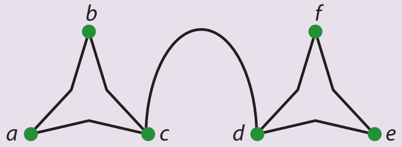
a 4   **b** 5   **c** 6

8 Consider the graph $K_{3,4}$. Let *a* and *b* be two vertices in the subset of three non-adjacent vertices. Find the number of walks of length x between these vertices when x is

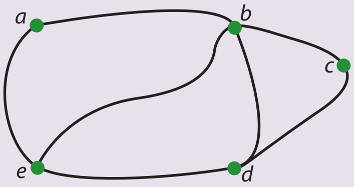
a 4   **b** 5   **c** 6   **d** 7

9 In each of the following, determine whether the given graph has a Hamiltonian cycle. If it does, find one such cycle. If it does not, justify why not. For those graphs that do not have a cycle, do any of them have a Hamiltonian path? If yes, find it and if not, justify why not.

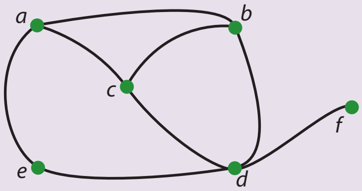
a



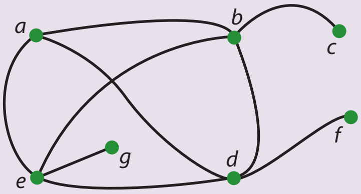
b



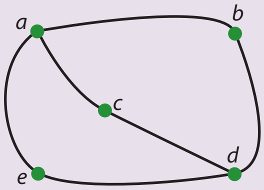
c



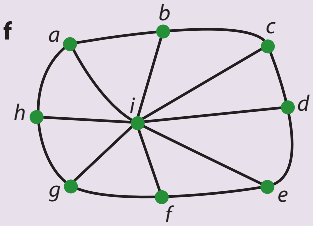
d



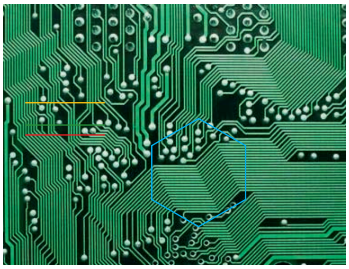
e



f



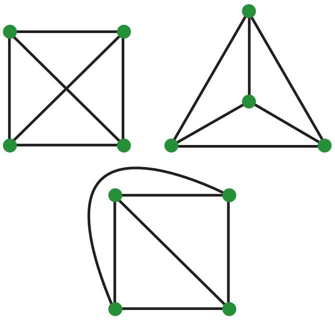
3.5 Planar graphs



One of the applications of graph theory is in the design of electronic components. In cases of computer chips, electronic components are assembled using printed circuits, where the conducting strips are printed onto boards of insulating material. The conducting strips may not cross, since that would lead to a malfunction of the component because of short circuits. Complex circuits where crossing strips are unavoidable have to be printed on several boards which are then packed together. Naturally, manufacturers want to print circuits onto the minimum number of boards, for obvious reasons. This is an application where graphs that represent components of circuits have to be **planar**.

Definition 19

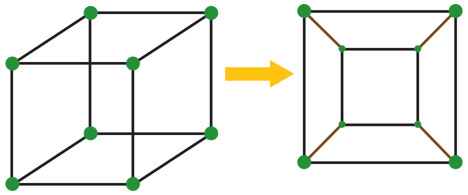
A **planar graph** is a graph that can be represented by a diagram in which no edges cross. Such a diagram is called a **plane diagram** (also known as **planar representation** or **embedding**). For example, K_4 is a planar graph.



For instance, two diagrams of K_4 are shown left. The first is not a plane diagram, while the second and third are.

Example 21

Is the graph known as the 3-cube, Q_3 shown below, planar?

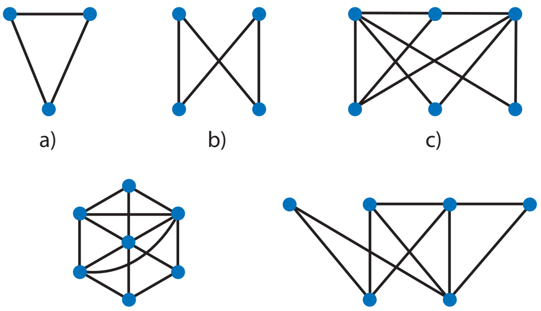


Solution

Q_3 is planar because it can be drawn without any edges crossing, as you can see in the accompanying plane diagram.

Example 22

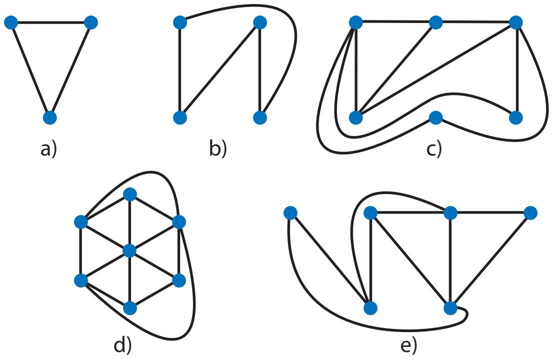
Below are the plane graphs of a few graphs. Show that they are planar.



a) b) c) d) e)

Solution

Here are the plane graphs redrawn to show that no two edges in any of the graphs cross. Hence, they are planar.



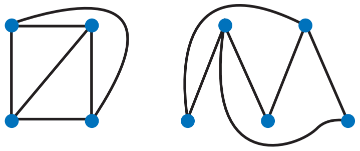
a) b) c) d) e)

Example 23 (Important)

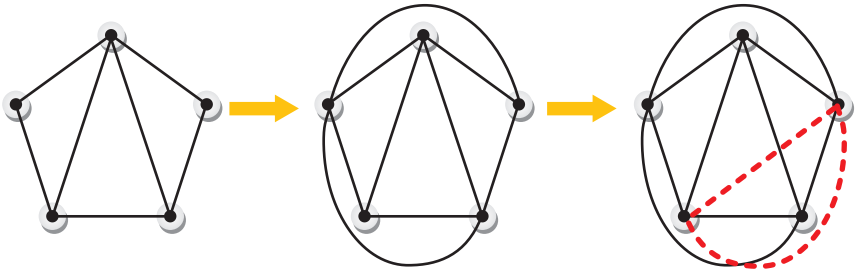
Investigate which of the complete graphs K_n and complete bipartite graphs $K_{m,n}$ are planar.

Solution

It is obvious that the following complete graphs are planar: K_1, K_2, K_3, $K_{2,1}$, and $K_{2,2}$ (as shown in Example 22). It is not very difficult to find the planar embedding for K_4 and $K_{3,2}$, as shown in the following figure.

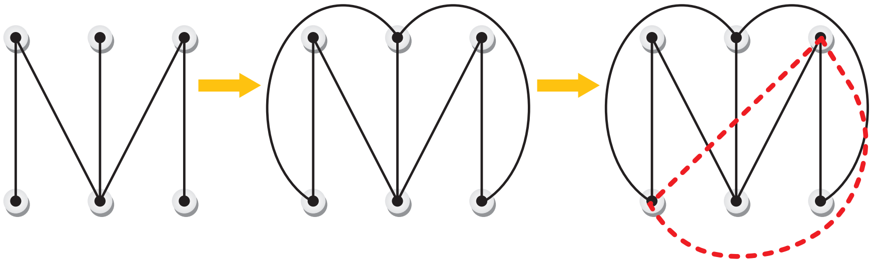


Whether K_5 and $K_{3,3}$ are planar needs to be further investigated. Start with K_5. After drawing the pentagon and all the diagonals from one vertex, proceed with drawing one edge at a time.



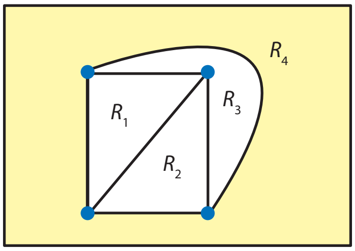
It becomes clear that in order to draw the last edge we must cross one of the previously drawn edges; therefore, it is not possible to find a planar representation of K_5.

Apply a similar approach to find a plane diagram of $K_{3,3}$.



You can see that before reaching the last edge, there is no way to draw any edge left without crossing some other edge. Thus, $K_{3,3}$ is not planar.

Euler's formula



A planar representation of a graph partitions the plane into separate regions. For example, the graph diagram K_4 is given left, and, as you notice, it splits the plane into four **regions** (known as **faces** in IB documents). Euler showed that all graph diagrams of the same graph partition the plane into the same number of regions. He accomplished this by finding a relationship between the number of regions, the number of edges and the number of vertices of a planar graph.

Theorem 11 (Euler's formula)

Let $G = (V, E)$ be a **connected planar simple graph** (**multigraph**) where $|V| = v, |E| = e$, and f is the number of faces or regions this graph's **planar embedding** establishes in the plane, then

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

Proof (By induction)

$P(e)$: For every embedding of a connected planar graph with e edges, v vertices, and f faces, $v - e + f = 2$.

Basis step: $P(0)$: The formula is true for a graph with zero edges. This means the graph is made of one vertex only. $v = 1, f = 1$ (since the vertex does not partition the plane!) and $e = 0$. Since $1 - 0 + 1 = 2$, so $P(0)$ is true. We can also consider $P(1)$. That means one edge, thus $v = 2$ and $f = 1$. Thus $2 - 1 + 1 = 2$, which indicates that $P(1)$ is true. (If the edge is a loop, it is a similar argument with $f = 2$, $v = 1$, and $e = 1$.)

Inductive step: Let $k > 1$ be given such that $P(k)$ is true. That is, we have a connected planar graph with k edges, v vertices, and f faces where the formula is true, $v - k + f = 2$. Now, consider a graph G with $k + 1$ edges, v vertices, and f faces. G either has a cycle or does not have one.

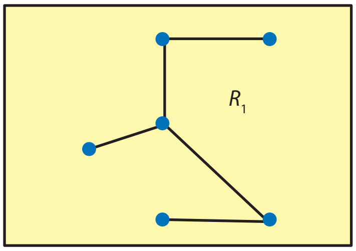


Figure 3.1

There will be more about this in the next chapter.

Case 1: G has no cycle. Since there are no cycles, the graph is not closed and there is only one unbounded face. (See Figure 3.1.) $v = k + 2$. In an open graph, every edge has two vertices, but since it is connected, every two edges share one vertex, and hence each edge contributes one to the number of vertices available, except either the first or last edge, and hence

$$v - e + f = k + 2 - (k + 1) + 1 = 2.$$

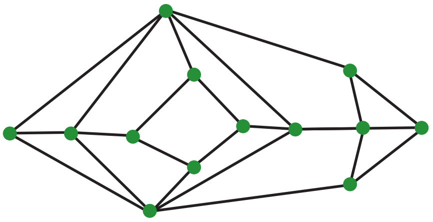
Case 2: G has a cycle. Let a be an edge in this cycle. Now, create a graph G_1 by deleting the edge a from G. (Deleting an edge merges two regions R_1 and R_2, for example, together.) This subgraph contains k edges and $f - 1$ faces. Using the fact that $P(k)$ is true and can be applied to G_1, then

$$v - k + f - 1 = 2 \Rightarrow v - (k + 1) + f = v - e + f = 2.$$

Thus, by the principle of mathematical induction, $P(0)$ is true, and assuming $P(k)$ to be true, we showed that $P(k + 1)$ is true, and thus the relation is true for all $e \in \mathbb{N}$.

Example 24

Verify Euler's formula for the connected planar graph given right.



Solution

The graph has 13 vertices, 23 edges, and 12 regions.

So, $13 - 23 + 12 = 25 - 23 = 2$.

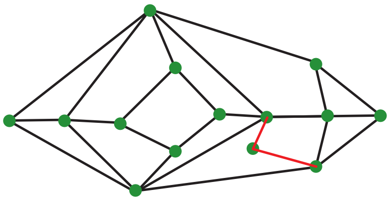
Example 25

A connected planar graph has 24 edges, dividing the plane into 12 regions. How many vertices does this graph have? Create such a graph.

Solution

$$v - 25 + 13 = 2 \Rightarrow v = 14.$$

We took the liberty of using the previous graph and added one vertex!



Theorem 12

If G is a connected simple planar graph with e edges and $v > 2$ vertices, then $e \leqslant 3v - 6$.

Proof

Given that we need at least three edges to form two regions or faces[1] in a simple graph then $2e \geqslant 3f$. Then, by using Euler's formula, we obtain the following:

$$\left.\begin{array}{l} 2 + e - v = f \\ 2e \geqslant 3f \end{array}\right\} \Rightarrow 2e \geqslant 3(2 + e - v) \Rightarrow 2e \geqslant 6 + 3e - 3v \Rightarrow e \leqslant 3v - 6$$

Example 26

Show that K_5 is not planar.

Solution

K_5 is a simple connected graph with $e = 10$ and $v = 5$. If it were planar, then

$$e = 10 \leqslant 3v - 6 = 15 - 6 = 9,$$

which is not true. Thus, K_5 is not planar.

[1] There are some other considerations we chose not to include here. For more information, see Ralph Grimaldi, *Discrete and Combinatorial Mathematics*, 5th edition (Addison-Wesley, 2003).

Theorem 13

If G is a connected simple planar graph with e edges and $v > 2$ vertices, and no circuits of length 3, then $e \leqslant 2v - 4$.

Proof

The proof is similar to that of Theorem 12. Since there are no circuits of degree 3, then we need at least four edges to form two regions. Hence, $2e \geqslant 4f$. Thus,

$$\left.\begin{array}{l}2+e-v=f\\ 2e\geqslant 4f\end{array}\right\}\Rightarrow 2e\geqslant 4(2+e-v)\Rightarrow 2e\geqslant 8+4e-4v\Rightarrow 2e\leqslant 4v-8\Rightarrow e\leqslant 2v-4$$

Example 27

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

Solution

$K_{3,3}$ is a simple connected graph with no circuit of length 3. $v = 6$ and $e = 9$. If it were planar, then

$$e = 9 \leqslant 2v - 4 = 12 - 4 = 8,$$

which is not true. Thus, $K_{3,3}$ is not planar.

Important

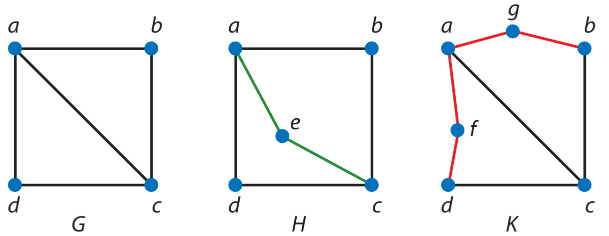
Since $K_{3,3}$ is a simple connected graph, if we were to apply Theorem 12, then we have $e = 9 \leqslant 3v - 6 = 18 - 6 = 12$, which is true! It would be **an error to conclude that $K_{3,3}$ is planar**. This is using the converse of the theorem without proving it. Unfortunately, the theorem we proved is necessary but not sufficient. That is, if the graph is planar, then the relation is true.

Note: Since K_5 and $K_{3,3}$ are not planar, it is obvious that all the graphs containing K_5 or $K_{3,3}$ as subgraphs are also not planar. Moreover, all the graphs that contain a subgraph that can be obtained from K_5 or $K_{3,3}$ using certain permitted operations are not planar.

Homeomorphic graphs

If we remove an edge, let's call it $\{A, B\}$, from a graph and we add another vertex C together with the edges $\{A, C\}$ and $\{B, C\}$, such an operation is called an **elementary subdivision**. Graphs are called **homeomorphic** if they can be obtained from the same graph by a sequence of elementary subdivisions.

To understand the idea consider the graphs in the following figure.



Graph H is obtained from G by one elementary subdivision: remove edge ac from G, then add the edges ae and ec to the graph. Graph K is obtained from G by two elementary subdivisions: remove ab and add ag and gb, and remove ad and add af and fd. Thus, H and K are homeomorphic.

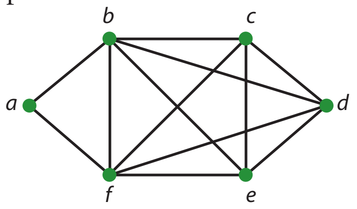
The following theorem is a useful result of the previous discussion.

Theorem 14 (Kuratowski's theorem)

A graph $G = (V, E)$ is not a planar graph if and only if it contains a subgraph homeomorphic to K_5 or $K_{3,3}$.

Example 28

Is the following graph planar?



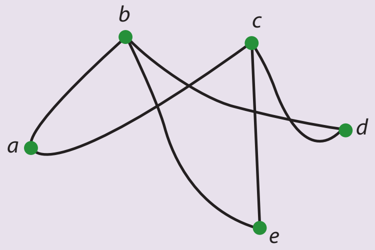
Solution

The graph is not planar since K_5 is a subgraph. $(bcdef)$ is K_5.

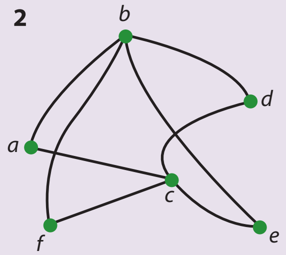
Exercise 3.5

For each graph in questions 1–4, decide whether the graph is planar. If it is, give a reason for your decision and draw a planar representation. If it is not, justify why not.

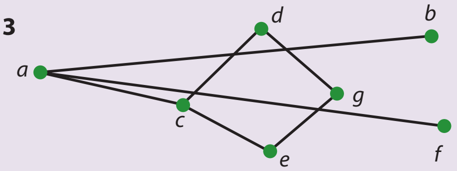
1



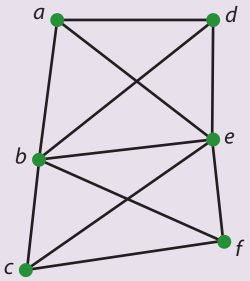
2



3



4



5 A connected planar graph contains 10 vertices and partitions the plane into seven regions. What is the number of edges in the graph?

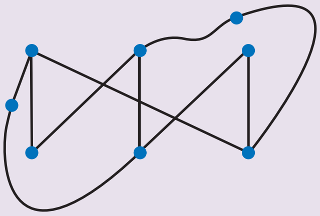
6 What is the maximum number of edges in a simple connected planar graph with 7 vertices? 8 vertices?

7 Find the minimum number of vertices in a simple connected planar graph with 14 edges? 21 edges?

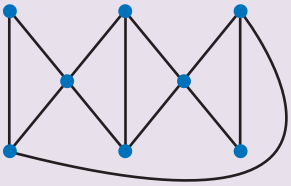
8 A connected planar graph has 8 vertices with 3 degrees each. How many regions are created by a planar embedding of this graph?

In questions 9–10, determine whether the graphs are planar.

9



10



Practice questions 3

1 Explain whether or not it is possible to have a cycle of odd length in a bipartite graph.

2 a A complete graph K_n contains subgraphs isomorphic to K_m, where $m < n$. How many isomorphic subgraphs does K_n contain if:

i $m = 2$ ii $m = 3$ iii $m, m = 1, \dots, n$

b For what value(s) of m would the number of isomorphic graphs be the largest?

3 Given a complete graph K_5, find the number of trails no longer than 3 between two vertices.

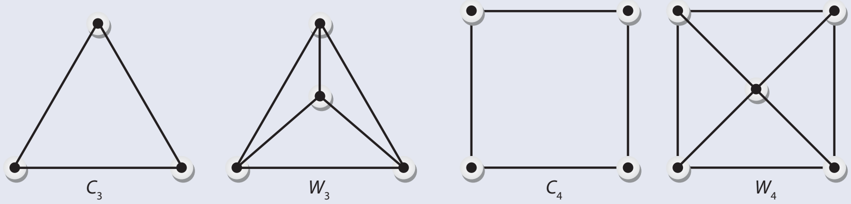
4 Given the complete graph K_4, and a walk of length l between any two vertices in the graph, find the number of different walks when

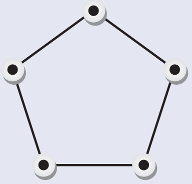
a $l = 2$ b $l = 3$.

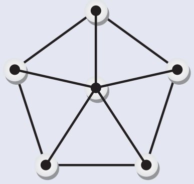
5 Given the complete bipartite graph $K_{3,3}$ and a walk of length l between any two non-adjacent vertices in the graph, find the number of different walks when

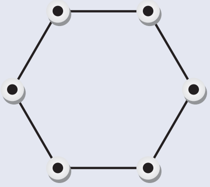
a $l = 3$ b $l = 4$.

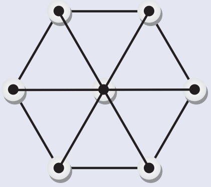
6 **Cycle** C_n, $n \geqslant 3$, is a graph in which every vertex has an order of 2.
Wheel W_n, $n \geqslant 3$, is a graph that consists of a cycle C_n and an additional point that is connected to all the vertices in the cycle. Below are some examples of cycles and wheels:




C_5


W_5


C_6


W_6

a Show that the number of edges in a wheel W_n is twice the number of edges in a cycle C_n .

b Are any of these graphs, C_n or W_n, isomorphic to a complete graph K_n?

c Show that in C_4 there are 2^{n-1} paths of length n between

i adjacent vertices when n is odd

ii non-adjacent vertices when n is even.

7 Show that a cycle graph C_n, $n \geqslant 3$, is bipartite if and only if n is even.

8 Explain why no wheel graph W_n, $n \geqslant 3$, can be bipartite.

9 Draw the complementary graph of C_5. Is the complementary graph isomorphic to the original graph? If yes, construct an isomorphism between those two graphs.

10 A graph is called **self-complementary** if it is isomorphic to its complementary graph. Is it possible to find a self-complementary graph with

a 4 vertices **b** 6 vertices?

If possible, draw the graph and its complementary graph.

11 A parent-teacher organization (PTO) at an international school has six people working for it. They are Adam, Bernard, Cecile, Donatella, Eva, and Flor. They can communicate in at least one language according to the following table.

Name	English	Spanish	French	German
Adam	✓	✓	✓	
Bernard	✓		✓	✓
Cecile			✓	
Donatella		✓	✓	
Eva				✓
Flor		✓		

a Draw a graph indicating which people can communicate with each other.

b Cecile ordinarily communicates with Flor with the help of Donatella. Unfortunately, Donatella has gone to visit her mother. Can Cecile still communicate with Flor? Write down how it can be done.

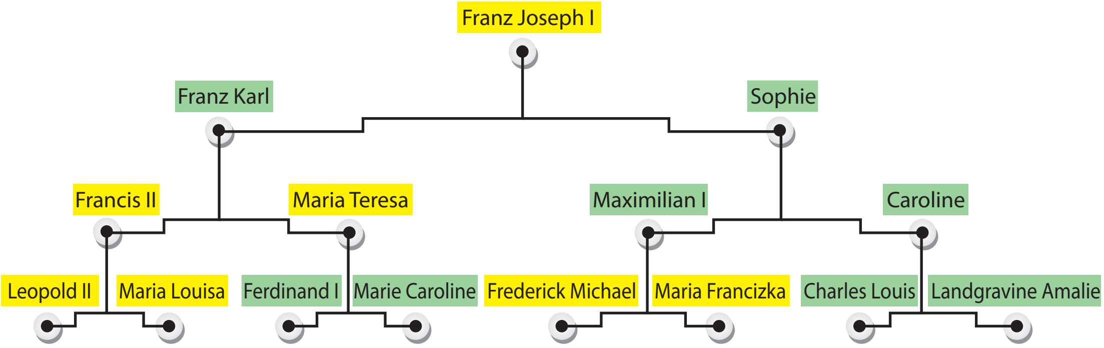
c Who is the most important person without whom it is not possible to communicate with all the members of the PTO? Give your reasons.

4 Trees and Algorithms

4.1 Introduction

Trees are among the most, if not the most, important class of graphs and they make fine modelling tools. In 1847, Gustav Kirchhof, a German scientist, used them to solve systems of equations for electrical networks. In 1857, the English mathematician Arthur Cayley used them to count the different isomers of the saturated hydrocarbons. Today, trees are widely used in mathematics, computer science, and many other fields including social sciences.

For example, a common representation of the genealogical charts of a family is called a family tree. In the form of a graph, vertices represent the family members, whilst edges represent the parent-child relationship. Here is a tree that represents the ancestors of the Austrian Emperor Franz Joseph I.



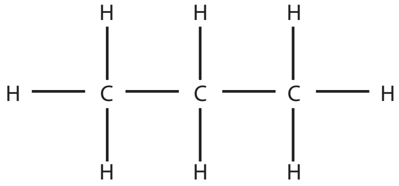
4.2 Trees

You are familiar with trees in graph theory. In Chapter 3, we discussed several instances of connected graphs that do not contain cycles. These are trees. As in graph theory, tree terminology is unfortunately not standard. We will use the IBO terminology in this publication.

Definition 1

Let $T = (V, E)$ be a **connected simple graph**. If T contains no **cycles**, it is a **tree**. A **subtree** is a **subgraph** of a **tree** that is a **tree** itself.

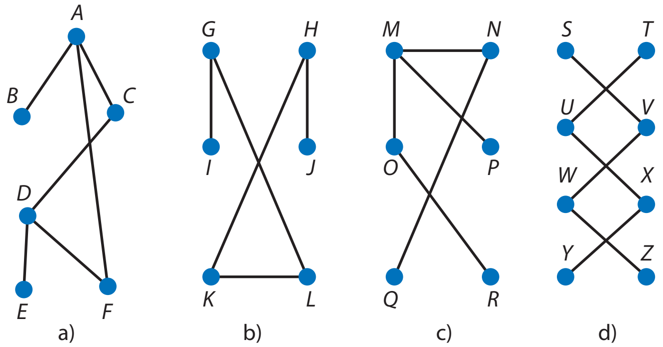
For example, the compound propane (C_3H_8) has this structure:



The structure has no cycle, so it is a tree.

Example 1

Which of the graphs are trees? Give your reasons.



Solution

Graphs b) and c) are trees. Graph a) contains a cycle, *ACDFA*, while graph d) is not connected.

Theorem 1

A graph $T = (V, E)$ is a **tree** if and only if there is **a unique simple path** between any pair of vertices.

Proof

(⇒) If graph T is a tree, then it is connected with no cycles; thus, for any two vertices, there is a simple path between those two vertices. The uniqueness of the path can be proven by contradiction. Assume that there are two different paths between two vertices, but then those two paths together would form a cycle which is a contradiction, since T is a tree.

(⇐) Now, assume that there is a unique simple path between any two vertices of the graph T. Given that there is a path then graph T is connected. Now, if graph T contains a cycle, then between two vertices in that cycle we can find two different paths, which contradicts the uniqueness of the path.

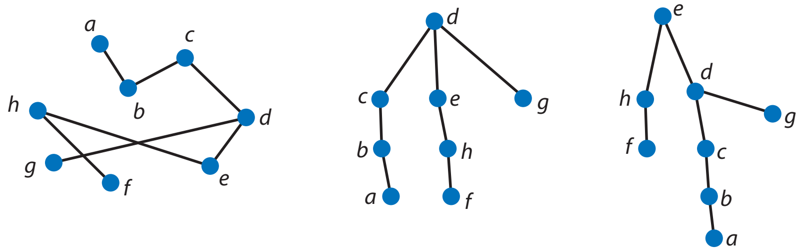
In many applications of trees, such as the family tree we discussed earlier, organizational trees, computer file systems, networks, etc., a vertex is designated as the **root.** Since there is a unique path from 'the root' to each vertex of the tree by Theorem 1, we direct each edge away from the root in a manner described by Figure 4.1. A tree with its root produces a graph called a **rooted tree**.

Definition 2

Let $T = (V, E)$ be a **tree**. Let v_i be a vertex such that every edge is directed away from it. T is called a **rooted tree**.

As you notice from the definition above, we can change any tree into a rooted tree by the choice of the **root**.

Figure 4.1



'Unrooted' *T*

With root *d*

With root *e*

In a rooted tree, the starting vertex is the **root** while the other vertices are called **parent**, **child**, **siblings**, **ancestors**, and **descendants**. A vertex of a tree with no children is called a **leaf**. Vertices that have children are called **internal vertices**.

In Figure 4.1 above, for the tree with root *d*, *b* is a parent of *a* and *a* is a child of *b*. Vertices *c*, *e*, and *g* are siblings, since they have the same parent *d*. Ancestors of *f* are *d*, *e*, and *h*, whereas *a* and *f* have no descendants – therefore each of them is a leaf. We can say that all the vertices in the tree are descendants of the root. An internal vertex in a rooted tree is said to be at a level *i* when the path connecting it to the root is *i*. For example, in the tree with root *d*, *c*, *e*, and *g* are at level 1, while *a* and *f* are at level 3. In the tree with root *e*, *h* and *d* are at level 1, while *a* is at level 4.

Note: All vertices in a rooted tree have each a degree at least 2, except for the leaves. Each leaf has a degree of 1.

Theorem 2

A **tree** $T = (V, E)$ with n vertices has $n - 1$ edges.

Proof

We will conduct the proof by mathematical induction.

Statement: $S(n)$: a tree with n vertices has $n - 1$ edges.

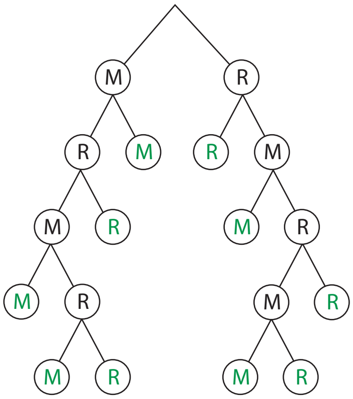
Basis step: When a tree has only one vertex, it has no edges. The statement is thus true for $n = 1$.

Inductive step: Assume that every tree with k vertices has $k - 1$ edges.

Now, consider a tree that has $k + 1$ vertices. Let vertex a be a leaf of T and let vertex b be the parent of a. Removing vertex a from the tree removes the edge $\{a, b\}$ too and leaves us with a **subtree** that has k vertices. By assumption, this subtree has $k - 1$ edges. However, T has one more edge than its subtree, and therefore has k edges. Thus, tree T that has $k + 1$ vertices has $(k + 1) - 1$ edges.

Conclusion: Since the statement is true for $n = 1$ and $S(k) \Rightarrow S(k + 1)$, by the principle of mathematical induction, the statement is true for all $n \in \mathbb{Z}^+$.

Example 2



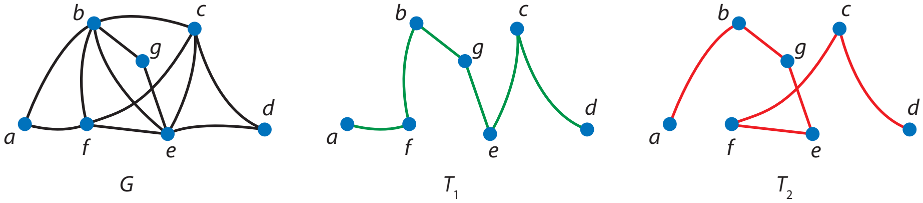
Marco and Roberto play a tennis game. They agree that whoever wins a total of three games first or two games in a row will be declared the winner. How many outcomes are possible, and what is the maximum number of games they will play?

Solution

The situation can be represented by a tree. There could be 10 possible outcomes corresponding to the vertices with degree 1 in the tree. The number of possible games corresponds to the layers of the tree we have, that is, five games.

4.3 Spanning trees

All connected graphs have trees that span them. Consider the following situation: In a small mountainous area, winter is harsh and snow sometimes makes it difficult to keep all the towns connected to the rest of the world. Because of the cost involved and the amount of equipment needed for the task, the authorities try to make sure that a minimum number of roads between the towns are accessible by ploughing as few roads as possible. Graph G below shows the road network on the left and two possible networks of ploughed roads to the right (T_1 and T_2). These subgraphs of G are called **spanning trees** of G.



Definition 3

Let $G = (V, E)$ be a **connected graph**. A subgraph H of G is a **spanning tree** of G if H is a tree which contains every vertex of G.

Theorem 3

Every connected graph has a spanning tree.

Proof

Let G be a connected graph. If G has no cycles, then it is a tree and we are done.

If G is not a tree, it must contain at least one cycle. Remove an edge from the cycle. The graph is still connected. If the new graph is acyclic (with no cycles), then it is a tree, and hence a spanning tree since it visits all vertices. Otherwise, it must have another cycle. Repeat the process with another edge from a cycle, until a subgraph T is acyclic. Since T is acyclic, connected, and contains every vertex, then it is a spanning tree.

How to find a spanning tree

Spanning trees can be constructed in two ways, either by removing edges (vertices are not removed) which form cycles or by building a tree one edge at a time. The two methods are described below.

Method 1: Edge removal

Assume that $G = (V, E)$ is a **connected graph**. Edges are removed one at a time in such a way that the resulting graph always remains connected. If this is done until no further edges can be removed, then the resulting graph is a **spanning tree**.

Method 2: Edge addition

Assume that $G = (V, E)$ is a **connected graph**. Start with the subgraph containing all the vertices from the set V. Adjoin the edges, one edge at a time, in such a way that the resulting graph has no cycle. If this is done until no further edge can be added, then the resulting graph is a **spanning tree**.

We will present here three algorithms for constructing spanning trees. They all proceed by successively adding edges that have not already been used. We will consider non-programming sets of instructions for these algorithms. One of these is Kruskal's algorithm which makes use of Theorem 2 of Section 4.2.

Kruskal's algorithm

Given that a graph $G = (V, E)$ is a simple connected graph, and $|V| = n$, find a spanning tree T for G.

Algorithm

Set the counter $i = 0$. (i is the number of edges of the sought tree. Every time we add an edge, we increase this number by 1.)

Step 1: Select an edge, e_1. If e_1 does not create a cycle, add it to the tree, set $i = 1$, and add e_1 to the tree T.

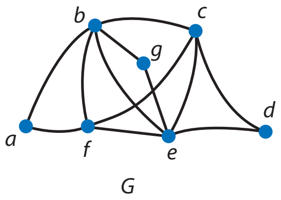
Step 2: For $1 \leqslant i \leqslant n - 2$, if edges $e_1, e_2, \ldots, e_i$ have been selected, then select edge e_{i+1} from the remaining edges so that the subgraph determined by $e_1, e_2, \ldots, e_{i+1}$ contains no cycles.

Step 3: Replace i by $i + 1$.

If $i = n - 1$, the subgraph T determined by $e_1, e_2, \ldots, e_{i+1}$ is connected with $n - 1$ edges and n vertices, and hence is a spanning tree.

If $i < n - 1$, return to step 2.

Example 3

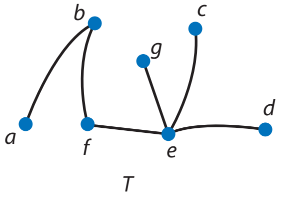


Apply Kruskal's algorithm to find a spanning tree for graph G given left.

Solution

We will construct a spanning tree using the steps in Kruskal's algorithm and summarize the steps in the table below. Observe that the number of vertices is seven.

Edge in G	Cycle formed?	Edges in tree	Number of edges in tree	Notes
ab	no	*ab*	1	
bf	no	*ab, bf*	2	
fa	yes	*ab, bf*	2	no edges added
fe	no	*ab, bf, fe*	3	
eg	no	*ab, bf, fe, eg*	4	
gb	yes	*ab, bf, fe, eg*	4	no edges added
ec	no	*ab, bf, fe, eg, ec*	5	
ed	no	*ab, bf, fe, eg, ec, ed*	6	stop, $i = 7 - 1$



The figure left gives the spanning tree so constructed. Notice though that this is not a unique tree and we could have created a different one if we made different choices at f, for example.

The depth-first search algorithm (DFS)

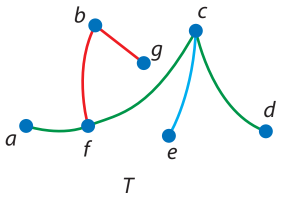
Here is an outline of the steps in this algorithm.

1. Start at a vertex v_i, and mark it as visited.
2. Pick a vertex v_{i+1}, adjacent to v_i and not yet visited.
3. Add edge v_iv_{i+1} to the tree, and replace i by $i + 1$.
4. Repeat steps 2 and 3, until you reach a vertex that has no adjacent vertices.
5. Backtrack to a vertex that has adjacent vertices that have not been visited, and repeat step 4.
6. Stop when all vertices have been visited.

Example 4

Refer to the same graph G given in Example 3. Find a spanning tree using DFS.

Solution



1. Start at a, $i = 1$.
2. Go to f.
3. $T = \{af\}$, $i = 2$.
4. Go to c: $T = \{af, fc\}$, $i = 3$.

   Go to d: $T = \{af, fc, cd\}$, $i = 4$ (path marked in green).
5. Backtrack to c and

   go to e: $T = \{af, fc, cd, ce\}$, $i = 5$ (new edge in blue).

   Backtrack to f and

   go to b: $T = \{af, fc, cd, ce, fb\}$, $i = 6$.

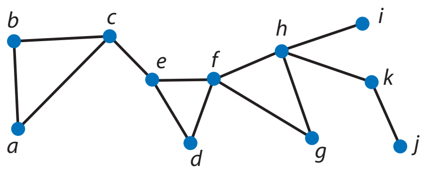
   go to g: $T = \{af, fc, cd, ce, fb, bg\}$, $i = 7$ (in red).

   Stop, all vertices added.

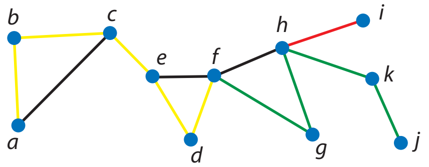
The figure is shown right.

Example 5

Find a spanning tree using DFS for the graph below.



Solution



1. Start at f, $i = 1$.
2. Go to g.
3. $T = \{fg\}$, $i = 2$.
4. Continue to h, k, and j; now $i = 5$ (in green).
5. Backtrack to h, then go to i (in red).

   Now, backtrack to f, then go to d, e, c, b, and a (in yellow).

On the left is the resulting spanning tree.

BFS as given here is in outline only. If you are interested in a detailed algorithm, check the algorithm given at the end of this section.

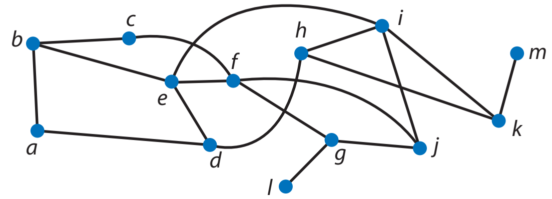
The breadth-first search algorithm (BFS)

Here is an outline of the steps in this algorithm. In this algorithm, we visit the vertices, level by level, until all vertices are visited.

1. Start at a vertex v_i, and mark it as visited.
2. Pick a vertex v_{i+1}, adjacent to v_i and not yet visited.
3. Add edge v_iv_{i+1} to the tree, and replace i by $i + 1$.
4. Visit all unvisited vertices adjacent to v_i.
5. Repeat step 4 until all vertices are visited.
6. When $i = n$, stop. All vertices are added.

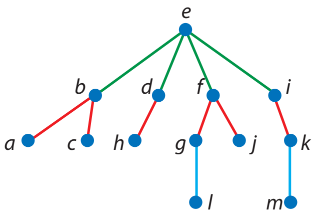
Example 6

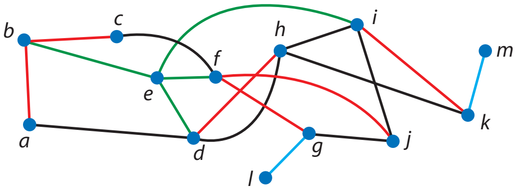
Find a spanning tree using BFS for the graph below.



Solution

1. Start at *e*.
2. Add *b*, *d*, *f*, and *i*. There are no more vertices adjacent to *e*. These are at level 1.
3. Go to *b*, add *a* and *c*. There are no more vertices adjacent to *b*.
4. Go to *d*, add *h*. No more vertices adjacent to *d*.
5. Go to *f*, add *g*, and *j*.
6. Go to *i*, add *k*. Now level 1 vertices are exhausted. Go to level 2 vertices.
7. At *a*, *c*, *h*, and *j* we cannot add any new vertices. At *g* add *l* and at *k* add *l*, and stop.





On the left is a plan of the algorithm, with the corresponding spanning tree.

BFS algorithm

```
procedure BFS(G: Connected graph with vertices $v_1, v_2, \ldots, v_n$)[1]
T := tree consisting only of vertex $v_1$
L := empty list
Put $v_1$ in L (list of unprocessed vertices)
while L is not empty
    remove the first vertex, v, from L
    for each neighbour w of v
        if w is not in L and not in T then
            add w to the end of list L
            add w and edge (v, w) to T
```

[1] Kenneth Rosen, *Discrete Mathematics and its Applications*, 7th edition (McGraw-Hill Higher Education, 2012) p. 759

Exercise 4.1–4.3

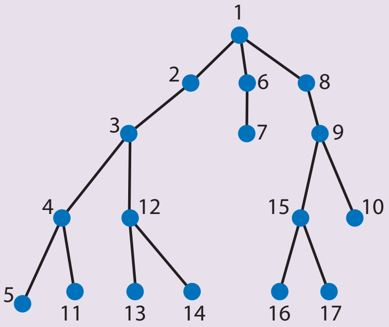
1 Consider the tree on the right.

a List the leaves of this tree.

b List the parents of 4, 8, and 15.

c List the descendants of 3, 7, and 15.

d List the siblings of 4, 7, and 9.



2 Let $T(u, e)$ and $S(v, f)$ be two trees, where u and v are the set of vertices and e and f are the sets of edges for the two trees. If $|e| = 17$ and $|v| = 2|u|$, find $|u|$, $|v|$, and $|f|$.

3 $G = (V, E)$ is a connected undirected graph with $|E| = 30$. What is the maximum number of vertices?

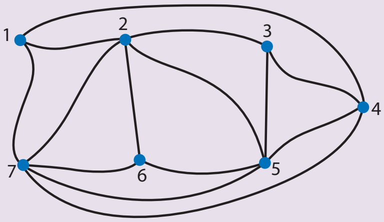
4 $T = (V, E)$ is a tree with n vertices, where $n \geqslant 2$. How many different paths are there in T?

5 a Find two non-isomorphic spanning trees for $K_{2,3}$. How many such trees are there?

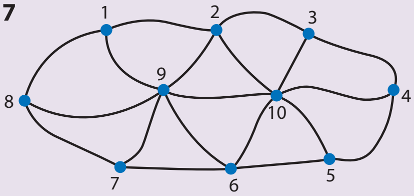
b How many non-isomorphic spanning trees are there for $K_{2,n}$, $n \in \mathbb{Z}^+$?

In questions 6–8, find a spanning tree for the graph shown. In each question use an edge removal process.

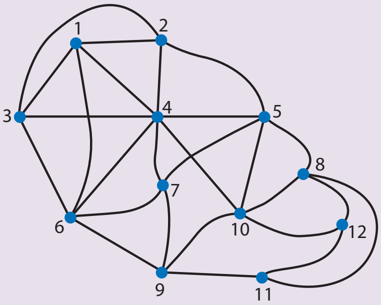
6



7



8



In questions 9–11, use Kruskal's algorithm to produce a spanning tree for each graph.

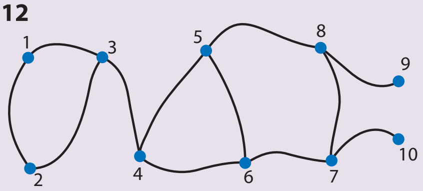
9 Find a spanning tree for the graph in questions 6.

10 Find a spanning tree for the graph in question 7.

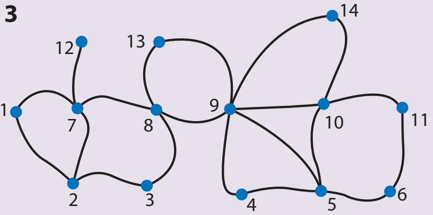
11 Find a spanning tree for the graph in question 8.

In questions 12–14, use DFS to produce a spanning tree for each graph. Consider 1 to be the root.

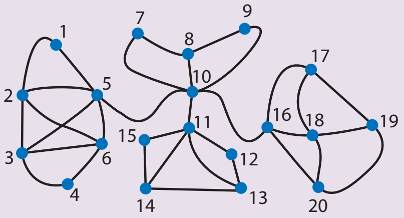
12



13



14



In questions 15–17, use **a** BFS and **b** Kruskal's algorithm to produce a spanning tree for each graph. Consider 1 to be the root.

15 Find a spanning tree for the graph in question 12.

16 Find a spanning tree for the graph in question 13.

17 Find a spanning tree for the graph in question 14.

18 **Cycle** C_n, $n \geqslant 3$, is a graph in which every vertex has an order of 2.
Wheel W_n, $n \geqslant 3$, is a graph that consists of a cycle C_n and an additional point that is connected to all the vertices in the cycle.

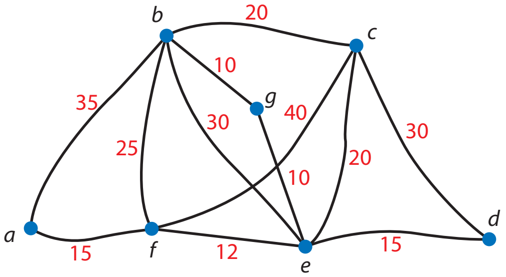
Use **a** DFS and **b** BFS to find a spanning tree for each of the following:

i W_6 starting at the centre vertex

ii K_5

iii $K_{3,4}$ starting at a vertex with degree 3

4.4 Weighted graphs and greedy algorithms



Several real situations can be modelled using graphs with weights assigned to their edges.

Consider the roads in the mountainous area discussed in Example 3. However, now we have the distances between the towns (see left). To minimize cost, we will have to minimize the total distance travelled. Airlines use such graphs to represent distances and times between different airports; networks utilize such graphs to represent the response time between different nodes; and there are many other applications. These graphs are called **weighted graphs**.

Definition 4

Let $G = (V, E)$ be a **graph**. If a numerical value or a weight is assigned to every edge of G, then we say that G is a **weighted graph**.
The **weight of a path** would be the sum of all the weights of all the edges in that path.

Representation

A convenient way of representing the weights that are assigned to the different edges is to use a special type of adjacency matrix called the **cost adjacency matrix** C_G. The entry (i, j) corresponds to the weight of the path from vertex i to vertex j. So, for example, the entry corresponding to (a, b) in the cost adjacency matrix for the graph above is 35. Below is the cost adjacency matrix for that graph.

$$C_G = \begin{array}{c|ccccccc} & a & b & c & d & e & f & g \\ \hline a & - & 35 & - & - & - & 15 & - \\ b & 35 & - & 20 & - & 30 & 25 & 10 \\ c & - & 20 & - & 30 & 20 & 40 & - \\ d & - & - & 30 & - & 15 & - & - \\ e & - & 30 & 20 & 15 & - & 12 & 10 \\ f & 15 & 25 & 40 & - & 12 & - & - \\ g & - & 10 & - & - & 10 & - & - \end{array}$$

We use the convention that where there is no connection, we put a dash (–). (In some books, 0 is used instead.)

The cost adjacency matrix is a good tool for storing data and retrieving weights of edges when needed, without getting lost in looking at the numbers next to each edge.

Weighted graphs are associated with spanning trees that have a minimum weight. In the examples in this section, we are interested in finding a spanning tree with minimum weight. Such trees are called **minimal** (or **minimum**) **spanning trees**. There are a few algorithms that help us find such trees. These are called **greedy algorithms**. Two of these will be discussed in this section: **Kruskal's algorithm** and **Prim's algorithm**.

Kruskal's algorithm

Kruskal's algorithm for minimal spanning trees is an extension of his algorithm for spanning trees, introduced on page 1628. In this algorithm, we keep track of the weight of the edge. Here is an outline:

Given that a graph $G = (V, E)$ is a simple, weighted, connected graph, and $|V| = n$, find a spanning tree T for G.

Algorithm

Set the counter $i = 0$. (i is the number of edges of the sought tree. Every time we add an edge, we increase this number by 1.)

Step 1: Select an edge, e_1, where e_1 does not create a cycle and has the smallest possible weight, add it to the tree, set $i = 1$, and add e_1 to the tree T.

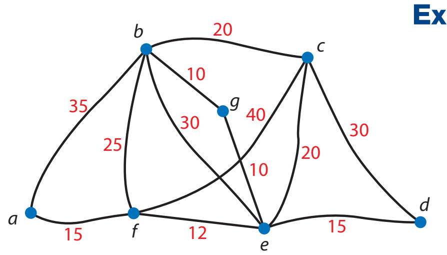
Step 2: For $1 \leqslant i \leqslant n - 2$, if edges $e_1, e_2, \ldots, e_i$ have been selected, then select edge e_{i+1} from the remaining edges so that the subgraph determined by $e_1, e_2, \ldots, e_{i+1}$ contains no cycles and the weight of e_{i+1} is the smallest possible.

Step 3: Replace i by $i + 1$.

If $i = n - 1$, the subgraph T determined by $e_1, e_2, \ldots, e_{i+1}$ is connected with $n - 1$ edges and n vertices, and hence is a spanning tree.

If $i < n - 1$, return to step 2.

Example 7

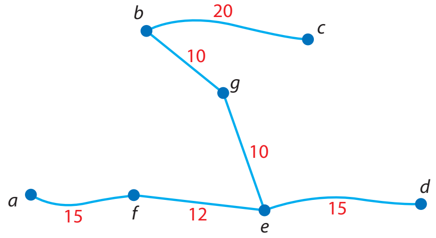


Apply Kruskal's algorithm to find a minimal spanning tree for graph G given left.

Solution

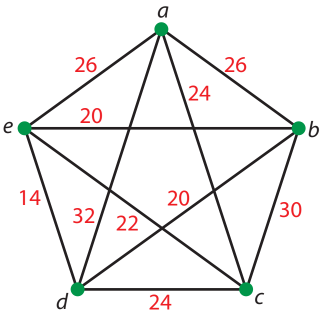
Here too we can use a table to summarize our steps. However, we will not use a table as we want you to experience applying the algorithm in as many different ways as possible.

1. Select edge bg as it has the lowest weight (ge too); $i = 1$, weight is 10.
2. Now select ge with smallest possible weight of 10, no cycle formed, add it to the tree; weight is 20, $i = 2$.
3. Now select fe with weight 12, no cycle, add it to T; weight is 32, $i = 3$.
4. Select af, then ed, add to T; weight is $32 + 15 + 15 = 62$, $i = 3 + 1 + 1 = 5$.
5. Select bc (or ed), add to T; weight is $62 + 20 = 82$, $i = 6$. Stop.



The tree is $T = \{bg, ge, fe, af, ed, bc\}$ with minimal weight of 82. The minimal spanning tree is shown in the figure left.

Example 8



Find a minimal spanning tree for the network left.

Solution

We will arrange the weights in non-decreasing order to make it easier to choose the edges to be added.

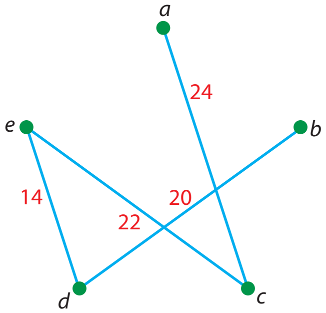
Weight	14	20	20	22	24	24	26	26	30	32
Edge	*de*	*db*	*eb*	*ec*	*dc*	*ac*	*ea*	*ab*	*bc*	*ad*

This procedure of applying Kruskal's algorithm is very helpful especially in graphs with a relatively small number of edges.

1. Select de, weight 14, add to T, $i = 1$.
2. Select db, no cycle formed, weight 20, add to T, $i = 2$.
3. Select eb, cycle formed, reject.

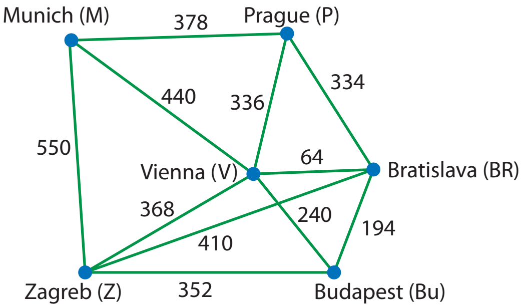
4. Select *ec*, no cycle formed, weight 22, add to *T*, $i = 3$.
5. Select *dc*, cycle formed, reject.
6. *ac*, no cycle formed, weight 24, add to *T*, $i = 4$. Stop.
7. Tree is formed and has a weight of $14 + 20 + 22 + 24 = 80$.

The diagram to the right shows the resulting minimal spanning tree.



Example 9

Use Kruskal's algorithm to find a minimum spanning tree for the graph below.



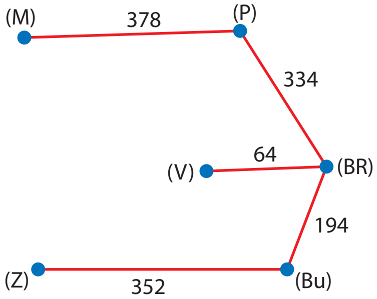
Solution

We will list all the edges in a table and then sort them in non-descending order. Then we decide whether or not we are going to include them in the minimum spanning tree.

Edge	Weight
V-Br	64
V-Bu	240
V-P	336
V-Z	368
V-M	440
Br-Bu	194
Bu-Z	352
Z-M	550
M-P	378
P-Br	334
Br-Z	410

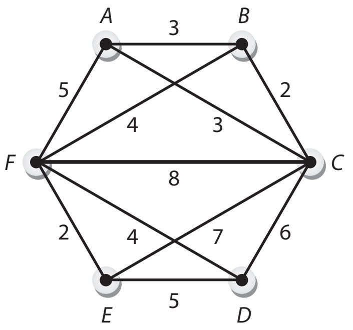
Edge	Weight	Decision
V-Br	64	yes, $i = 1$
Br-Bu	194	yes, $i = 2$
V-Bu	240	no, the cycle V-Br-Bu-V
P-Br	334	yes, $i = 3$
V-P	336	no, the cycle V-Br-P-V
Bu-Z	352	yes, $i = 4$
V-Z	368	no, the cycle V-Bu-P-V
M-P	378	yes, $i = 5$, STOP
Br-Z	410	
V-M	440	
Z-M	550	

The three edges left form a cycle with the edges already included in the spanning tree, and hence are not included. Also, after we included the fifth edge, we stop since a tree with six vertices contains five edges. We know that any additional edge to the tree will form a cycle with some of the existing edges.



By Kruskal's algorithm, the minimum spanning tree appears right.

So, the **minimum spanning tree** has a **weight** of $64 + 194 + 334 + 352 + 378 = 1322$.



Example 10

Use Kruskal's algorithm to find the weight of a minimum spanning tree in the graph left.

Solution

Edge	Weight
$\{A, B\}$	3
$\{A, C\}$	3
$\{A, F\}$	5
$\{B, C\}$	2
$\{B, F\}$	4
$\{C, D\}$	6
$\{C, E\}$	7
$\{C, F\}$	8
$\{D, F\}$	4
$\{D, E\}$	5
$\{E, F\}$	2

Edge	Weight	Decision
$\{B, C\}$	2	yes, $i = 1$
$\{E, F\}$	2	yes, $i = 2$
$\{A, B\}$	3	yes, $i = 3$
$\{A, C\}$	3	no, creates cycle $BCAB$
$\{B, F\}$	4	yes, $i = 4$
$\{D, F\}$	4	yes, $i = 5$, STOP
$\{A, F\}$	5	
$\{D, E\}$	5	
$\{C, D\}$	6	
$\{C, E\}$	7	
$\{C, F\}$	8	

So, the **minimum spanning tree** has a **weight** of $2 + 2 + 3 + 4 + 4 = 15$.

It is also possible that instead of the edge $\{A, B\}$ we include the edge $\{A, C\}$.

Notice here that edge $\{E, F\}$ was added, even though it was not adjacent to any existing edge in the tree. The algorithm will guarantee that the tree will eventually be formed by focusing on $n - 1$ edges with no cycles.

Prim's algorithm (Optional)

This topic is no longer in the IB Syllabus for 2014.

Prim's algorithm is similar to Kruskal's with the exception that it requires the added edges to be adjacent to existing edges of the tree.

Algorithm

Set the counter $i = 0$. (i is the number of edges of the sought tree. Every time we add an edge, we increase this number by 1.)

Step 1: Select an edge, e_1, where e_1 does not create a cycle and has the smallest possible weight, add it to the tree, set $i = 1$, and add e_1 to the tree T.

Step 2: For $1 \leqslant i \leqslant n - 2$, if edges $e_1, e_2, \ldots, e_i$ have been selected, then select edge e_{i+1} from the remaining edges which is adjacent to one of the edges in the tree and so that the subgraph determined by $e_1, e_2, \ldots, e_{i+1}$ contains no cycles and the weight of e_{i+1} is the smallest possible.

Step 3: Replace i by $i + 1$.

If $i = n - 1$, the subgraph T determined by $e_1, e_2, \ldots, e_{i+1}$ is connected with $n - 1$ edges and n vertices, and hence is a spanning tree.

If $i < n - 1$, return to step 2.

Example 11

Use Prim's algorithm to find a minimum spanning tree in the graph in Example 9. The data from the figure can be stored into the following cost adjacency matrix.

$$\begin{array}{c|cccccc}
 & \text{V} & \text{Br} & \text{Bu} & \text{Z} & \text{M} & \text{P} \\
\hline
\text{V} & - & 64 & 240 & 368 & 440 & 336 \\
\text{Br} & 64 & - & 194 & 410 & - & 334 \\
\text{Bu} & 240 & 194 & - & 352 & - & - \\
\text{Z} & 368 & 410 & 352 & - & 550 & - \\
\text{M} & 440 & - & - & 550 & - & 378 \\
\text{P} & 336 & 334 & - & - & 378 & -
\end{array}$$

Solution

Again we are going to start with the Vienna–Bratislava edge that has a length of 64 and then we will add one edge at a time. Once we reach five edges in the set we will stop. (*wt* corresponds to weight.)

Step 1: $T = \{\{\text{V}, \text{Br}\}\}, \quad wt(\{\text{V}, \text{Br}\}) = 64$

Step 2: $T = \{\{\text{V}, \text{Br}\}, \{\text{Br}, \text{Bu}\}\}, \quad wt(\{\text{Br}, \text{Bu}\}) = 194$

Step 3: $T = \{\{\text{V}, \text{Br}\}, \{\text{Br}, \text{Bu}\}, \{\text{Br}, \text{P}\}\}, \quad wt(\{\text{Br}, \text{P}\}) = 334$

Step 4: $T = \{\{\text{V}, \text{Br}\}, \{\text{Br}, \text{Bu}\}, \{\text{Br}, \text{P}\}, \{\text{Bu}, \text{Z}\}\}, \quad wt(\{\text{Bu}, \text{Z}\}) = 352$

Step 5: $T = \{\{\text{V}, \text{Br}\}, \{\text{Br}, \text{Bu}\}, \{\text{Br}, \text{P}\}, \{\text{Bu}, \text{Z}\}, \{\text{P}, \text{M}\}\}, \quad wt(\{\text{P}, \text{M}\}) = 378$
STOP

So, we have the same **minimum spanning tree** with a **weight** of 1322.

Notice how in Example 11 step 2, we added {Br, Bu} because it is adjacent to {V, Br} and in step 4 {Bu, Z} because it is adjacent to {Br, Bu}. This is not a requirement of Kruskal's algorithm. In this specific example, both algorithms happened to add the edges in the same order. This is not always the case. Notice how in Example 10 step 2, we added $\{E, F\}$ even though it is not adjacent to $\{B, C\}$, which is in the tree already. To show the difference between the two algorithms, the next example will apply Prim's algorithm to the same graph.

Example 12

Apply Prim's algorithm to the graph given in Example 10. For demonstration purposes, the cost adjacency matrix is produced here.

$$C_G = \begin{array}{c|cccccc}
 & A & B & C & D & E & F \\
\hline
A & - & 3 & 3 & - & - & 5 \\
B & 3 & - & 2 & - & - & 4 \\
C & 3 & 2 & - & 6 & 7 & 8 \\
D & - & - & 6 & - & 5 & 4 \\
E & - & - & 7 & 5 & - & 2 \\
F & 5 & 4 & 8 & 4 & 2 & -
\end{array}$$

Solution

Note: {*A*, *C*} is added as it is adjacent to {*B*, *C*}. Notice that at this stage in Kruskal's algorithm, we add {*E*, *F*} instead because it is the next 'lightest' edge.

Since there are two edges with the same weight of 2, we can start with either of them. We will start with the edge $\{B, C\}$.

Step 1: $T = \{\{B, C\}\}, \quad wt(\{B, C\}) = 2$

Step 2: $T = \{\{B, C\}, \{A, C\}\}, \quad wt(\{A, C\}) = 3$

Step 3: $T = \{\{B, C\}, \{A, C\}, \{B, F\}\}, \quad wt(\{B, F\}) = 4$

Step 4: $T = \{\{B, C\}, \{A, C\}, \{B, F\}, \{F, E\}\}, \quad wt(\{F, E\}) = 2$

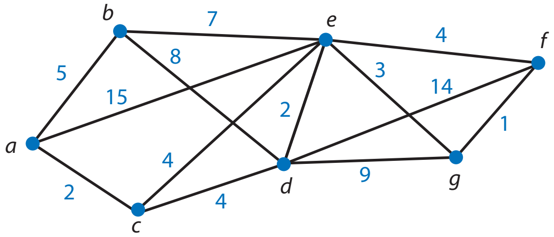
Step 5: $T = \{\{B, C\}, \{A, C\}, \{B, F\}, \{F, E\}, \{F, D\}\}, \quad wt(\{F, D\}) = 4$

STOP

So, the **minimum spanning tree** has the same **weight** of 15, but the process of adding edges to the tree had a different order.

Note: Kruskal's algorithm appears to be the easier of the two. However, this is only true for small graphs. As the graph size increases, spotting a cycle in Kruskal's algorithm is more difficult than in Prim's algorithm.

Example 13



Apply Kruskal's and Prim's algorithms to find a minimum spanning tree for the graph left.

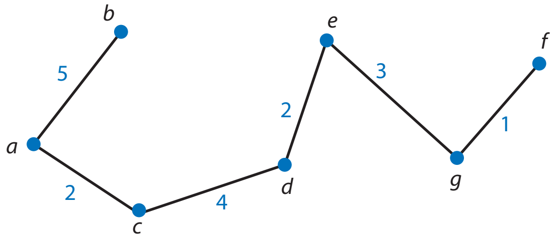
Solution

In both cases, since we have seven vertices, we will stop after finding six edges. We will set up a table of weights that will help us in finding the spanning trees we need.

Weight	1	2	2	3	4	4	4	5	7	8	9	14	15
Edge	*fg*	*de*	*ac*	*eg*	*ef*	*cd*	*ce*	*ab*	*be*	*bd*	*dg*	*df*	*ae*

Kruskal's algorithm

Weight	Edge	Cycle	Tree	Total weight	*i*
1	*fg*	no	*fg*	1	1
2	*de*	no	*fg, de*	3	2
2	*ac*	no	*fg, de, ac*	5	3
3	*eg*	no	*fg, de, ac, eg*	8	4
4	*ef*	yes, reject	*fg, de, ac, eg*	8	4
4	*cd*	no	*fg, de, ac, eg, cd*	12	5
4	*ce*	yes, reject	*fg, de, ac, eg, cd*	12	5
5	*ab*	no	*fg, de, ac, eg, cd, ab*	17	6
		Stop	Tree found	17	



On the left is the minimum spanning tree.

Prim's algorithm

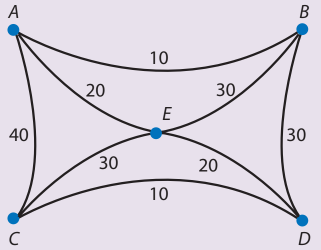
Weight	Edge	Adjacent	Cycle	Tree	Total weight	*i*
1	*fg*		no	*fg*	1	1
2	*de*, *ac*	no		*fg*	1	1
3	*eg*	yes	no	*fg*, *eg*	4	2
2	*de*	yes	no	*fg*, *eg*, *de*	6	3
2	*ac*	no		*fg*, *eg*, *de*	6	3
4	*ef*	yes	yes, reject	*fg*, *eg*, *de*	6	3
4	*cd*	yes	no	*fg*, *eg*, *de*, *cd*	10	4
2	*ac*	yes	no	*fg*, *eg*, *de*, *cd*, *ac*	12	5
4	*ce*	yes	yes, reject	*fg*, *eg*, *de*, *cd*, *ac*	12	5
5	*ab*	yes	no	*fg*, *eg*, *de*, *cd*, *ac*, *ab*	17	6
			Stop	Tree found	17	

Notice that we found a minimum spanning tree with the same weight as Kruskal's. In this specific example, it turned out to be the same tree. However, this must not be the case. The only common result should be the weight of the tree. Also worth noting here is that in Kruskal's algorithm, once you finish investigating a minimum weight you move to the next level, while in Prim's algorithm, if the adjacency test fails, then you need to revisit the level at a later stage, as happened to edges *ac* and *de* (weight of 2) and *ce* (weight 4).

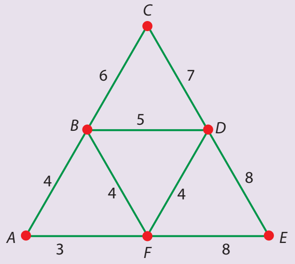
Exercise 4.4

For questions 1–5, use Kruskal's algorithm to find a minimum spanning tree (mst) for each given weighted graph.

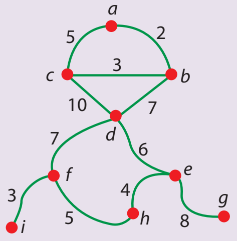
1



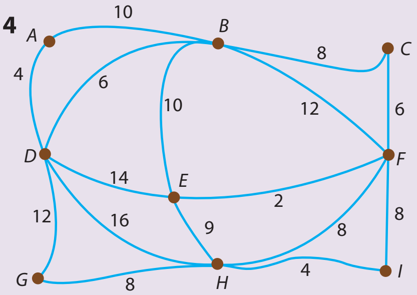
2



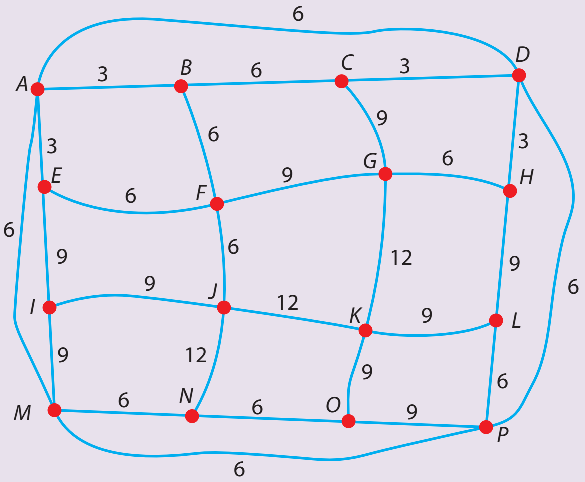
3



4



5



For questions 6–10 (optional), use Prim's algorithm to find a minimum spanning tree (mst) for each given weighted graph.

6 Find a mst for the graph in question 1.

7 Find a mst for the graph in question 2.

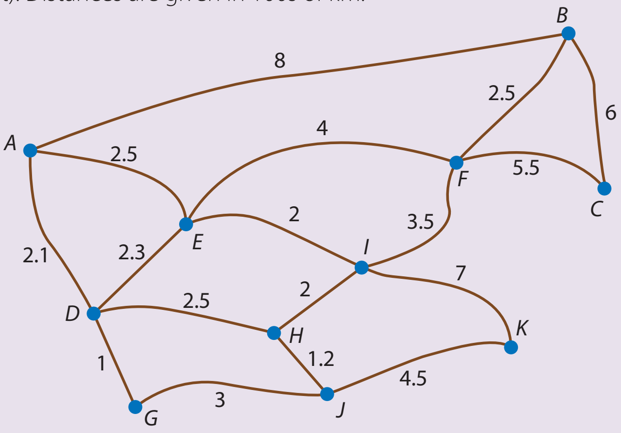
8 Find a mst for the graph in question 3.

9 Find a mst for the graph in question 4.

10 Find a mst for the graph in question 5.

11 (Optional) Describe the differences between the results of questions 1 and 6, 2 and 7, 3 and 8, 4 and 9, as well as 5 and 10.

12 The following is the network for a large bus company. To minimize cost, some routes must be discontinued. Find out which routes should be kept to ensure that transport between all the cities is still possible (though not necessarily direct). Distances are given in 100s of km.

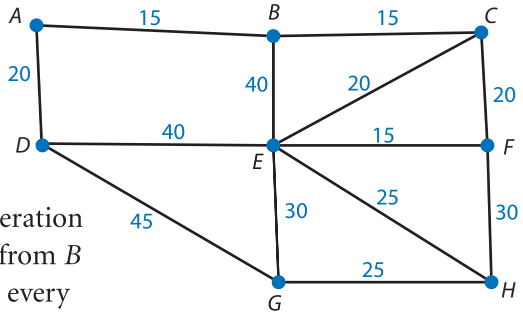


4.5 Shortest path, route inspection and the travelling salesman problem

A **shortest path** is a path from one vertex to another in a weighted graph, using the smallest possible weight. As a path, no edges or vertices are visited more than once. The shortest path, especially in complex networks, is not always evident. That is why Edsger Dijkstra, a Dutch mathematician, in his shortest-path algorithm created a way for finding the shortest path. In this section, we will discuss the algorithm and apply it to a few situations. However, you need to keep in mind that in textbook examples, the solution may be readily obvious by inspection or trial and error. However, by learning the algorithms, like many other aspects of graph theory, you are developing the skills which can later be used in more complex situations. So, even if you can immediately spot the solution to a problem, we strongly recommend that you follow the algorithm's steps in order to understand how to apply it later.

Example 14

In the weighted graph right we are required to find a path between vertices *A* and *H* which has the smallest total weight.

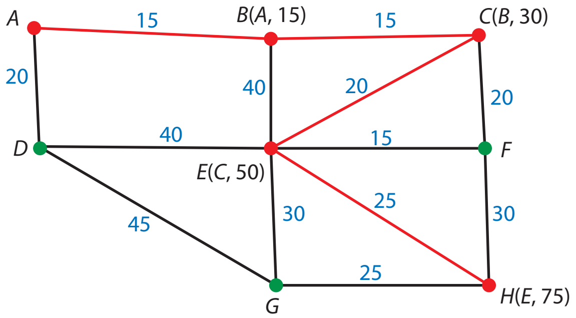


Solution

We can proceed from *A* to the 'nearest' vertices, taking into consideration the least weight possible. So, from *A* we can go to *B* or to *D*. Then from *B* we can go to *C* or *E*, while from *D* we can go to *E* or *G*. Arriving at every new vertex, we look at the total weight of the path. If there is a new path to arrive at the old vertex, we consider the total weight; if it is smaller than the one we already have, we cross out the old path and adopt the new one instead. The whole process is given in the table below. (Several ways of arranging your work are available and will be demonstrated.)

Step 1	Step 2	Step 3	Step 4	Step 5
A	*B*(*A*, 15)	*C*(*B*, 30)	*F*(*C*, 50)	
			E(*C*, 50)	
	D(*A*, 20)	~~*E*(*B*, 55)~~		*H*(*E*, 75)
		G(*D*, 65)	~~*H*(*G*, 90)~~	

Note that for every vertex we visit, we label it with a temporary label, which includes the previously visited vertex and the total weight, so far. In the third step, we labelled *E*(*B*, 55) because, so far, this is the smallest weight (coming through *B*), but then in the fourth step, once we reached *E* with a smaller weight of 50, we cross out *E*(*D*, 55). The same happens to the paths of the vertex *H* in the fifth step.



So, the path with the smallest weight is *ABCEH*.

Example 14 demonstrates the general rule used in Dijkstra's algorithm. It proceeds by finding the shortest path from *A* to its adjacent vertices, then the shortest path to a second 'level' set of vertices, and so on until the length of the shortest path to *H* is found.

The algorithm performs a sequence of iterations. A key set of vertices is assembled by adding one vertex at each iteration. A labelling process is executed at each iteration. In this labelling process, a vertex *w* is labelled with the length of a shortest path from *A* to *w* that contains only vertices from the key set. The vertex added to the set is one with the minimal label among those vertices not already members of the set. In the next few paragraphs, we give a formal statement of the algorithm followed by a description of the algorithm.

Dijkstra's algorithm[1]

```
procedure Dijkstra (G: weighted connected simple graph, with all
        weights positive)
{G has vertices $a = v_0, v_1, ..., v_n = z$ and weights $w(v_i, v_j)$,
        where $w(v_i, v_j) = \infty$ if $\{v_i, v_j\}$ is not an edge in G.}
for $i := 1$ to $n$
        $L(v_i) := \infty$
$L(a) := 0$
$S := \varnothing$
{The labels are now initialized so that the label of a is zero and all other
        labels are $\infty$, and S is the empty set.}
while $z \notin S$
begin
        $u :=$ a vertex not in S with $L(u)$ minimal
        $S := S \cup \{u\}$
        for all vertices v not in S
            if $L(u) + w(u, v) < L(v)$ then $L(v) := L(u) + w(u, v)$
        {This adds a vertex to S with minimal label and updates the labels
            of vertices not in S.}
end {$L(z)$ = length of shortest path from a to z}
```

Interpretation of Dijkstra's algorithm

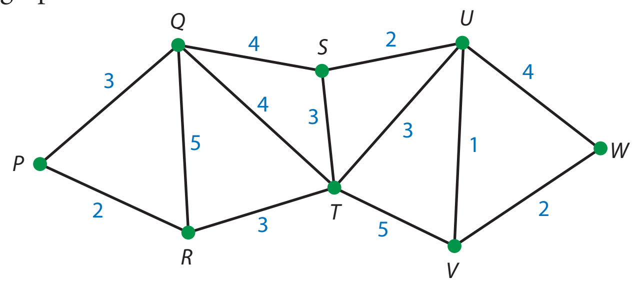
We need to find the shortest path from *a* to *z*. The algorithm begins by labelling *a* with 0 and the other vertices with ∞. We use the notation $L(v)$ to represent the shortest path from the source, *a*, to the present vertex *v*. *S* is the key set containing all vertices with minimum path length discovered so far. We begin with $S = \varnothing$. Every iteration will update the set *S* by adding a new vertex *u* with the smallest label. Once this is done, we update the labels of all vertices not in *S*, say *v*, such that $L(v)$ is the length of the shortest path to *v* through vertices already in *S*. This process is iterated successively adding vertices to the key set until *z* is added. In the following example, we will demonstrate the use of this algorithm. There are several interpretations of how to keep track of the successive steps; we will use the following convention:

[1] Kenneth Rosen, *Discrete Mathematics and its Applications*, 5th edition (McGraw-Hill, 2003) p. 597.

each vertex, *v*, is labelled with an ordered pair (*x*, *l*), where *x* represents the vertex just preceding *v* and *l* is the shortest length of the path from *a*. All labels are temporary, until the algorithm identifies their path as shortest and they are changed into permanent labels, which we will denote by circling the vertex. Any temporary label that does not become permanent will be crossed out. We will also use tables to demonstrate the steps.

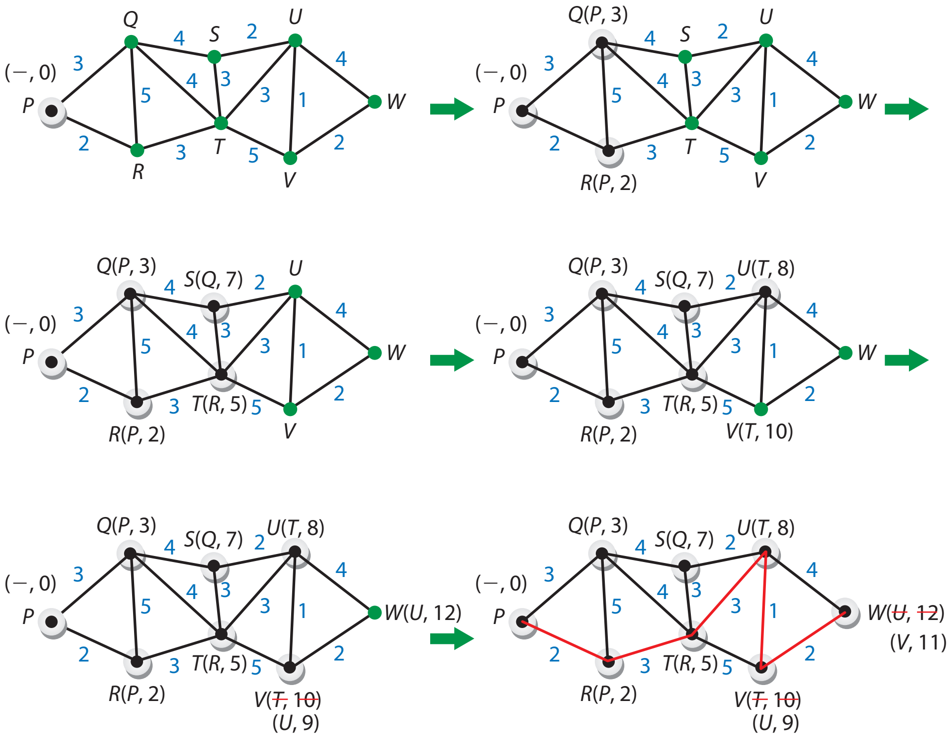
Example 15

Use Dijkstra's algorithm to find the shortest path between *P* and *W* in the following graph.



Solution

Note that only in this example will we draw the graph at different stages. You would not have to do that if you were performing the algorithm. In the diagrams below, we use the convention that if a vertex is not labelled, then it has the label (−, ∞).



Below is the table with the steps. Each cell contains the length of the path and the preceding vertices. The highlighted cells are the ones describing the shortest path. Each cell also lists the path lengths that are calculated at this stage.

Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
$L(P) = 0$						
$L(R) = \infty$	$L(R) = 2, \{P\}$					
$L(Q) = \infty$	$L(Q) = \infty$	$L(Q) = 3, \{P\}$				
$L(T) = \infty$	$L(T) = \infty$	$L(T) = 5, \{P, R\}$				
$L(S) = \infty$	$L(S) = \infty$	$L(S) = \infty$	$L(S) = 7, \{P, Q\}$			
$L(U) = \infty$	$L(U) = \infty$	$L(U) = \infty$	$L(U) = \infty$	$L(U) = 8, \{P, R, T\}$		
$L(V) = \infty$	$L(V) = \infty$	$L(V) = \infty$	$L(V) = \infty$	$L(V) = 10, \{P, R, T\}$	$L(V) = 9, \{P, R, T, U\}$	
$L(W) = \infty$	$L(W) = \infty$	$L(W) = \infty$	$L(W) = \infty$	$L(W) = \infty$	$L(W) = 12, \{P, R, T, U\}$	$L(W) = 11, \{P, R, T, U, V\}$

Step 1: We start by labelling $P\,(-, 0)$ since there is no vertex to precede it. Make it permanent.

Step 2: From A there are two unlabelled vertices, Q and R. Since $L(P) = 0$, vertex R gives the smallest $L(P) + w(P, R) = 0 + 2$, then we label $R(P, 2)$ and we add it to the path S. Make it permanent.

Step 3: Now S has two vertices, P and R. They have two unlabelled adjacent vertices, Q and T. Vertex Q has the smallest $L(P) + w(P, Q) = 0 + 3 = 3$ $(L(R) + w(R, T) = 2 + 3 = 5$, and $(L(R) + w(R, Q) = 2 + 5 = 7)$. We make $Q(P, 3)$ permanent.

Step 4: Now S has three vertices, P, R, and Q. They have two unlabelled adjacent vertices, S and T. Similar to the previous process, we make $T(R, 5)$ permanent.

Step 5: Now S has four vertices, P, R, Q, and T. They have three unlabelled adjacent vertices, S, U, and V. Similar to the previous process, we make $S(Q, 7)$ permanent.

Step 6: Now S has five vertices, P, R, Q, T, and S. They have one unlabelled adjacent vertex, W. Similar to the previous process, we make $U(T, 8)$ permanent and *update* $L(V)$.

Step 7: Similar to above, we make $V(U, 9)$ permanent, and *update* $L(W)$ and make it permanent.

So, the shortest path is *PRTUVW* and it has a length of 11.

A practical interpretation of Dijkstra's algorithm

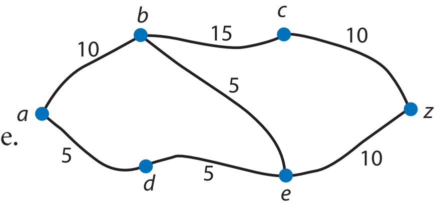
To find a shortest path from vertex a to vertex z in a weighted graph, proceed as follows:

1. Set $v_1 = a$ and assign to this vertex the label $(-, 0)$. Assign every other vertex a temporary label of ∞, where ∞ is reckoned to be larger than any real number!

2. Until *z* has been assigned a *permanent* label, do the following:
   (i) Take the vertex v_i that most recently acquired a permanent label, say *d*. For each vertex that is adjacent to v_i which has not yet received a permanent label, if $d + w(v_iv) < L(v)$, the current temporary label of *v*, update $L(v)$ to $d + w(v_iv)$.
   (ii) Take a vertex *v* that has a temporary label smallest among all temporary labels in the graph and make its temporary label permanent. If there are several vertices *v* that tie for the smallest temporary label, make any choice.

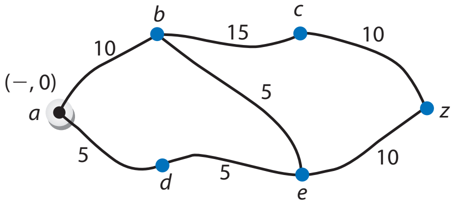
Example 16

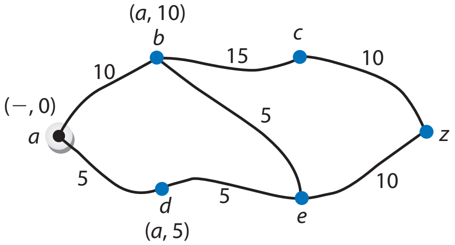
Find a shortest path from *a* to *z* in the graph on the right.



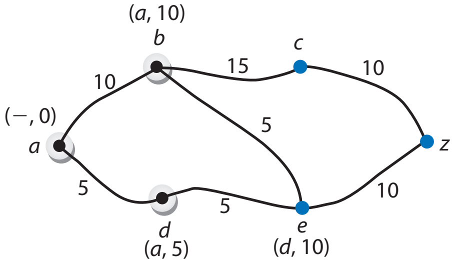
Solution

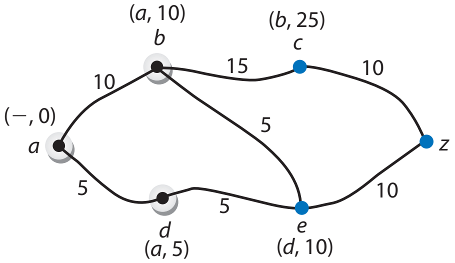
We will follow the algorithm by labelling the graph without a table this time.



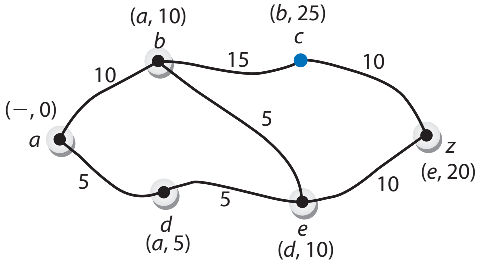


First we label and make *a* permanent. Next, we label vertex *d* with (*a*, 5) to indicate the length of the path and that it is visited through *a*. Similarly, we label *b* (*a*, 10).





Next we make *d* permanent and update vertex *e*. Then we make *b* permanent and update vertices *c* and *e* (no change in *e*).



Next we make *e* permanent and update *z*. At this point, we can make the label at *z* permanent; a shortest path has been found.

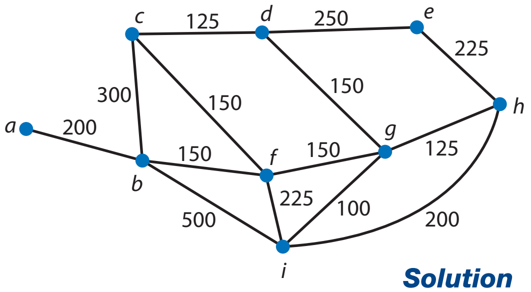
Notice in the above example that it is not necessary to change the label of a vertex v if $d + w(v_i v) \ngtr L(v)$, and that it is also unnecessary to make all vertices in the graph permanent as long as they don't contribute towards a shortest path.

The Chinese postman problem

This is also known as the **route inspection problem**. Contrary to its name, this has little to do with a 'real' Chinese postman. The reference is to the Chinese mathematician Kwan Mei-Ko who, in 1962, posed an inspection problem in terms of a postman covering each road of a network exactly once and coming back to his starting point.

We will start this subsection with an example.

Example 17



A cable network has to be inspected for possible faulty wires. The diagram left represents a sketch of the wires along with the length of each section (in metres) and the junction names. We would like to inspect every cable at least once and come back to the starting junction, a.

Solution

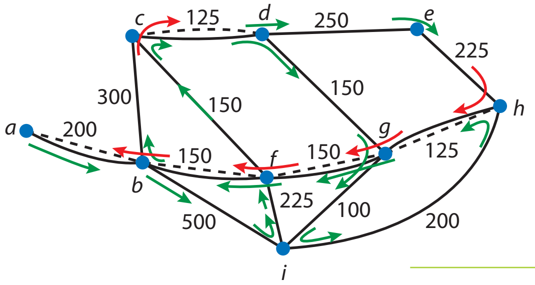
The problem is similar to finding an Eulerian circuit. However, this is not possible since we have four vertices with odd degree: a, c, d, and h.

Since we are starting at a, ab has to be retraced. This makes b also with odd degree. Knowing that we have to get back to b to reach a, leaves us now with four vertices with odd degree. To be able to inspect the cables, we need to retrace some of the paths between these junctions. We will consider all possible pairings that result in shortest lengths.

bc and dh: $300 + 275 = 575$

bd and ch: $425 + 400 = 825$

bh and cd: $425 + 125 = 550$



So, bh and cd is the shortest, and hence we will retrace these paths.

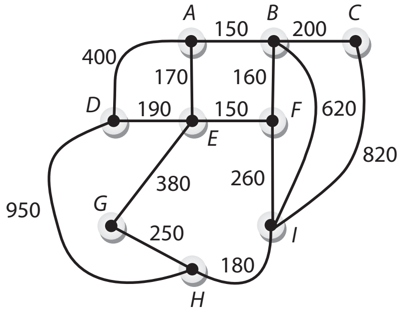
The original network has 2850 metres, and we will retrace $ab = 200$ and $bh + cd = 550$, giving a total length of 3600 metres. Such a route is: *abifcdgihgfbcdehgfba*. The route is given left.

As you may have noticed, when the number of edges to be inspected is high, the process will be tedious to follow. The algorithm proposed by Kwan Mei-Ko makes the process more systematic.

Chinese postman algorithm

1. Find all vertices of odd degree.
2. For each pair of odd vertices, find the path of shortest length.
3. Pair up all odd vertices from step 2, so that the sum of the lengths is minimum.
4. In the original graph, duplicate the shortest-length paths found in step 3.
5. Find an Eulerian circuit containing all edges of the 'new graph'.

Example 18



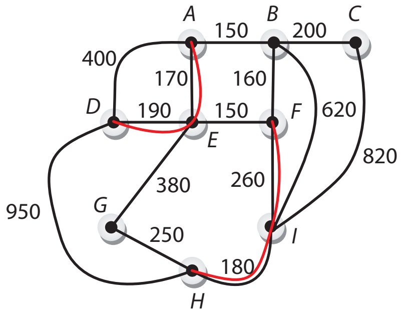
A guard patrols a campus of a large school as given by the graph right. The weights of the edges are distances given in metres. If the guard must pass through each street at least once during his shift, find the minimum distance he will cover.

Solution

The sum of all the distances in the graph is 4880 metres.

There are four odd vertices: *A*, *D*, *F*, and *H*. We need to investigate all the possible pairings and then choose the shortest paths between pairs of vertices.

Pairing	Shortest path	Distance (m)
A, *D*	*AED*	360
A, *F*	*ABF*	310
A, *H*	*ABFIH*	750
D, *F*	*DEF*	360
D, *H*	*DEGH*	820
F, *H*	*FIH*	440

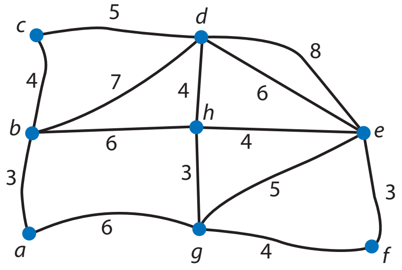


Now, we need to look at the pairings that will include all four vertices and give us the minimum sum of the distances. The pairings are *A*, *D* and *F*, *H*, and the paths that we will repeat are *AED* and *FIH* with their distances of 360 and 440 metres.

So, the minimum distance the guard will cover in one shift is 4880 + 360 + 440 = 5680 metres. We leave tracing a path with length of 5680 for you as an exercise.

Note: The IB syllabus limits the number of odd vertices to two. So, the focus in the exercises and examples will be limited to that case.

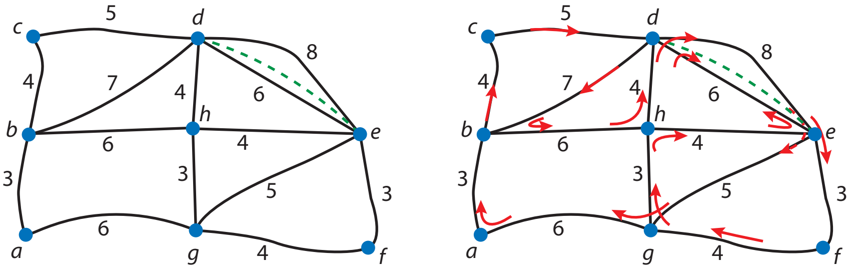
Example 19



A truck has to visit a neighbourhood with a street network as shown. What is a possible route that minimizes the distance it has to travel? Distances are in kilometres.

Solution

Vertices *d* and *e* are odd. So, we first duplicate the shortest path between them which is 6, and then try to find the minimum distance to be travelled.



Since all vertices are even by now, the graph is Eulerian. We can use the algorithm developed in Example 6 of Chapter 3, or any other method, to find the circuit. If we start at *b*, we can create a cycle *bcdb*, which can be joined at *b* with *hdefgh*, which can be joined at *h* with *edegab*. Our route is then *bcdbhdefghedegab* with length of 68 + 6 (retracing *de*) = 74 km. This is not unique. You can find other circuits with the same minimum length of 74 km.

The travelling salesman problem

Given a set of cities and the cost of travel between each pair of them, the travelling salesman problem, or TSP for short, is to find the cheapest way of visiting all of the cities and returning to your starting point.

The simplicity of the statement of the problem is misleading. The TSP is one of the most considered problems in computational mathematics and yet no successful solution method is known for the general case.

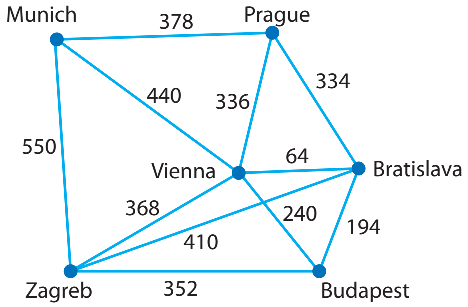
The TSP naturally arises in many transportation and logistics applications; for example, practical uses for the TSP include routing trucks for package pickups and material handling in warehouses. Other applications involve the scheduling of service calls at communications businesses.

Although transportation applications are the most natural setting for the TSP, the simplicity of the model has led to many interesting applications in other areas. A classic example is the scheduling of a machine to drill holes in a circuit board. In this case the holes to be drilled are the cities, and the cost of travel is the time it takes to move the drill head from one hole to the next. The technology for drilling varies from one industry to another, but whenever the travel time of the drilling device is a significant portion of the overall manufacturing process, then the TSP can play a role in reducing costs.

Basically, the travelling salesman problem is related to the search for Hamiltonian cycles in a graph. We will start with a simple example.

Example 20

A travelling salesman lives in Vienna. He needs to go on a business trip by car, visiting the following cities: Prague, Munich, Zagreb, Budapest, and Bratislava. On the figure right the distances between the cities are given in kilometres. (Not all routes have been included in the diagram.)



Solution

There are several possible Hamiltonian cycles and for each we calculate the total distance travelled.

Cycle	Distance (km)
V Br P M Z Bu V	1918
V P M Z Bu Br V	1874
V M Z Bu Br P V	2206
V Z Bu Br P M V	2066
V Bu Br P M Z V	2064

The shortest cycle is the second one from Vienna to Prague, Munich, Zagreb, Budapest, Bratislava, and back to Vienna, which has a total length of 1874 km. Since every cycle can have two directions, it is possible to visit all the cities in reverse order.

The solution presented in the example is a trial and error approach. Are there other approaches?

Remembering that a Hamiltonian cycle is a cycle that visits every vertex in a connected graph exactly once, we see that the classical TSP is a Hamiltonian cycle with minimum length. However, similar to the Chinese postman problem, we allow vertices to be visited more than once.

As you observed in the previous example, if you can inspect all possible routes involved in the TSP, you will be able to find the minimum total weight. However, as the number of vertices increases, checking all possibilities becomes impractical, if we don't say impossible. There is an assumption that the graph under consideration is *complete*, and as such, theoretically, the number of possible routes to inspect for a graph with n vertices will be $\frac{(n-1)!}{2}$ (considering routes in reverse order). For example,

if you have five cities, then the number of routes to be inspected will be 12 and if you have 10 cities the number will jump to 181 440. If there are 20 cities, the number will be $6.0 \cdot 10^{17}$. Even if you have a fast computer that can calculate 1 000 000 routes per second, it will take such a computer approximately 19 years to finish the task! So far, there is no known solution to the general TSP problem. Mathematicians resorted to finding near-minimum-weight solutions. Many algorithms have been developed. The nearest neighbour algorithm, nearest insertion algorithm, cutting-plane methods, and branch-and-cut methods are a few such algorithms. The IB syllabus does not require you to use such algorithms and thus we will not discuss these concepts in detail in this publication. We will just demonstrate the use of two of the algorithms without requiring you to do them.

In discussing the TSP in this publication, we will consider complete graphs with at least three vertices. Such graphs will have a Hamiltonian cycle. Moreover, since the number of vertices is finite, then the number of Hamiltonian cycles will also be finite. Thus, there must be at least one with minimum weight.

Also, since the weights of the edges in the complete graphs represent the shortest distances between the nodes of the original route network, the complete graph must satisfy the **triangle inequality**. As you recall from geometry, the sum of two sides of a triangle must be larger than or equal to the third side. Thus, for every choice of three vertices, v_i, v_j, and v_k, the following must be true:

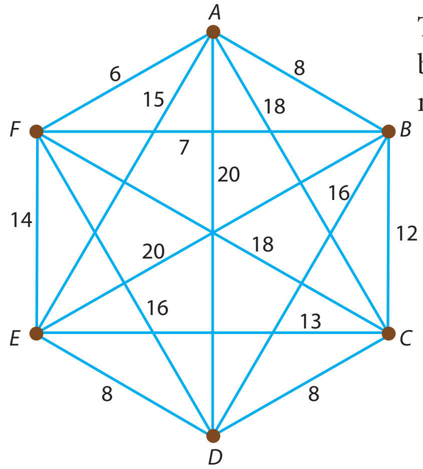
$$w(v_i, v_j) + w(v_i, v_k) \geqslant w(v_j, v_k)$$

The nearest neighbour algorithm

1. Choose a starting vertex.
2. Consider the edge of smallest weight incident to this vertex. If the other end of this edge is not visited yet, add it to the tour.
3. Repeat step 2 until all vertices have been visited.
4. Add the edge connecting the last visited vertex to the starting vertex.

where $w(v_i, v_j)$ is the weight of the corresponding edge.

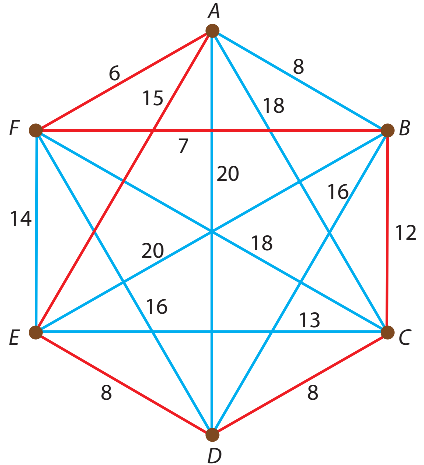
This algorithm will *sometimes* produce a minimal Hamiltonian cycle, but, in general, it may produce cycles with a considerably greater than minimum weight.



We will use the complete graph on the left to demonstrate the algorithm. The salesman is to start and end at A.

Starting at A, the first edge is AF since 6 is the minimum among 6, 8, 15, 18, and 20.

With the same reasoning, FB is chosen, with weight 7. BC is next with a weight of 12, followed by CD with 8, DE with 8 and finally, we get back to A with 15. The whole route has a total weight of 56. See the figure below.



The nearest insertion algorithm

This algorithm creates a cycle in the graph and then enlarges it to include a vertex which is closest to the given cycle.

1. Choose a starting vertex, u_1.
2. Consider the edge of smallest weight incident to this vertex. Add it to the cycle C. The vertex at the other end of this edge is added to C; call it u_2.
3. Select an edge with minimum weight that joins a vertex in C to one not in C; call the new vertex v.
4. Next, we enlarge the cycle to include the new vertex v. Now we consider the following expression:

   $x = w(u_i, v) + w(u_j, v) - w(u_i, u_j)$

   We choose the pair of vertices u_i and u_j for which x is minimum. We then include the edges (u_i, v) and (u_j, v), and we remove (u_i, u_j). (x represents the increase in the weight of the cycle when we add v.)
5. Repeat steps 3 and 4 until we include all vertices in the cycle.

Let us apply the algorithm to the previous graph.

We start with AF as it is the smallest, then we add B. Now we have a cycle AFB as shown in the first diagram overleaf.

Now consider all possible cycle expansions by comparing the x values for adding any of the remaining vertices. Here are the values:

Consider vertex C:
$AC + CF - AF = 18 + 18 - 6 = 30$, $AC + CB - AB = 18 + 12 - 8 = 22$,
$BC + CF - FB = 18 + 12 - 7 = 23$

Consider vertex D:
$20 + 16 - 8 = 28$, $16 + 20 - 6 = 30$, $16 + 16 - 7 = 25$

Consider vertex E:
$14 + 15 - 6 = 23$, $15 + 20 - 8 = 27$, $14 + 20 - 7 = 27$

So, 22 is the minimum, and since it corresponds to the connection of C and A and B, we add AC and BC and remove AB. Now the cycle is $AFBCA$ as shown in the second diagram.

Repeat the same steps for the new cycle:

Consider vertex D:
$8 + 16 - 12 = 12$, $8 + 20 - 18 = 10$, $16 + 20 - 6 = 30$, $16 + 16 - 7 = 25$

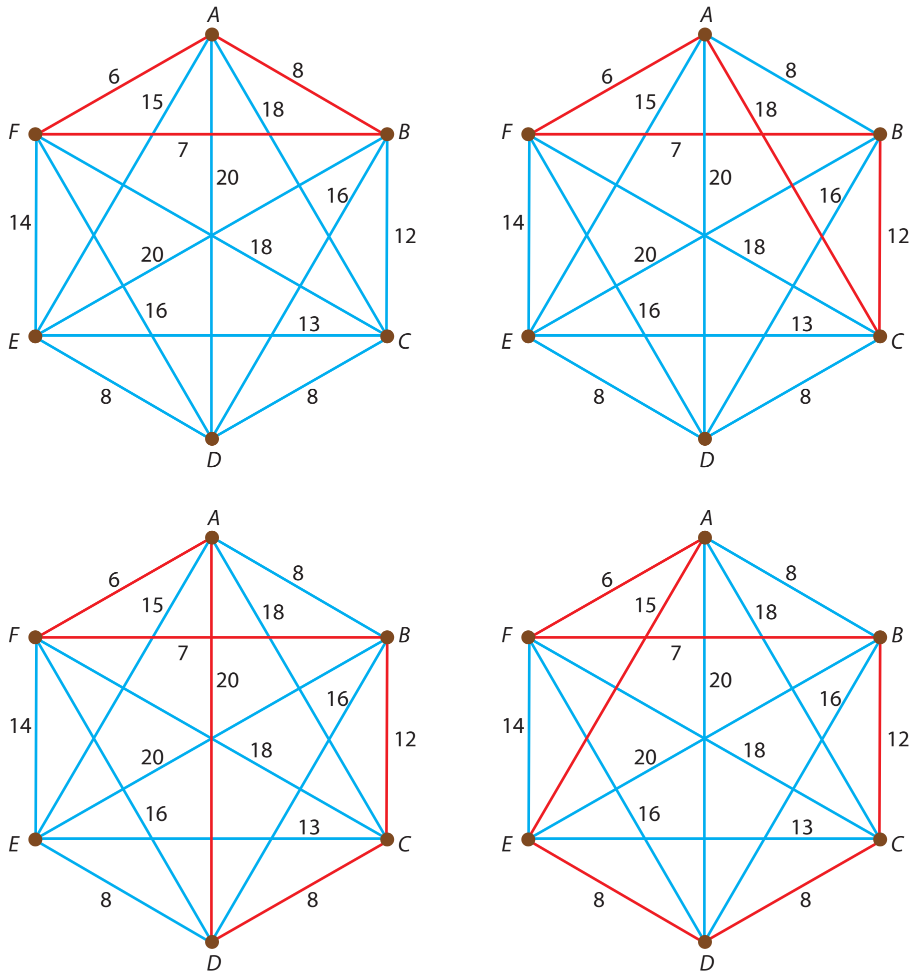
Consider vertex E:
$14 + 15 - 6 = 23$, $14 + 20 - 7 = 27$, $15 + 13 - 18 = 10$, $20 + 13 - 12 = 21$

Thus, DC and DA are added and AC removed. (Notice that we could have added E at this stage instead of D. See the third diagram.)

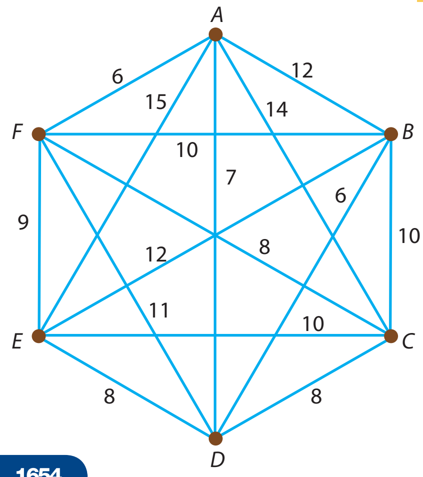
Lastly, consider *E*:

$15 + 8 - 20 = 3$, $14 + 20 - 7 = 27$, $14 + 15 - 6 = 23$, $20 + 13 - 12 = 21$, $13 + 8 - 8 = 13$

Thus, we add *ED* and *EA* and remove *DA*. The route now has a weight of 56 as before.



Caution: The equality between the routes created by these two algorithms are not always equal. And neither of them will definitely produce a minimum weight Hamiltonian cycle.

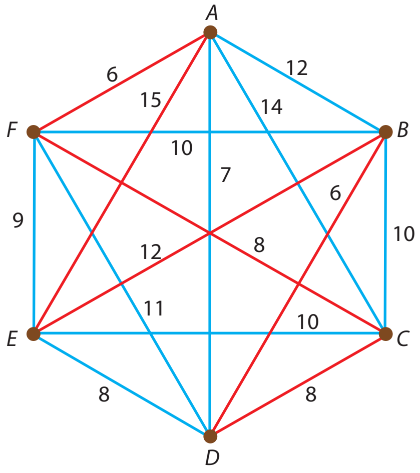


Example 21

Consider the graph in the figure left and use the nearest neighbour and nearest insertion algorithms to find a minimum TSP tour.

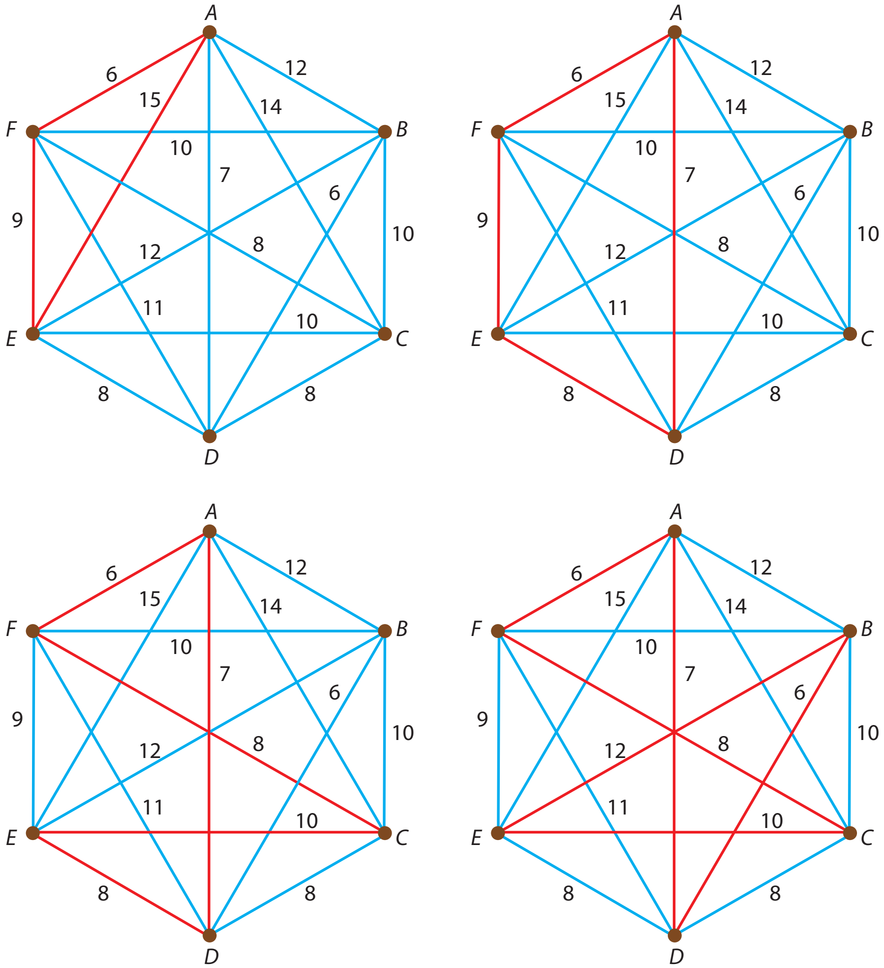
Nearest neighbour algorithm

Starting at *A*, the next vertex must be *F*. From *F*, the edge with smallest weight leads to *C*, then similarly from *C* to *D*, then to *B*, to *E*, and finally back to *A*. The total weight is 55.



Nearest insertion algorithm

First cycle could be *AFE* with weight of 30. Considering *x* values for possible expansion, we find that can be achieved by adding vertex *D* with $x = 0$. We add *AD* and *DE* and remove *AE*. The weight so far is 30. Applying the algorithm again, we can add *C* to the cycle by adding *FC* and *CE* and removing *FE*. The cycle *AFCEDA* has a weight of 39 so far. Lastly, we expand the cycle to include *B* by adding *BD* and *BE*, they have an *x* value of 10, and removing *ED*. So, the cycle now is *AFCEBDA* with a total weight of 49, which is less than what we achieved with the nearest neighbour algorithm.



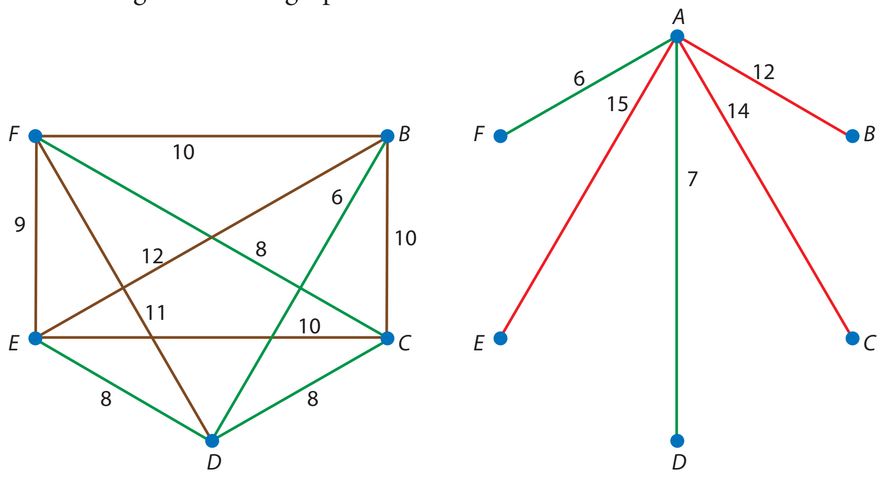
Lower and upper bounds

As Example 21 shows, the nearest neighbour algorithm, for example, did not lead us to a Hamiltonian cycle with minimum possible weight. As you observed, we were able to have an improved cycle. How far can we go?

A lower bound can be found by using algorithms that help us find minimal spanning trees. The argument is as follows: If we have a minimum weight Hamiltonian cycle in a complete graph, then we can remove one vertex v, for example, and all edges incident to it. Then we have a minimal spanning tree passing through the rest of the vertices. The weight of the Hamiltonian cycle is the weight of this minimal spanning tree plus the total weight of the edges we just removed. This argument leads us to the following **lower bound algorithm**.

1. Choose a vertex v in the complete graph and find the total of the two smallest edge weights incident to v.
2. Find the total weight of a minimum spanning tree going through all the remaining vertices.
3. The sum of the row totals is a lower bound.

Let us take the graph in Example 21 for instance. Remove A and its incident edges from the graph.



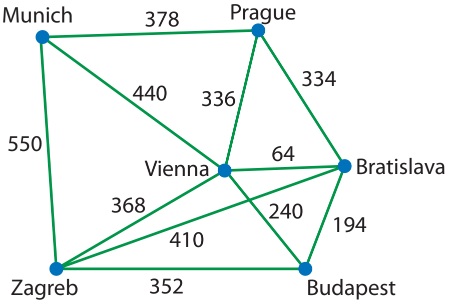
A minimum spanning tree for the remaining vertices is marked in green and has a weight of 30. The two edges with minimum total weight are AF and AD with a weight of 13. Hence, a lower bound for the cycle is 43, which is less than the smallest we found, 49.

So, now we can say that the minimum weight for a Hamiltonian cycle lies between 43 and 49.

As you notice from above, we used the weight of the Hamiltonian cycle we found earlier as an upper bound. There are a few ways of looking at an upper bound. One is to say the upper bound is the length of any cycle you manage to find, or, in general, is twice the length of a minimal spanning tree. The reason for this is a worst-case scenario. That is, the travelling salesman would visit every city and return that way, tracing each edge of the spanning tree twice.

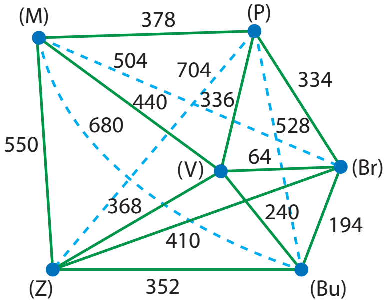
Example 22

We will try to find a lower bound, an upper bound, and a possible shortest route for the Vienna salesman in Example 20.



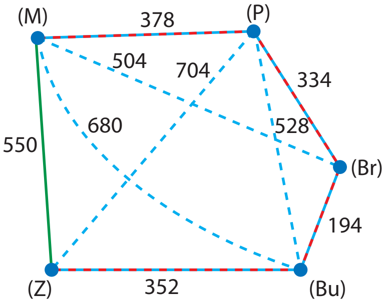
Solution

Note: As you may notice, the graph is not complete. However, in TSP we are allowed to add new edges which represent the minimum weight between two vertices that are not adjacent in the original graph. For example, Budapest and Prague are not directly connected; however, a path of minimum length of 334 + 194 = 528 through Bratislava can be added. Similarly, Budapest–Munich can have an extra edge of length 440 + 240 = 680 added, as well as Prague–Zagreb with 704 and Munich–Bratislava with 504. The new complete graph is given right.



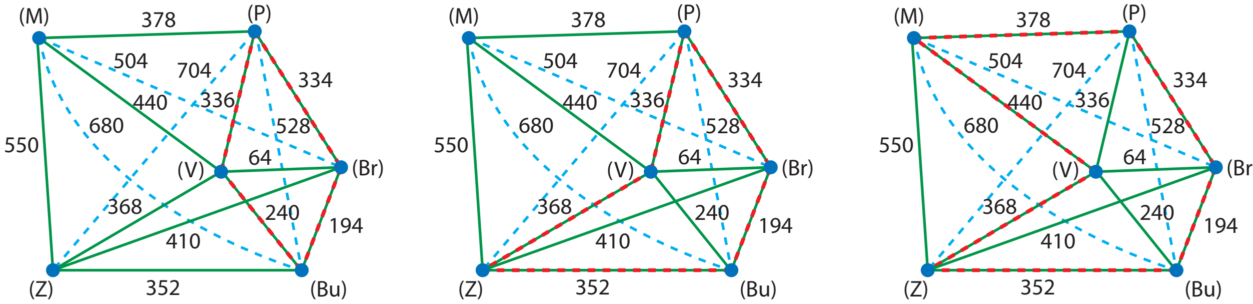
To find a lower bound, remove Vienna, for example, and all edges incident to it, and then find a minimum spanning tree for the rest. The tree weight is 1258.

The minimum total weight of two of the edges from Vienna is 64 + 240 = 304 and together with the minimum spanning tree this gives us a lower bound of 1562. Notice that if we remove another city, we may receive a different lower bound! An upper bound may be the route weight of 1874 that we found earlier.



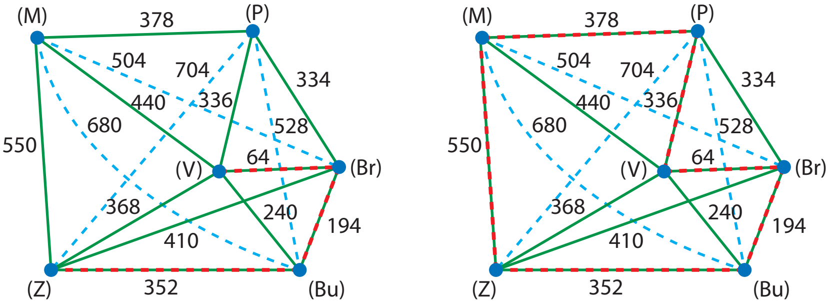
Thus, we are confident that our minimal Hamiltonian cycle would be between 1562 and 1874.

Apply the nearest insertion algorithm. You will expand the cycles; starting with V, Br, Bu, you will get the following sequence:



Unfortunately, the algorithm here did not yield the best results. The length of the route is 2066, which is greater than the upper bound.

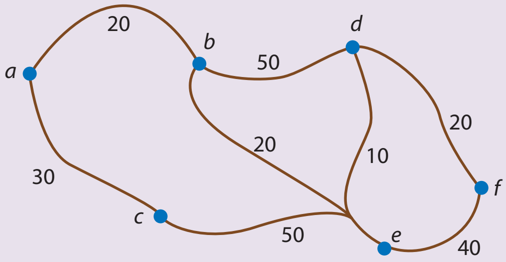
Applying the nearest neighbour algorithm yields the following:



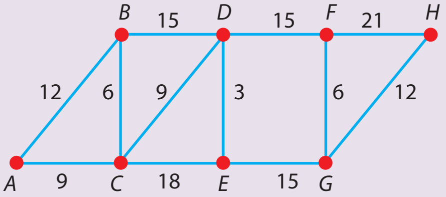
The total weight of this route is 1874, the same as that obtained by the 'brute force' method we used at the outset of this section and which we used as a lower bound. Notice here that the nearest neighbour algorithm gave better results than the nearest insertion algorithm. This again points to the fact that we do not have a unique solution to the TSP.

Exercise 4.5

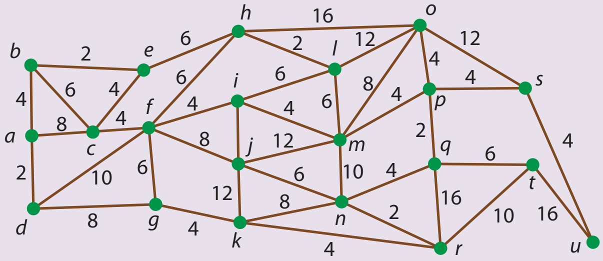
1 Find the length of the shortest path between *a* and *f* in the following weighted graph. Write down the path you suggest.



2 Find the length of the shortest path between *A* and *H* in the following weighted graph. Write down the path you find.



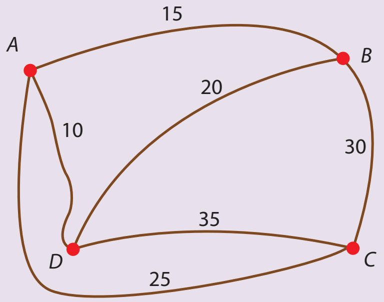
3 A circuit board has the following sub-network with the time, in millionths of a second, it takes a DC signal to flow through. Find the minimum time it takes a signal to go from *a* to *u*. Write down the path that gives this time.



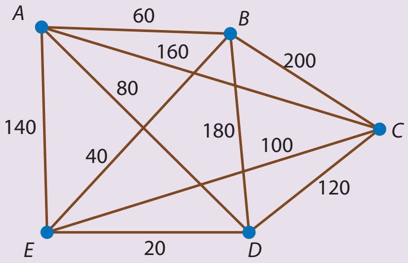
4 In question 1, find the shortest route between *a* and *d*.

5 In question 2, find the shortest route between *A* and *F*, and between *B* and *H*.

6 Solve the TSP for the following graph.

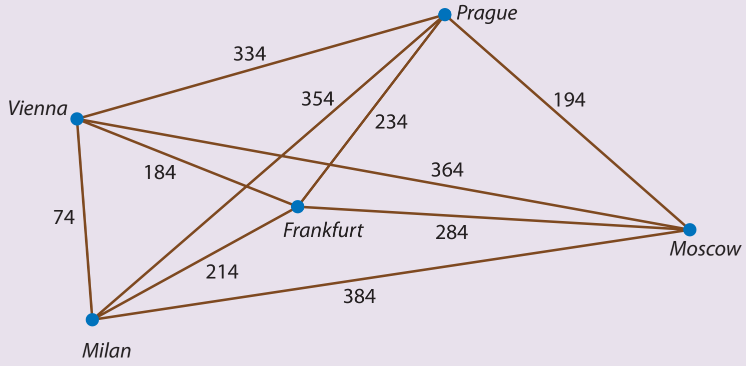


7 Solve the TSP for the following graph.



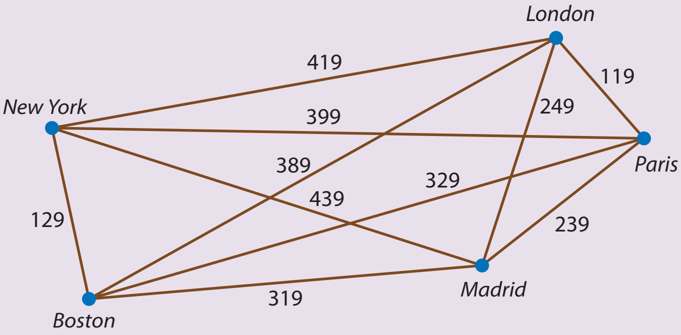
8 The flight paths between cities is given by the graph below. The weight on each edge is the cheapest possible two-way flight between the two cities. The prices are in Euros.

Find the route with the minimum total cost for a tourist who wants to visit each of the cities once.



9 The flight paths between cities is given by the graph below. The weight on each edge is the cheapest possible two-way flight between the two cities. The prices are in Euros.

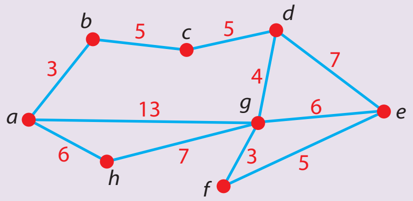
Find the route with the minumum total cost for a tourist who wants to visit each of the cities once.



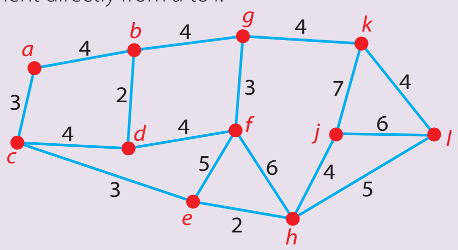
10 The nodes *A*, *B*, *C*, *D*, and *E* in a network have to be connected with the minimum length of cable. The distances between the nodes are given below. Find the most efficient connection route.

	A	*B*	*C*	*D*	*E*
A		100	90	80	110
B	100		130		120
C	90	130		120	
D	80		120		130
E	110	120		130	

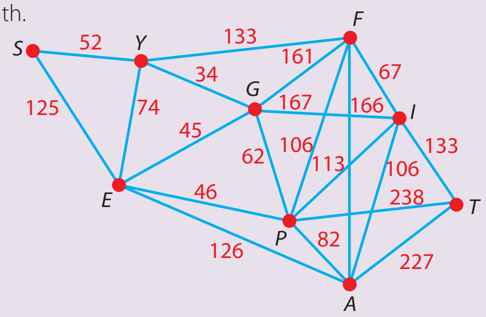
11 Use a shortest path algorithm to find the shortest route from *a* to *e*.



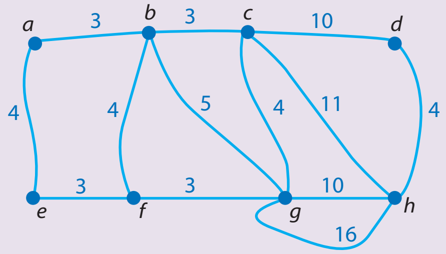
12 The graph below is the network of a transport company where the weights of the edges are distances in 10 km units. A shipment has to be transported from *a* to *i*. However, a part of the shipment has to be delivered to *f* first. Find the most efficient route for this delivery. Compare your result to the distance travelled when delivering the whole shipment directly from *a* to *i*.



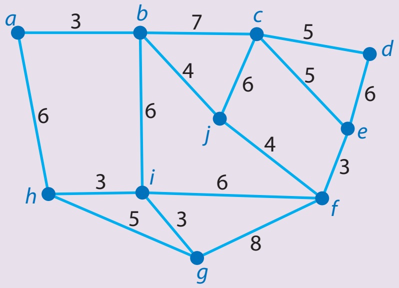
13 You are in charge of organizing the campaign tour for a politician. The following is a map showing the distances between the different cities that he must visit. He is based in *E* and needs to return there at the end of the tour. Find a suitable tour of minimum length.



14 A road sweeping truck has to sweep all the streets in a block of the city whose map is supplied. Distances are in 100s of metres. Find a route of minimum total length.



15 A local telephone network has to be inspected for possible defects. Find the shortest possible inspection tour to ensure that all cables have been checked. The sketch gives the length of each cable in 100s of metres.



Review questions 4

1 Show that if we delete an edge from a tree the remaining graph is not connected. The two unconnected components are subtrees to the original tree.

2 Show that if we add an edge between two non-adjacent vertices in a tree then the new graph contains only one cycle.

3 Show that a graph G contains a subgraph that is a tree if and only if it is connected and contains at least two vertices.

4 Let $T = (V, E)$ be a tree. Given that $|E| = 43$, find $|V|$.

5 **a** Let T be a tree with seven vertices. Find the number of all possible paths between the vertices in the tree (or subtrees).

b Find the formula for the number of all possible subtrees in a tree with n vertices.

6 Given that T is a tree, show that T contains at least two vertices of a degree 1.

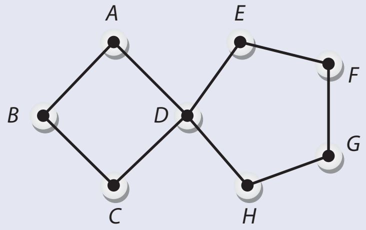
7 Given a complete graph with four vertices K_4, is it possible to find a spanning tree whose complement is also a spanning tree? Is it possible to find such a spanning tree in K_5?

8 Show that a complete bipartite graph $K_{m,n}$ contains a spanning tree with $m + n - 1$ edges.

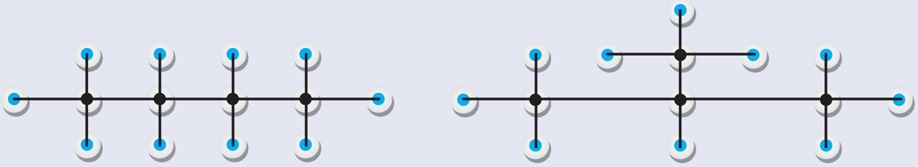
9 Given a complete bipartite graph $K_{2,2}$, is it possible to find a spanning tree whose complement is also a spanning tree? Is it possible to find such a spanning tree in $K_{2,3}$?

10 Draw all possible non-isomorphic trees with five vertices.

11 Find how many different spanning trees (some might be isomorphic) there are in the following graph.



12 The following graphs represent two molecules of chemical isomers of the saturated hydrocarbon C_4H_{10} (butane and isobutane). Each vertex that has a degree of 4 represents a carbon atom, C, whilst each vertex that has a degree of 1 represents a hydrogen atom, H. Explain why these two graphs are non-isomorphic.

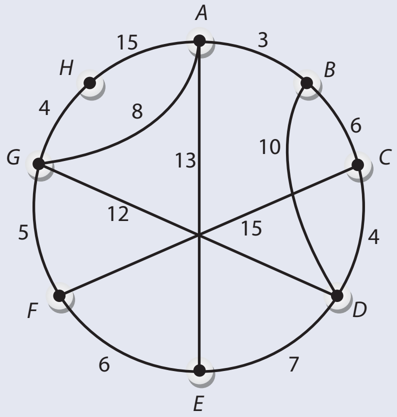


13 Given that a molecule (a tree) of a saturated hydrocarbon contains *n* carbon atoms (vertices of a degree 4), find how many hydrogen atoms (vertices of degree 1) there are.

14 Molecules of chemical isomers of the saturated hydrocarbon C_5H_{12} are called pentane, isopentane, and neopentane. Draw the trees representing those tree molecules, and give reasons why those three trees are mutually non-isomorphic.

15 Show that a complete binary tree with *n* internal vertices has *n* + 1 leaves.

16 Use Dijkstra's algorithm to find the shortest path between the vertices *B* and *F* in the following weighted graph.



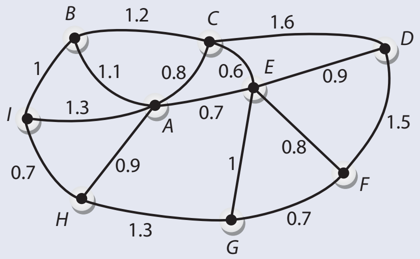
17 The bus routes connecting various cities and the cost of the tickets in dollars are given in the table below.

Cities	A	B	C	D	E	F
A	–	25	42	–	55	28
B		–	15	63	–	17
C			–	12	20	–
D				–	22	40
E					–	10
F						–

a Draw the weighted graph that represents all the routes between the cities.

b Jerry would like to travel from A to D. Determine the cheapest route and find how much will Jerry pay for his travel.

18 Ravi and his band have an upcoming concert in a club. He needs to display the concert posters in his neighbourhood. The following graph represents the plan of the posts where the posters can be displayed. Ravi's home is denoted by the vertex *A*. The distances between the posts are given in kilometres.



Find the shortest distance Ravi will need to travel in order to put the posters on all the posts before returning home.

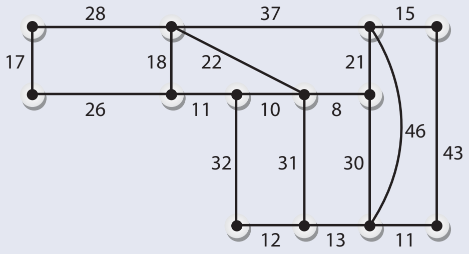
19 Jenny collects air miles and has earned 230 000 free miles through her air company. The cost of the plane tickets in free miles between the cities she visits is given in the matrix below. Each entry represents thousands of miles.

$$C_G = \begin{bmatrix} 0 & 0 & 10 & 0 & 20 & 25 & 10 \\ 0 & 0 & 10 & 18 & 54 & 0 & 0 \\ 10 & 10 & 0 & 8 & 0 & 50 & 0 \\ 0 & 18 & 8 & 0 & 0 & 0 & 45 \\ 20 & 54 & 0 & 0 & 0 & 28 & 32 \\ 25 & 0 & 50 & 0 & 28 & 0 & 16 \\ 10 & 0 & 0 & 45 & 32 & 16 & 0 \end{bmatrix}$$

a Draw a weighted graph representing the possible flights between the cities with the corresponding cost in free miles.

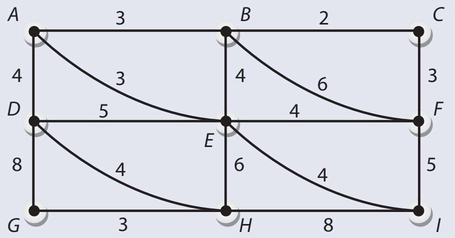
b Jenny would like to make a round trip and visit all the cities. What is the cheapest route and will she have enough miles for such a trip or she will need to buy some additional miles to pay for the trip?

20 Jack is a security guard. During the night shift he must patrol every single corridor of a warehouse. The plan of the corridors is given below. The time needed to patrol each corridor is given in minutes.



Is it possible for Jack to patrol the whole warehouse during his night shift from 10 p.m. till 6 a.m.? If yes, how many minutes will he have for a break? If not, how much longer would he need to stay in order to fulfil his duty?

21 Apply Kruskal's and Prim's algorithms to find the minimum spanning tree for the following graph.



Show all the decision steps in both algorithms. Draw the minimum spanning tree and state its weight.

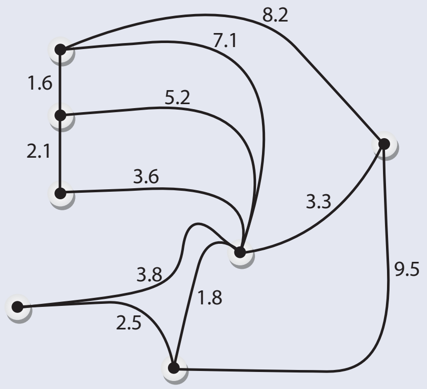
22 Adapt Kruskal's and Prim's algorithms to devise an algorithm to determine the maximum spanning tree in question 21.

23 Information on the distances between the cities in a country are provided in the table below. Each distance is given in kilometres.

Cities	P	Q	R	S	T	U
P	–	–	–	–	–	–
Q	200	–	–	–	–	–
R	292	487	–	–	–	–
S	333	465	222	–	–	–
T	86	282	203	257	–	–
U	333	509	133	97	235	–

The government would like to construct a system of highways to connect all the cities. Determine which highways should be built so that the cost of the construction is minimal. Assume that the cost of a kilometre of highway is constant.

24 Peter needs to install sockets that will be connected by an optical cable in his apartment so that he can watch TV and use the phone and internet in the rooms. The positions of the sockets are shown on the following graph. The distances between the sockets are given in metres.



Given that the cost of optical cable is 70 cents per metre, find the minimum price Peter will pay for buying the cable.

25 Christian plays a computer game in which he must enter rooms in order to collect some points. The points in the first level of the game are given in the following matrix.

$$C_G = \begin{bmatrix} 0 & 2 & 3 & 4 & 0 & 0 & 2 & 0 & 0 \\ 2 & 0 & 3 & 2 & 3 & 0 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 & 4 & 0 & 0 & 3 \\ 4 & 2 & 0 & 0 & 2 & 0 & 4 & 0 & 0 \\ 0 & 3 & 0 & 2 & 0 & 5 & 0 & 4 & 0 \\ 0 & 0 & 4 & 0 & 5 & 0 & 0 & 0 & 2 \\ 2 & 0 & 0 & 4 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 6 & 0 & 5 \\ 0 & 0 & 3 & 0 & 0 & 2 & 0 & 5 & 0 \end{bmatrix}$$

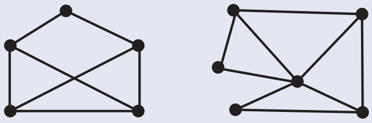
In order to advance to the higher level of the game, he must visit all the rooms in the shortest possible time. Find the maximum possible points Christian can collect at the first level.

Practice questions 4

- Practice questions 1–10 cover work from Chapters 3–4 inclusive.

1 a Prove that if two graphs are isomorphic they have the same degree sequence.

b Are the following graphs isomorphic? Justify your answer.



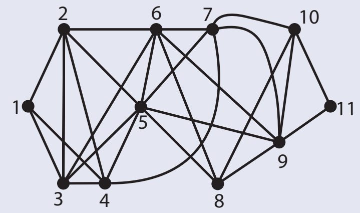
2 In an offshore drilling site for a large oil company, the distances between the planned wells are given below in metres.

	1	2	3	4	5	6	7	8	9	10
2	30									
3	40	60								
4	90	190	130							
5	80	200	10	160						
6	70	40	20	40	130					
7	60	120	50	90	30	60				
8	50	140	90	70	140	70	40			
9	40	170	140	60	50	90	50	70		
10	200	80	150	110	90	30	190	90	100	
11	150	30	200	120	190	120	60	190	150	200

a It is intended to construct a network of paths to connect the different wells in a way that minimizes the sum of the distances between them.

Use Prim's algorithm to find a network of paths of minimum total length that can span the whole site.

b Pipes are laid under water. Well 1 has the largest amount of oil to be pumped per day, and Well 11 is designed to be the main transportation hub. The only possible connections to be made between wells are shown in the diagram below.

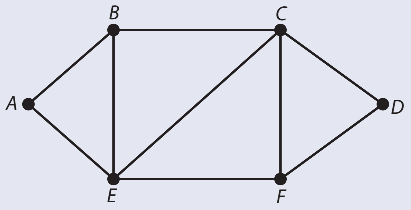


The associated cost for each pipe, in 100-thousand dollars, is given in the table below. Use Dijkstra's algorithm to find the path with minimum cost that can transport oil from Well 1 to Well 11.

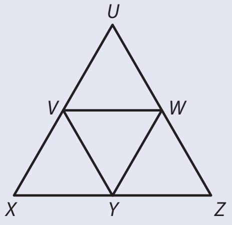
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2	6									
3	3									
4	8	7	2							
5		14	12	6						
6		16	19		7					
7				24	20	29				
8					23	15				
9					56	30	41	50		
10							42	25	40	
11									32	22

3 a Define the isomorphism of two graphs *G* and *H*.

b Determine whether the two graphs below are isomorphic. Give a reason for your answer.



Graph *G*



Graph *H*

c Find an Eulerian trail for graph *G* starting with vertex *B*.

d State a result which shows that graph *H* has an Eulerian circuit.

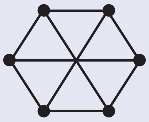
4 a Define the following terms.

i A bipartite graph.

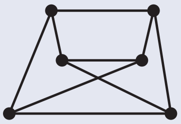
ii An isomorphism between two graphs, *M* and *N*.

b Prove that an isomorphism between two graphs maps a cycle of one graph into a cycle of the other graph.

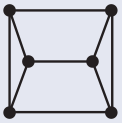
c The graphs *G*, *H*, and *J* are drawn below.



G



H



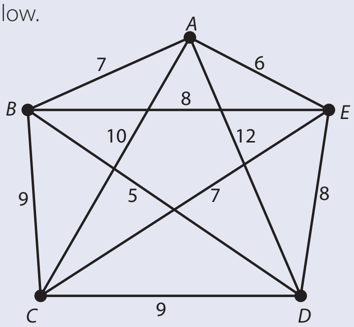
J

i Giving a reason, determine whether or not *G* is a bipartite graph.

ii Giving a reason, determine whether or not there exists an isomorphism between graphs *G* and *H*.

iii Using the result in part **b**, or otherwise, determine whether or not graph *H* is isomorphic to graph *J*.

5 Let *G* be the graph below.



a Find the total number of Hamiltonian cycles in *G* starting at vertex *A*. Explain your answer.

b **i** Find a minimum spanning tree for the subgraph obtained by deleting *A* from *G*.

ii Hence, find a lower bound for the travelling salesman problem for *G*.

c Give an upper bound for the travelling salesman problem for the graph above.

d Show that the lower bound you have obtained is not the best possible for the solution to the travelling salesman problem for *G*.

6 **a** Show that the sum of the degrees of all the vertices of a graph is even.

b There are nine men at a party. By considering an appropriate graph, show that it is impossible for each man to shake hands with exactly five other men.

c For a connected planar graph, prove Euler's relation, $v - e + f = 2$.

7 Consider the following adjacency matrices for the graphs G_1 and G_2.

$$\begin{array}{c|ccccc} & p & q & r & s & t \\ \hline p & 0 & 1 & 0 & 1 & 0 \\ q & 1 & 0 & 2 & 0 & 1 \\ r & 0 & 2 & 0 & 1 & 0 \\ s & 1 & 0 & 1 & 0 & 1 \\ t & 0 & 1 & 0 & 1 & 0 \end{array} \qquad \begin{array}{c|ccccc} & p & q & r & s & t \\ \hline p & 0 & 0 & 0 & 1 & 1 \\ q & 0 & 0 & 0 & 1 & 0 \\ r & 0 & 0 & 0 & 1 & 0 \\ s & 1 & 1 & 1 & 0 & 0 \\ t & 1 & 0 & 0 & 0 & 0 \end{array}$$

$$G_1 \qquad\qquad\qquad G_2$$

a Draw the graphs of G_1 and G_2.

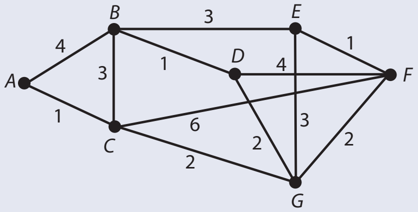
b For each graph, giving a reason, determine whether or not it

i is simple **ii** is connected

iii is bipartite **iv** is a tree

v has an Eulerian trail, giving an example of a trail if one exists.

8 Let H be the weighted graph drawn below.



a **i** Name the two vertices of odd degree.

ii State the shortest path between these two vertices.

iii Using the route inspection algorithm, or otherwise, find a walk, starting and ending at A, of minimum total weight which includes every edge at least once.

iv Calculate the weight of this walk.

b Write down a Hamiltonian cycle in H.

9 A graph G has e edges and n vertices.

a Show that the sum of the degrees of the vertices is twice the number of edges.

b Deduce that G has an even number of vertices of odd degree.

c **i** Graph G is connected, planar and divides the plane into exactly four regions. If $(n - 1)$ vertices have degree 3 and exactly one vertex has degree d, determine the possible values of (n, d).

ii For each possible (n, d), draw a graph which satisfies the conditions described in **i**.

10 **a** **i** Let M be the adjacency matrix of a bipartite graph. Show that the leading diagonal entries in M^{37} are all zero.

ii What does the (i, j)th element of $M + M^2 + M^3$ represent?

b Prove that a graph containing a triangle cannot be bipartite.

c Prove that the number of edges in a bipartite graph with n vertices is less than or equal to $\frac{n^2}{4}$.

Questions 1–10 © International Baccalaureate Organization

Answers

Chapter 1

Exercise 1.1–1.2

1 9 **2** 30 **3–6** Proof

7 a) Q = 30, R = 8
b) Q = −6, R = 70
c) Q = −5, R = 25

8–15 Proof

16 a) Proof b) $q = -8, r = 1$

17–18 Proof

19 $x = 4, y = 8$ **20** $x = 3, y = 9$

21 $\varnothing$ **22** Proof

23 True **24** True

25 True **26** True

27 False **28** False

29 True **30** a) 1 b)–d) proof

31 Proof **32** Proof

33 Proof **34** Proof

Exercise 1.3

1 4 **2** 1 **3** 17 **4** 68
5 77 **6** 1

7 $x = -17, y = 7$ **8** $x = -1, y = 1$

9 $x = -535, y = 132$ **10** $x = 9, y = 4$

11 $x = -1769, y = -29$ **12** $x = 5, y = 4$

13 No **14–16** Proof

17 8968 **18** 125 328

19 2100

20 (12, 360), (24, 180), (36, 120), (60, 72)

21 lcm(a, b) = ab **22** No **23–30** Proof

Exercise 1.4

1–5 Proof

6 For example, they end with 1 or 7.
7 a) $3 \cdot 29$ b) 19^2 c) $3^3 \cdot 5 \cdot 7$
d) $7 \cdot 11 \cdot 13$ e) $2^4 \cdot 19 \cdot 23$

8 a) gcd = 1, lcm = $3 \cdot 29 \cdot 19^2$
b) gcd = 1, lcm = $19^2 \cdot 7 \cdot 11 \cdot 13$
c) gcd = 1, lcm = $3 \cdot 29 \cdot 19^2 \cdot 7 \cdot 11 \cdot 13$
d) gcd = 1, lcm = $2^4 \cdot 3^3 \cdot 5 \cdot 7 \cdot 19 \cdot 23 \cdot 29$

9 6, 10, 15, 42, 70

10–12 Proof

13 $x \nmid y$, gcd = $3^2 \cdot 13$, lcm = $3^2 \cdot 5 \cdot 11^2 \cdot 13$

14 $x \nmid y$, gcd = $2^2 \cdot 23$, lcm = $2^3 \cdot 5^3 \cdot 23^2$

15 $x|y$, gcd = $3^2 \cdot 11 \cdot 23$, lcm = $3^2 \cdot 7 \cdot 11 \cdot 23$

16 $x \nmid y$, gcd = 5, lcm = $2 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17$

17–22 Proof

23 1, 3

24 1, 2

25 When a is odd, always; when a is even, only when c is even.

Chapter 2

Exercise 2.1

1 a) True b) False c) False d) True

2 19 **3** Proof **4** 2

5 5 **6** 17 **7** 9

8 38 **9** 19 **10** 5

11 6 **12** 6 **13** 15

14 12 **15** 1 **16** 16

17 5 **18** 11 **19–33** Proof

34 1, 18 **35** −18, 5, 28

36 3, 7, 11, 21, 33, 77, 231

37–39 Proof **40** 11, 39, 21 **41–42** Proof

Exercise 2.2

1 a) No solution b) Solution c) No solution

2 a) $x = 7 - 7t, y = 10 - 13t$
b) $x = 1 + 35t, y = -6 - 221t$
c) $x = -141 + 349t, y = 120 - 297t$

3 a) $x = 8 - 11t, y = 1 - 5t$, with $t \in \{..., -2, -1, 0\}$
b) No positive solutions
c) (1, 66), (12, 4)

4

Apples	16	34	52
Oranges	71	46	21

5 7 of the €4.98 posters and 11 of the €5.98 posters.

6 $10d + 25q = 455$; minimum = 20, maximum = 44

7

Chicken	3	10	17
Geese	9	5	1

8 (Calves, lambs, piglets): (5, 41, 54), or (10, 22, 68), or (15, 3, 82)

9 €3.96

10 23

11 Minimum number of sheep required = 16. Transaction is not possible.

12 $(1 + 2t, -1 - 3t)$ **13** $(1 - 2t, 1 - 3t)$

14 $(6 + 14t, -7 - 17t)$

15 $(1 - 4t, 2 - 11t)$ or $(1 + 4t, 2 + 11t)$

16 None

17 None

18 $(345 + 503t, -275 - 401t)$

19 $(6 + 7t, -11 - 13t)$

20 $(4 + 5t, -7 - 9t)$

21 $(5 + 11t, -3 - 7t)$

22 $(13 + 19t, -6 - 9t)$

23 $(1 + 3t, 16 - 2t), 0 \leqslant t < 8$

24 $(4 + 4t, 12 - 3t), 0 \leqslant t < 4$

25 $(3 + 3t, 8 - 2t), 0 \leqslant t < 4$

26 None

27 None

28 $(2 + 5t, 9999 - 3t)$

29 None

30 None

31 $(1 + 7t, 9 + 2t)$

32 $(3 + 17t, 2 - 22t)$

33 $(20 + 40t, -6 - 11t)$

34 $(21, 19)$ or $(72, 8)$

35–36 Proof

Exercise 2.3

1 $2 + 7k$

2 $2 + 3k$

3 $33 + 40k$

4 $41 + 49k$

5 $111 + 888k$

6 $75 + 80k$

7 $5 + 7k$

8 $2 + 3k$

9 $16 + 24k$

10 No solution

11 $812 + 1001k$

12 $10 + 45k$

13 No solution

14 $k \in (0, 4, 8, 12, \ldots, 32\}; 4$

15 11 (mod 12)

16 151 (mod 414)

17 34 (mod 35)

18 13 (mod 55)

19 6 (mod 210)

20 559 (mod 1430)

21 (2 (mod 5), 2 (mod 5))

22 No solution

23 (k (mod 5), $2 + k$ (mod 5))

24 (k (mod 7), $4 + 4k$ (mod 7))

Exercise 2.4

1 $(5600)_7$

2 $(1071)_{10}$

3 $(1562773)_8$

4 $(235056)_{10}$

5 $(5018)_{10}$

6 $(11111011010)_2$

7 $(77F394FB)_{16}$

8 $(33047851104)_{10}$

9 $(479)_{16}$

10 $(74E)_{16}$

11 $(1111111011001101110)_2$

12 $(11111011110111110101100111011011 0001001)_2$

13 a) When n is even.
b) When either a is a multiple of 3 or n is a multiple of 3.
c) When a is even.

Exercise 2.5

1 9

2 3

3 5

4 10

5 10

6 3 (mod 17)

7 9 (mod 17)

8 9 (mod 17)

9 5 (mod 11)

10 9 (mod 13)

11 1

12 a) 8 (mod 11), 11 (mod 13), 10 (mod 17)

b) 1064 (mod 2431)

13 1

14 10

15 8

16–20 Proof

Exercise 2.6–2.8

1 $b_1 = 6, b_2 = 15, b_3 = \frac{75}{2}, b_4 = \frac{375}{4}, b_5 = \frac{1875}{8}$; linear homogeneous of degree 1.

2 $a_1 = -2, a_2 = 4, a_3 = -8, a_4 = 16, a_5 = -32$; linear homogeneous of degree 2.

3 $a_1 = 5, a_2 = 10, a_3 = 40, a_4 = 320, a_5 = 5120$; not linear.

4 $b_1 = 1, b_2 = 8, b_3 = 43, b_4 = 218, b_5 = 1093$, not homogeneous.

5 $b_n = \frac{5}{2} b_{n-1}; b_1 = 4 \Rightarrow b_n = 4\left(\frac{5}{2}\right)^{n-1}$

6 $a_n = 5a_{n-1} + 3; a_1 = 3 \Rightarrow a_n = 3\left(5^{n-1}\right) + \frac{3}{4}\left(5^{n-1} - 1\right) = \frac{3}{4}(5^n - 1)$

7 $a_n = a_{n-1} + n; a_1 = 4 \Rightarrow a_n = 3 + \frac{n(n+1)}{2}$

8 $b_n = -\frac{11}{10} b_{n-1}; b_1 = 10 \Rightarrow b_n = 10\left(-\frac{11}{10}\right)^{n-1}$

9 $a_n = a_{n-1} - 2; a_1 = 0 \Rightarrow a_n = 2 - 2n$

10 $b_n = nb_{n-1}; b_1 = 8 \Rightarrow b_n = 8n!$

11 $\begin{cases} b_n = 4b_{n-1} + 5b_{n-2}; b_1 = 6, b_2 = 6 \Rightarrow r^2 - 4r - 5 = 0 \Rightarrow r \in \{-1, 5\} \\ \quad \Rightarrow b_n = \frac{2}{5} \cdot 5^n - 4(-1)^n \end{cases}$

12 $\begin{cases} a_n = -3a_{n-1} - 2a_{n-2}; a_1 = -2, a_2 = 4 \Rightarrow r^2 + 3r + 2 = 0 \\ \quad \Rightarrow r \in \{-1, -2\} \Rightarrow a_n = (-2)^n \end{cases}$

13 $\begin{cases} a_n = 2a_{n-1} - 2a_{n-2}; a_1 = 1, a_2 = 4 \Rightarrow r^2 - 2r + 2 = 0 \\ \quad \Rightarrow r \in \{1 - i, 1 + i\} \Rightarrow a_n = \left(\sqrt{2}\right)^n \left(-\cos\left(n\frac{\pi}{4}\right) + 2\sin\left(n\frac{\pi}{4}\right)\right) \\ \quad \Rightarrow \text{or } a_n = \frac{-1-2i}{2}(1+i)^n + \frac{-1+2i}{2}(1-i)^n \end{cases}$

14 $u_n = au_{n-1} + b \Rightarrow u_n = a^{n-1} \cdot u_1 + b \cdot \frac{a^{n-1} - 1}{a - 1}, a \neq 1.$

Practice questions 2

1 Proof

2 a) $x = 11, y = -6$ b) Proof

3 32

4 a) 235 b) 105 441 c) 9025

5 a) $\{1, 2, 3, 6\}$ b) 6
c) $6k - 4$ or $6k - 2, k \in \mathbb{Z}^+$

6 a) 1
b) (i) $x = 119 - 73k, y = -70 + 43k$ (ii) $(-27, 16)$

7 a) Proof
b) $x = 11 + 378n, y = -8 - 275n$, where $n \in \mathbb{Z}$

8 Definition and proof

9 a) (i) Proof (ii) (0, 5), (2, 3), (4, 1) (mod 6)
b) Proof

10 a) Proof b) $x \equiv 18 \pmod{35}$

Chapter 3

Exercise 3.1 and 3.2

1 a) (i) 4 (ii) 9 (iii) {5, 6, 5, 6}
b) (i) 4 (ii) 6 (iii) {3, 3, 3, 3}
c) (i) 5 (ii) 5 (iii) {2, 1, 3, 2, 2}

2 a) No b) Yes, K_5

3 $n - 1$

4 $\dfrac{n(n-1)}{2}$

5 a) $v = 7, e = 12$ b) $v = 30, e = 221$
c) $v = m + n, e = mn$

6 8, 16

7 a) 8 b) Yes; $r = 2, |v| = 14$, or $r = 4, |v| = 7$
c) $\left\lfloor \dfrac{p}{2} \right\rfloor$ d) Proof

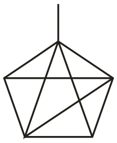
8–9 Proof

10 a, c

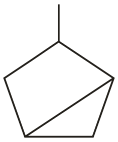
11 12

12 a) No, $|E|$ is not even.

b) Yes,



c) Yes,

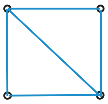


Exercise 3.3

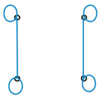
1 a) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 0 & 3 \\ 1 & 2 & 3 & 0 \end{pmatrix}$ b) $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$

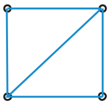
2 a)



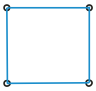
b)



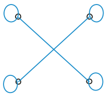
c)



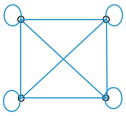
d)



e)



f)

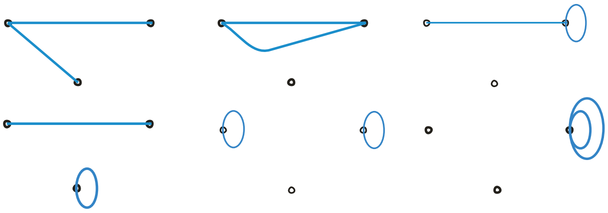


Graphs a) and c), and b) and e), are isomorphic.

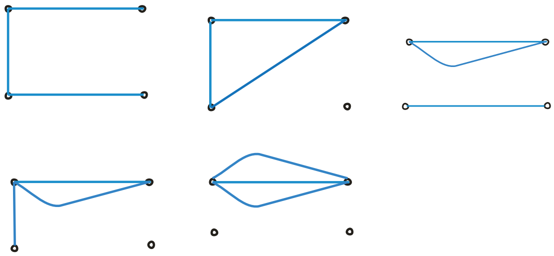
3 Isomorphic. Label the nodes, in both graphs, clockwise *a*, *b*, *c*, *d*, *e*, *f*, *g*. The correspondence *a*↔*g*, *b*↔*f*, *c*↔*e*, *d*↔*d*, *e*↔*c*, *f*↔*b*, *g*↔*a* is a homomorphism because when you rearrange the vertices in the second graph, you will have the same adjacency matrix as the first one.

4 a) No b) No c) No d) Yes

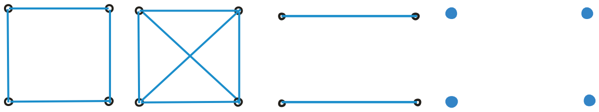
5 2 without loops, 6 with loops



6 5 without loops, 15 with loops



7



Exercise 3.4

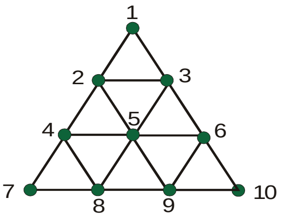
1 Vertices have even degrees.
a) 123174263456751 b) 1234543251

2 a) 1234214241 b) 12345241
c) Vertices 2 and 5 have degree 5 each.

3 a) When n is odd. b) When m and n are both even.

4 Graph 1(a) Hamiltonian: 12345671; graph 1(b) Hamiltonian: 123451.
Graph 2(a) Hamiltonian: 12341; graph 2(b) Hamiltonian path: 12345; graph 2(c) neither.

5 a) (10, 9, 6, 5, 9, 8, 5, 4, 8, 7, 4, 2, 5, 3, 2, 1, 3, 6, 10)
b) (10, 9, 8, 7, 4, 5, 2, 1, 3, 6, 10)
c) An Eulerian circuit is always possible ($n \geqslant 3$), because the degree of every vertex is even. A Hamiltonian cycle is also possible using the same plan as above: visit all vertices except one side, and then go back along that side.



6 Length 1 = 0; length 2 = 2; length 3 = 3, and length 4 = 10.

7 a) 51 between vertices not on the main diagonal, 52 for vertices on the diagonal
b) 205 between vertices not on the main diagonal, 204 for vertices on the diagonal
c) 819 between vertices not on the main diagonal, 820 for vertices on the diagonal

8 a) 48 among vertices of the 3-part, and 36 among the 4-part
b) 144 from vertices of 3-part to vertices of 4-part
c) 576 among vertices of the 3-part, and 432 among the 4-part
d) 1728 from vertices of 3-part to vertices of 4-part

9 a) No cycle. If you start at the left, you will need to visit *c* and *d* twice. Path: *abcdef*.

b) Cycle: *abcdea*.

c) No cycle since *f* has degree 1. Path: *eabcdf*.

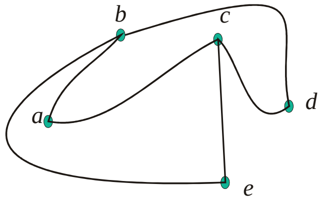
d) Neither cycle nor path as three vertices have degree 1.

e) No cycle, because in any of them *a* or *d* would have to be visited twice. Path: *eacdb*.

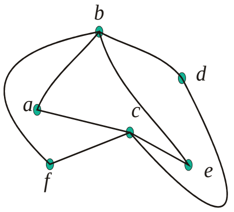
f) Cycle: *ahgfedcbia*.

Exercise 3.5

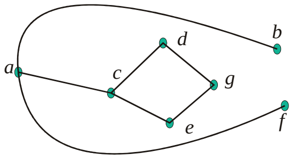
1 Planar. Redraw:



2 Planar. Redraw:



3 Planar. Redraw:



4 Not planar. *bf* and *ce* must cross, so must *ae* and *bd*.

5 15

6 15, 18

7 7, 9

8 6

9 Not planar

10 Planar

Practice questions 3

1 No, because there will be an edge connecting two vertices in the same component.

2 a) (i) $\binom{n}{2}$ (ii) $\binom{n}{3}$ (iii) $\binom{n}{m}$

b) $\frac{n+2}{2}$ or $\frac{n+1}{2}$

3 10

4 a) 2 b) 7

5 a) 0 b) 27

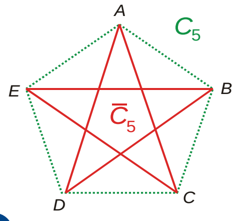
6 a) Proof b) Only C_3 is isomorphic to K_3 and W_3 to K_4.

c) Proof

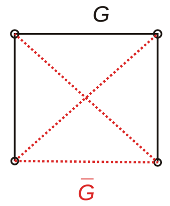
7 Proof

8 They contain odd cycles (size 3).

9 Yes; $A \leftrightarrow A$, $B \leftrightarrow C$, $C \leftrightarrow E$, $D \leftrightarrow B$, $E \leftrightarrow D$.

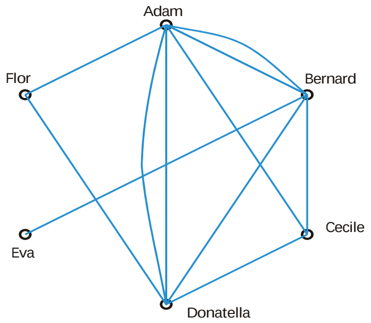


10 a) Yes:



b) No

11 a)



b) Yes, through Adam.

c) Bernard, as without him Eva is isolated.

Chapter 4

Exercise 4.1–4.3

1 a) 5, 7, 10, 11, 13, 14, 16, 17

b) 3, 1, 9

c) 3: 12, 13, 14; 7: no descendants; 15: 16, 17

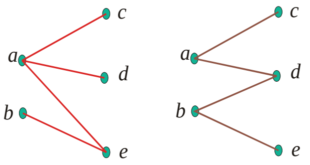
d) 4: 12; 7: no siblings; 9: no siblings

2 $|u| = 18$, $|v| = 36$, $|f| = 35$

3 31

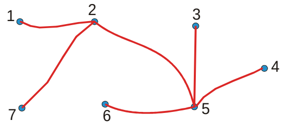
4 $\binom{n}{2}$

5 a) These are the only two non-isomorphic trees.

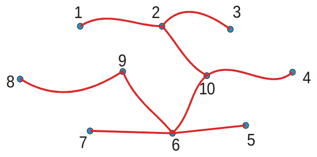


b) $\left\lfloor \frac{n+1}{2} \right\rfloor$

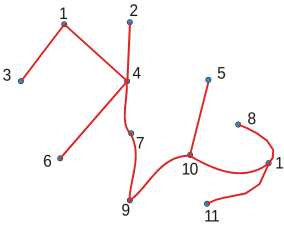
6



7



8



9 12, 23, 34, 45, 56, 67

10 12, 23, 34, 45, 56, 67, 78, 89, 9(10)

11 12, 24, 45, 58, 8(12), (12)(11), (11)9, 9(10), 47, 76, 63

12 13, 34, 45, 58, 89, 46, 67, 7(10), 12

13 17, 78, 89, 9(10), (10)(11), (11)6, 65, 54, (10)(14), 9(13), 83, 32

14 12, 23, 34, 46, 65, 5(10), (10)9, 98, 87, (10)(11), (11)(12), (12)(13), (13)(14), (14)(15), (10)(16), (16)(17), (17)(19), (19)(20), (20)(18)

15 a) 13, 12, 34, 45, 46, 67, 78, 7(10), 89

b) 12, 23, 34, 45, 56, 67, 78, 89, 7(10)

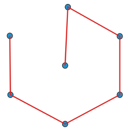
16 a) 12, 17, 7(12), 78, 83, 8(13), 89, 94, 95, 9(10), 9(14), (10)(11), (11)6

b) 12, 23, 38, 89, 94, 45, 56, 6(11), (11)(10), (10)(14), 9(13), 87, 7(12)

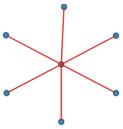
17 a) 12, 15, 23, 26, 34, 5(10), (10)7, (10)8, (10)9, (10)(11), (10)(16), (11)(12), (11)(13), (11)(14), (11)(15), (16)(17), (16)(18), (16)(20), (20)(19)

b) 12, 23, 34, 46, 65, 5(10), (10)7, 78, 89, (10)(11), (11)(12), (12)(13), (13)(14), (14)(15), (10)(16), (16)(17), (17)(18), (18)(20), (20)(19)

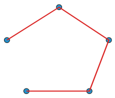
18 a) (i)



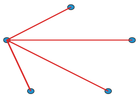
b) (i)



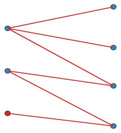
a) (ii)



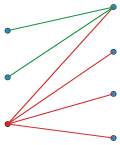
b) (ii)



a) (iii)

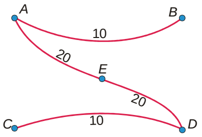


b) (iii)

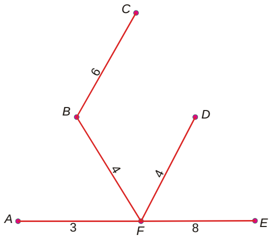


Exercise 4.4

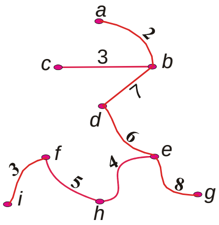
1



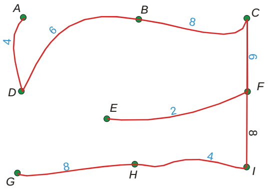
2



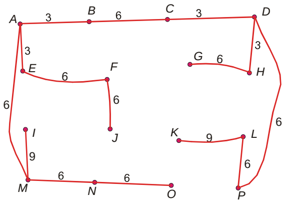
3



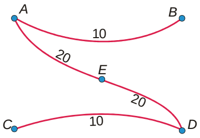
4



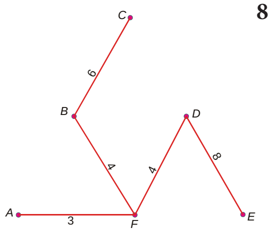
5



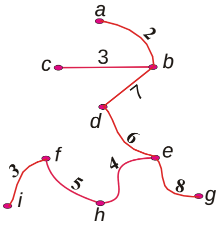
6



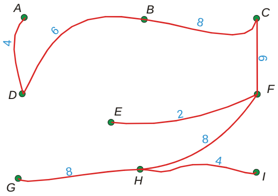
7



8



9

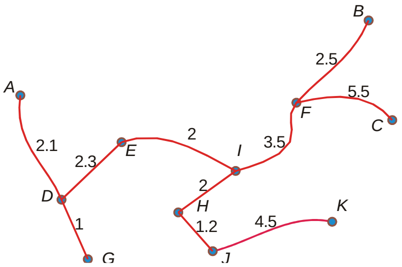


10 A few shapes are possible, one of which is similar to the answer to question 5.

11 1 and 6 have the same final tree. However, when building the tree using Kruskal's algorithm, *AB* and *CD* were added first. When using Prim's algorithm, *AB* was followed by *AE*, *ED*, and then *CD*.
With 2 and 7, there is no apparent difference. The different shapes are due to random choices.
3 and 8 have the same final tree too. Using Kruskal's algorithm, the order of addition to the tree is: *ab*, *bc*, *fi*, *he*, *fh*, *ed*, *bd*, and *eg*. Using Prim's algorithm, the order is: *ab*, *bc*, *bd*, *ed*, *he*, *fh*, *fi* and *eg*.
4 and 9 may have the same tree too. However, using Kruskal's algorithm, the order of edge addition is: *ef*, *ad*, *hi*, *cf*, *db*, *bc*, *fi*, and *gh*. Using Prim's algorithm, the order is: *ef*, *fc*, *fh*, *ih*, *cb*, *bd*, *da*, and *gh*.

5 and 10 may have the same tree too. However, using Kruskal's algorithm, the order of edge addition is: *AB*, *AE*, *CD*, *DH*, *BC*, …. Using Prim's algorithm, the order is: *AB*, *AE*, *BC*, *CD*, *DH*, ….

12



Exercise 4.5

1 70, *abedf* **2** 48, *ACDEGH*

3 32, *acfimpsu* **4** *abed*

5 *A–F*: *ACDF*; *B–H*: *BCDEGH*

6 *ADBCA*, 85

7 *EDCABE* or *DEBACD*, 400

8 Vienna–Frankfurt–Prague–Moscow–Milan–Vienna: €1070.

9 New York–Paris–London–Madrid–Boston–New York: €1215.

10 *DACBED*, 550

11 *age*, 19

12 *abdfhi*, 21; *acehi*, 13

13 Without visiting any city twice: *ESYFITAPGE*, 926. Visiting *Y* twice: *EGYSYFITAPE*, 871.

14 *abcdhghcgbfgfea*, 8300

15 *abcdecjfefibjfgihgha*, 9200

Review questions 4

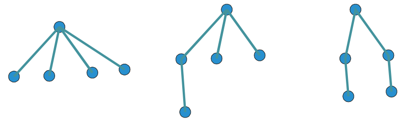
1–3 Proof **4** 44

5 a) 21 b) $\binom{n}{2}$

6 Proof **7** Yes; no

8 Proof **9** No; no

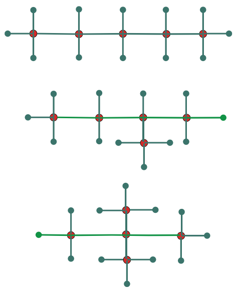
10



11 20

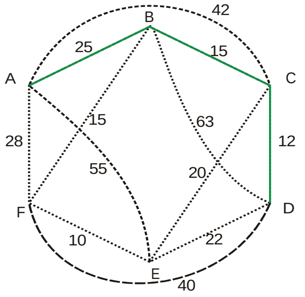
12 On the left there are 2 carbon atoms adjacent to 3 hydrogen atoms each, while on the right 3 carbon atoms have this property.

13 $2n + 2$ **14**



15 Proof **16** *BAGF*, 16

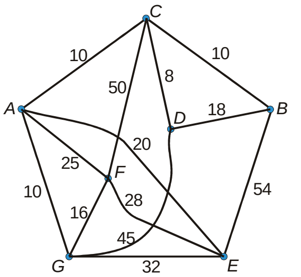
17 a)



b) ABCD, $52

18 *ACEDFGHIBA*, 8.6 km

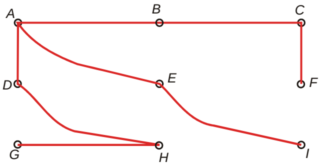
19 a)



b) Sample: *ACBDCAEFGA* with 130 000 free miles, which she can afford.

20 Yes; he will have a 20-minute break.

21 Sample for Kruskal's algorithm: *BC*, *AB*, *AE*, *CF*, *GH*, *AD*, *DH*, *EI*. Sample for Prim's algorithm: *BC*, *AB*, *AE*, *CF*, *AD*, *DH*, *GH*, *EI*. Weight = 26.



22 Sample for Kruskal's algorithm: *DG*, *HI*, *BF*, *EH*, *DE*, *FI*, *AD*, *FC*. Sample for Prim's algorithm: *DG*, *DE*, *EH*, *HI*, *IF*, *BF*, *AD*, *CF*. Weight = 45.

23 PT, SU, RU, PQ, TR, total distance of 719 km

24 1043 cents (10.43 dollars)

25 35

Practice questions 4

1 a) Proof

b) Not isomorphic; one has a vertex of degree 4, the other does not.

2 a)

Vertices added to the tree	Edge added	Weight
3	Ø	0
5	3, 5	10
6	3, 6	20
7	5, 7	30
10	6, 10	30
1	3, 1	40
2	1, 2	30
11	2, 11	30
9	1, 9	40
4	6, 4	40
8	7, 8	40
		310

b) Any of two paths: 1–3–4–5–6–8–10–11 or 1–3–4–5–6–9–11, with weight 80.

3 a) Student definition

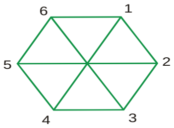
b) Not isomorphic; G has a vertex of degree 3, while H has not.

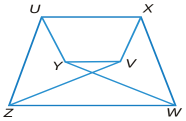
c) $BAEBCEFCDF$

d) All vertices have even degree.

4 a) Student definition b) Proof

c) (i) G is bipartite since if we label the vertices clockwise as 1, 2, 3, …, the two components will be {1, 3, 5} and {2, 4, 6}.





(ii) G and H are isomorphic: $1\leftrightarrow U$, $2\leftrightarrow X$, $3\leftrightarrow V$, $4\leftrightarrow Y$, $5\leftrightarrow W$, $6\leftrightarrow Z$.

(iii) No; H is bipartite, J is not.

5 a) 24

b) (i) $BDEC$

(ii) 33

c) $DBAEC$ is a minimum spanning tree of weight 26. Upper bound = 26 × 2 = 52.

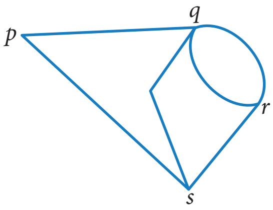
d) A minimum tour is 34; 33 cannot be achieved.

6 a) Every edge creates 2 degrees, with n edges there are $2n$ degrees.

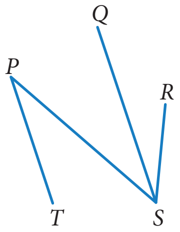
b) Each vertex will have a degree of 5, 45 in total, which is not even. Hence, it is not possible.

c) See Chapter 3, page 121.

7 a)



G_1



G_2

b) (i) G_1 is not simple, G_2 is simple.

(ii) Both are connected.

(iii) Both are bipartite. G_1: components are $\{p, r, t\}$ and $\{q, s\}$. G_2: components are $\{P, R, Q\}$ and $\{T, S\}$.

(iv) G_1 is not a tree, as it has a cycle. G_2 is a tree.

(v) G_1 contains an Eulerian trail: $rqpsrqts$. G_2 does not have an Eulerian trail since four vertices have odd degrees.

8 a) (i) D, E

(ii) EBD

(iii) Example: $ABEFGCBDBEGDFCA$

(iv) 36

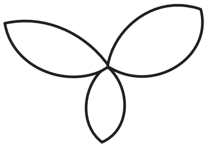
b) Example: $ABEFDGCA$

9 a) Every edge creates 2 degrees, with e edges there are $2e$ degrees.

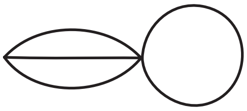
b) Student deduction

c) (i) (n, d) = (1, 6), (2, 5), (3, 4), (5, 2) or (6, 1)

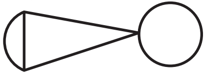
(ii) (1, 6)



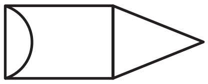
(2, 5)



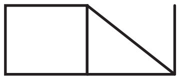
(3, 4)



(5, 2)



(6, 1)



10 a) (i) Proof

(ii) Number of paths from v_i to v_j with a maximum length of 3.

b)–c) Proof