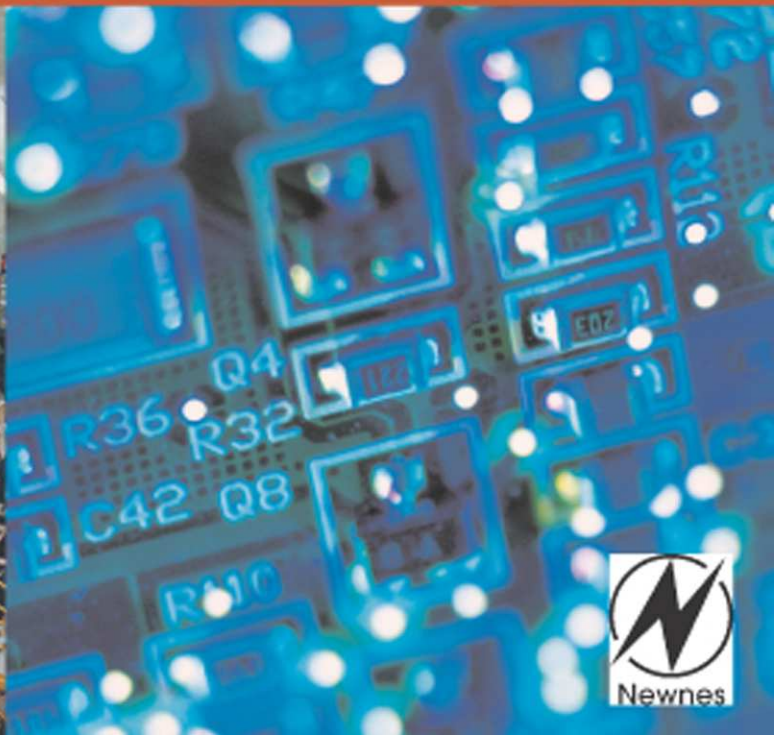




*Second edition*

# HIGHER NATIONAL ENGINEERING

*Mike Tooley & Lloyd Dingle*



# **Higher National Engineering**

This Page Intentionally Left Blank

# Higher National Engineering

**Mike Tooley and Lloyd Dingle**



AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD  
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Newnes is an imprint of Elsevier



Newnes  
An imprint of Elsevier  
Linacre House, Jordan Hill, Oxford OX2 8DP  
30 Corporate Drive, Burlington, MA 01803

First published 1998  
Reprinted with amendments 1999  
Reprinted 2000, 2001, 2002  
Second Edition 2004

© Michael H. Tooley and Lloyd B. Dingle 1998, 2004

All rights reserved. No part of this publication may be reproduced in any material form (including photocopying or storing in any medium by electronic means and whether or not transiently or incidentally to some other use of this publication) without the written permission of the copyright holder except in accordance with the provisions of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London, England W1T 4LP. Applications for the copyright holder's written permission to reproduce any part of this publication should be addressed to the publishers

**British Library Cataloguing in Publication Data**

A catalogue record for this book is available from the British Library

ISBN 0 7506 6177 1

For more information on all Newnes press publications  
please visit our website at [www.newnespress.com](http://www.newnespress.com)

Working together to grow  
libraries in developing countries

[www.elsevier.com](http://www.elsevier.com) | [www.bookaid.org](http://www.bookaid.org) | [www.sabre.org](http://www.sabre.org)

ELSEVIER

BOOK AID  
International

Sabre Foundation

Typeset by Charon Tec Pvt. Ltd, Chennai, India

Printed and bound in Great Britain

# Contents

|          |                                                                 |            |
|----------|-----------------------------------------------------------------|------------|
|          | <b>Introduction</b>                                             | <b>vii</b> |
|          | <b>Acknowledgements</b>                                         | <b>ix</b>  |
| <b>1</b> | <b>Business Management Techniques</b>                           | <b>1</b>   |
| 1.1      | Managing work activities                                        | 1          |
| 1.2      | Costing systems and techniques                                  | 22         |
| 1.3      | Financial planning and control                                  | 36         |
| 1.4      | Project planning and scheduling                                 | 48         |
| <b>2</b> | <b>Engineering Design</b>                                       | <b>59</b>  |
| 2.1      | The design specification                                        | 60         |
| 2.2      | The design report                                               | 70         |
| 2.3      | Computer technology and the design process                      | 83         |
| <b>3</b> | <b>Engineering Science</b>                                      | <b>102</b> |
| 3.1      | Static engineering systems                                      | 102        |
| 3.2      | Dynamic engineering systems                                     | 140        |
| 3.3      | DC and AC theory                                                | 175        |
| 3.4      | Information and energy control systems                          | 249        |
| <b>4</b> | <b>Electrical and Electronic Principles</b>                     | <b>285</b> |
| 4.1      | Circuit theory                                                  | 285        |
| 4.2      | Networks                                                        | 331        |
| 4.3      | Complex waves                                                   | 342        |
| 4.4      | Transients in $R$ – $L$ – $C$ circuits                          | 355        |
| <b>5</b> | <b>Mechanical Principles</b>                                    | <b>363</b> |
| 5.1      | Complex loading systems                                         | 363        |
| 5.2      | Loaded beams and cylinders                                      | 373        |
| 5.3      | Pressure vessels                                                | 385        |
| 5.4      | Dynamics of rotating systems                                    | 397        |
| <b>6</b> | <b>Analytical Methods</b>                                       | <b>431</b> |
| 6.1      | Algebra                                                         | 432        |
| 6.2      | Trigonometry                                                    | 459        |
| 6.3      | Calculus                                                        | 477        |
| 6.4      | Statistics and probability                                      | 515        |
| 6.5      | Advanced topics                                                 | 585        |
|          | <b>Appendix: Unit mapping and additional reference material</b> | <b>625</b> |
|          | <b>Abbreviations</b>                                            | <b>631</b> |
|          | <b>Answers</b>                                                  | <b>635</b> |
|          | <b>Index</b>                                                    | <b>649</b> |

This Page Intentionally Left Blank

# Introduction

This book has been written to help you achieve the learning outcomes of the core units of BTEC's new Higher National Engineering programme and the two authors have many years experience of teaching HNC and HND engineering students. In producing the book, the authors' principal aim has been that of capturing, within a single volume, the core knowledge required of all engineering students at HNC/HND level.

In order to reflect recent changes in the BTEC syllabus this second edition has been greatly extended and expanded. In particular, new sections have been included on **Managing work activities**, **DC and AC theory**, **Statistics and probability**, and **Engineering and scientific applications**. In addition, a number of small changes and corrections have been made.

The six core units covered in the book are:

- Business management techniques,
- Engineering design,
- Engineering science,
- Electrical and electronic principles,
- Mechanical principles,
- Analytical methods.

The book has been organised on a logical basis with each chapter devoted to a single core unit. We have, however, attempted to reduce duplication and some material is appropriate to more than one of the core units. Furthermore, to put mathematical concepts into context, we have developed a number of mathematical topics within the individual chapters. This material appears as *Mathematics in Action*, and features throughout the text. In addition, these mathematical topics are summarised in the chapter dealing with Analytical methods. You will also find that, where difficult concepts are introduced, we have included notes in the margin headed *Another View*. These will provide you with an alternative way of understanding them.

This book has been designed to provide you with a thorough introduction to each of the core units. Despite this, you are advised to make use of other reference books and materials wherever and whenever possible. You should also get into the habit of using all of the resources that are available to you. These include your tutor, your college or university library, computer centre, engineering laboratories, and other learning resources, such as the Internet (see Appendix). You should also become familiar with selecting materials that are appropriate to the topics that you are studying. In particular, you may find it useful to refer to materials that will provide you with several different views of a particular topic.

Throughout this book we have provided worked examples that show how the ideas introduced in the text can be put into practice. We have also



included problems and questions at various stages in the text. Depending on the nature of the topic, these questions take a variety of forms, from simple problems requiring short numerical answers to those that may require some additional research or that may require the use of an analytical software package in their solution. Your tutor may well ask you to provide answers to these questions as coursework or homework but they can also be used to help you with revision for course assessments.

**Business Management Techniques** are introduced in Chapter 1. This chapter will provide you with an introduction to standard costing techniques as well as an insight into the key functions that underpin financial planning and control, project planning, and scheduling. Chapter 2, **Engineering Design**, deals with the thought processes and procedural activities concerned with the design of engineering artefacts and systems. At this level the mathematical rigour often associated with ‘designing’ has been omitted. Instead, we have emphasised the production of the design specification, management report, and the use of computer technology as a design aid.

The scientific principles that underpin the design and operation of modern engineering systems are introduced in Chapter 3, **Engineering Science**. This chapter provides essential preparation to the two ‘principle’ units that follow. It also provides a valuable introduction to engineering science for anyone who has not studied engineering before.

**Electrical and Electronic Principles** are covered in Chapter 4. This chapter introduces electrical circuit theory, networks, and complex waveforms. **Mechanical Principles** are introduced in Chapter 5, which explains the principles that underpin the design and operation of mechanical engineering systems and deals with complex loading systems, loaded cylinders and beams, power transmission, and rotational systems.

Chapter 6, **Analytical Methods**, covers the essential mathematical principles, methods and applications that underpin a study of engineering at HNC/HND level. Each topic is presented in three parts: *formulae, methods, and applications*. This novel method of presentation is particularly effective because it will allow you to develop analytical and mathematical skills *alongside* the relevant engineering concepts.

Finally, we would like to offer a few words of practical advice to students. At the beginning of your HNC or HND course you will undoubtedly find that some topics appear to be more difficult than others. Sometimes you may find the basic concepts difficult to grasp (perhaps you have not met them before), you may find the analytical methods daunting, or you might have difficulty with things that you cannot immediately visualise.

No matter what the cause of your temporary learning block, it is important to remember two things: you would not be the first person to encounter the problem, and there is plenty of material available to you that will help you overcome it. All that you need to do is to recognise that it is a problem and then set about doing something about it. A regular study pattern and a clearly defined set of learning goals will help you get started. In any event, do not give up – engineering is a challenging and demanding career and your first challenge along the road to becoming a practising engineer is to master the core knowledge that engineers use in their everyday work. And that is what you will find in this book.

May we wish you every success with your Higher National studies!

*Mike Tooley and Lloyd Dingle*

# Acknowledgements

The authors would like to thank a number of people who have helped in producing this book. In particular, we would like to thank Richard Tooley for taking our rough drawings and turning them into 'real' artwork; Gerry Wood for comments and advice on the chapter dealing with Business Management Techniques; Matthew Deans, and all members of the team at Newnes for their patience and perseverance. Last, but by no means least, we would like to say a big 'thank you' to Wendy and Yvonne. But for your support and understanding this book would never have been finished!

This Page Intentionally Left Blank

# 1

# Business management techniques

## Summary

This unit is designed to give students an appreciation of the application of standard costing techniques as well as an insight into the key functions underpinning financial planning and control. The unit also aims to expand students' knowledge and interest in managerial and supervisory techniques by introducing and applying the fundamental concepts of project planning and scheduling.

## 1.1 MANAGING WORK ACTIVITIES

In order to be effective, an engineering business needs to manage its work activities in order to meet its organisational objectives. Such objectives may not be simply that of returning a profit to shareholders but will often be passed on other considerations; for example, include establishing a market lead, developing expertise in the application of new materials, pioneering new manufacturing processes, or diversifying interests into other spheres of engineering. This section looks at the various business functions of an engineering company, how these are organised and how they interact with one another. The various functions that we shall look at include planning, control, leadership and direction, marketing, finance, and research and development (R&D).

### Planning

Planning is absolutely fundamental to the correct functioning of an engineering company. If no planning is done then activities are almost certainly going to be very ineffective. Planning is the sum of the following activities:

- setting the *goals* for an engineering company;
- forecasting the *environment* in which the engineering company will operate;
- determining the *means* to achieve goals.

Setting goals or objectives must be the first step. It determines the direction an engineering company is going. It encourages all engineering company members to work towards the same ends, otherwise members

## 2 Higher National Engineering

are likely to set their own objectives which will conflict with each other. Good objectives make for rational engineering companies that are more co-ordinated and effective. The objectives must therefore be set by the most senior management group in an engineering company so that all of its staff can be given clear direction. If the goals are clearly stated, logical and appropriate for the business, then they act both as motivators and yardsticks for measuring success.

Once the engineering company has clear direction the next step is to analyse the environment and forecast its effect on the business. For example, a car manufacturer may set objectives so that, within 5 years, it will:

- achieve a 15% share of the European market;
- it will achieve a 5% share of the global market;
- be acknowledged as the market leader;
- be the accepted technological leader;
- be highly competitive on price;
- operate internationally.

When it forecasts its environment it might conclude that:

- new designs will be marketed by competitors;
- new lean-burn engines will become available;
- new battery technology will become available to support dual powered vehicles;
- there will be a sharp decline in the demand for luxury vehicles.

Its conventional petrol engines are technically very sound but are being threatened by the use of new highly efficient lean-burn engines introduced by competitors which are proving attractive to customers because they reduce the overall costs of ownership and running a vehicle. It has to decide how to deal with the threat, either to improve the efficiency of its existing engine designs or to follow the new trend and produce products based on new engine technology including the use of dual power.

Forecasting the environment allows the company to set new objectives and to prepare plans to meet its revised goals. Companies that fail to go through this process will go into decline in the long run because they are ignoring the changing world around them.

Once the goals have been refined and changed in the light of environmental forecasting then plans can be made to achieve the goals. Some plans will not change that much, whereas others will be dramatically affected by the changing environment. For this reason plans can be classified as follows:

- standing plans,
- single use plans,
- strategic plans,
- tactical plans.

*Standing plans* are those that are used many times, and remain relatively unaffected by environmental change. Examples are employment, financial, operating, and marketing policies and procedures. For example, hiring new employees involve standard procedures for recruitment and selection. Another example would be the annual routines for establishing budgets.

*Single use plans* are those that are used once only, such as those for the control of a unique project or specific budgets within an annual budget.

(Budgets themselves are single use plans, even though the procedures used for producing them are standing plans.)

*Strategic plans* are the broad plans related to the whole engineering company and include forecasting future trends and overall plans for the development of the engineering company. They are often in outline only and very often highly subjective, involving judgements made by top managers. For example, a plan may be made to build an entirely new factory based on forecasts of demand. This plan is strategic, and if it is wrong and the sales forecasts on which it is based do not materialise, the results for a company could be devastating.

*Tactical plans* operate within the strategic plan. The new lean-burn engine factory has to be brought into commission, and production has to be scheduled and controlled. Plans for the latter are tactical, since they focus on how to implement the strategic plan.

## Control

---

The prerequisite of control is planning. Controlling involves comparing events with plans, correcting deviations, and ensuring that the planned events happen. Sometimes deviations are so fundamental that they require a revision to the plan so that later events are controlled against a new plan. For example, the original sales forecast may turn out to be too optimistic, and production plans may have to be reduced to bring output into line with what sales are possible.

There are various ways in which control can be exercised. It can be predictive as in the case of a cash-flow forecast. This forecast may indicate a shortfall of cash in August but a surplus in September. The finance manager may need to arrange additional finance with the bank in August and then in September he might deposit surplus funds onto the money market. The point here is that variances are predicted in advance, thereby promoting cash control.

In the case of monthly comparisons between budgeted expenditures and actual expenditures an *over-spend* might be revealed. This triggers action that holds back expenditure until spending comes back into line with budget. This is historical control since action is based on a report of events in the recent past.

Concurrent control is *real time* such as that which might occur in controlling a continuous process or a production line. In this case the system has built-in *feedback* which enables it to remain in balance by regulating inputs and outputs. An example of concurrent control would be where a production process requires temperature regulation. The control system is designed to switch off heating when temperature reaches a threshold or switch on heating when it drops to a minimum level. The feedback is information on temperature. The same principle applies in some stock control systems, where stocks are maintained at pre-determined minimum and maximum levels, with supplies being switched on and off to maintain equilibrium.

## Leadership and direction

---

Planning and control activities are the tasks of *management*. However, they are only achieved through *people*. People will work effectively if they are led and directed effectively. This implies that top managers must

## 4 Higher National Engineering

be in touch with the business and be visible. They must have a clear vision for the future, reinforced by specific objectives which are communicated to their employees.

This approach to leadership is apparent in some of our best companies as exemplified by Marks & Spencer. Such companies have a clear *mission and objectives*, and have a visible committed top management. This philosophy permeates the whole engineering company stimulating better performance from all employees.

Motivating good performance from all employees is the responsibility of all managers. What motivates individuals and groups within commercial engineering companies is a complex and important subject, the detail of which is well beyond the scope of this book. However, it is still worth saying that managers must discover what it is that will stimulate employees to work productively.

In general people respond best to *considerate styles* of management, whereby their personal contributions are fully recognised. It is also true that there has to be an atmosphere of discipline coupled with a *work-oriented culture*. The task has to be accomplished, and being considerate does not extend to the toleration of slack or sloppy practices and behaviour.

Clear direction, sound and explicit guidelines, and well worked out procedures all of which are well communicated, ensure that the engineering company works smoothly.

### Allocation and supervision of work

---

This is the practical implementation of all that we have discussed in this section. An engineering company exists to fulfil the goals of its owners. It has to function in a co-ordinated and rational way. The people who are its members have to work together and understand their specific roles and functions. They need to receive directions and work has to be allocated. There has to be supervision of these activities. An engineering company is analogous to a machine or a living organism. In order to function properly everything has to work together smoothly. The managers have the task of ensuring that this work takes place according to plan and within the engineering company's stated objectives.

### Finance and accounting

---

The *finance director* will be responsible for managing a company's cash flow, short- and long-term investments, and supervise the company's accounting and budgetary system. His or her job is to ensure that the company has sufficient cash to support day-to-day operations. He or she should also ensure that cash that is not immediately required is made to work for the company. This is done by arranging short-term investments on the money markets or by switching funds into building society or bank deposit accounts. He or she may also be involved with international money flows and the 'hedging' of risk against exchange rate losses.

The *financial accountant* will be in charge of all recording in the financial accounting system. All business transactions are recorded in a system called *double entry accounts*. These records are called so because every business transaction has a twofold effect, and to make a complete record this twofold aspect has to be recorded. For example, if £100 was spent to buy raw materials, there has to be a record in the bank account of the money spent and then the existence of materials bought must be

recorded in a purchases account. This method of recording enables the business to produce a profit and loss account, and a balance sheet, which summarise all the transactions so that the financial performance can be tracked, and reported to shareholders and other interested parties, including the tax authorities.

The *management accountant* will administer the budgetary control and costing system. This system enables the business to forward plan its profit and loss, and its overall financial position, as well as being able to control and report on costs of operation. Thus a budgeted profit and loss account, and balance sheet is produced, called a *master budget*. This master budget summarises the budgets for all cost centres or operating divisions. The management accountant monitors actual results against the budget and will use data from the *double entry accounts* system referred to above. This monitoring enables the departmental managers to correct deviations to budget, control and manage the cost of running their departments.

Part of management accounting is the process of *investment appraisal* to plan the purchase of fixed assets and to ensure the best choices are made for new and replacement equipment. The management accountant will also prepare reports and analyses of various kinds for management.

The *cashier*, who is likely to report to the financial accountant will deal with all bank and cash transactions. In a large company the number of transactions is considerable, especially those that deal with receipts from customers and payments to suppliers. The work of this section is important because forecasting and monitoring cash flow is vital to the financial well-being of the business.

The *credit controller* will be concerned with authorising new credit customers and controlling the amount of credit granted, and reducing or preventing bad debts. A budget will be prepared giving planned levels for outstanding debtors. This figure really represents a short-term investment of capital in customers. This investment has to be managed within budgeted limits otherwise the company's finance costs would increase and there would be further risks of non-payment. The credit controller would monitor the financial stability of the existing customers or vet the standing of new customers. He might do this through a combination of bank references, credit agencies, and studying the customers' own published accounts.

Other activities carried out in accounting include payment of wages and salaries, depreciation of fixed assets, maintaining shareholder records, and paying shareholder dividends.

## Marketing

Marketing is said to be the most important function in the business, since if customers cannot be found for the company's products the company will go out of business, regardless of how financially well run or efficient it is.

Although sales is considered separately later, it is really part of the marketing function. Marketing is all about matching company products with customer *needs*. If customer needs are correctly identified and understood, then products can be made which will give the customer as much as possible of what he wants. Companies which view the customer as 'sovereign' are those companies that stay in business, because customers will continue to buy products that meet their requirements.

Hence marketing activities are centred around the process of filling customers' known needs and discovering the needs the customer does



## 6 Higher National Engineering

not yet know he has and exploiting this by finding out how to improve products so that customers will buy this company's products in preference to other goods. Some of the most important activities are:

- market research;
- monitoring trends in technology and customer tastes;
- tracking competitors' activities;
- promotion activities;
- preparing sales forecasts.

Remember that in some businesses the marketing activity is directed at end consumers, members of the public. This has to be done by national forms of advertising, such as television (TV) commercials, direct mail, newspapers or through major retailers selling to the consumer. The methods used may be somewhat different if the customers are other companies. Although the principle of meeting customer needs is the same, the approach taken may be much more technical and may include the services of sales engineers to provide technical backup and advice. The publicity methods are more likely to be centred around trade fairs, exhibitions, advertisements in the trade press or technical journals, for example. You should note these two distinct marketing approaches are respectively called consumer marketing and industrial marketing.

### Sales

---

The sales department is concerned with advertising and selling goods. It will have procedures for controlling sales and the documentation required. The documents used are the same as for purchasing, described below, except from the suppliers' viewpoint rather than from the customers'. The company may employ *commercial travellers*, *distributors*, or *agents*, or have its own *sales force*. It is involved with many possible ways of publicising the company's products such as trade fairs, wholesalers displays, press and TV advertising, special campaigns, promotional videos, the Internet, direct mail, and e-mail. The company will also be concerned about the quality of goods and services as well as administering warranty and guarantee services, returns and repairs, etc.

Sales will maintain contacts with customers that will entail the following customer services:

- technical support;
- after-sales service, service engineering;
- product information, prices, and delivery;
- maintaining customer records.

### Distribution: consumer markets

---

For companies operating in the consumer marketing field distribution can be accomplished through a variety of ways. This can include wholesalers, retailers, mail order, or direct selling through the company's own retail outlets. Some companies may use all of these methods. We will examine the wholesale and retail systems, as well as the pricing aspects.

#### ***Retail outlets***

These outlets sell directly to the consumer. Most of these are shops or mail order businesses. Retailers fall into several types, hypermarkets,

supermarkets, multiple shops, departmental stores, co-operative retail societies, independent retailers, voluntary retail chains, franchise outlets, discount stores, etc.

The purpose of retailing is to provide for the availability of goods close to where consumers live. Retailers also study consumer preferences and stock goods accordingly. They also keep manufacturers informed of what it is that consumers want so that supply matches demand.

### ***Wholesale outlets***

If a retailer requires a large range of goods in relatively small quantities it is not very convenient to buy directly from manufacturers. Think of the number of different manufacturers that a small independent grocer would have to deal with if he did deal directly with each manufacturer. Hence the continuing need for wholesalers, who stock goods from many manufacturers and can supply smaller quantities to retailers.

The wholesaler is a middle man and it is said that his presence puts up prices. This is not necessarily the case, since manufacturers can sell to him in bulk quantities and save on transport and administration costs. In effect, wholesalers operate as intermediate storage depots for retailers and therefore provide a useful service. They can usually provide retailers with credit terms of trading, often enabling small businesses to sell before they have to pay for goods, or at least to reduce the impact of the cost of carrying a large range of stock items.

They can also act as a buffer to smooth out demand for manufacture. If demand is seasonal they can buy regularly through the year, thus making it easy for manufacturers to make goods in economic runs and then store stock to meet heavy demand, but which does not place excessive loads on the manufacturer's capacity.

Wholesalers as such have been in decline in recent times, thus many manufacturers have started to deal directly, especially with large retailers, such as the supermarkets. However, he has had to take over the functions of storage, transport, and dealing directly with retailers.

### ***Mail order***

Companies may decide to deal directly with the public through mail order, thus by-passing the wholesaler and retailer. Mail order depends on a good postal service or the existence of transport operators who can provide a similar service. It has the advantage of being nationwide or even international, thus extending the potential market enormously.

In some cases there are very large mail order retailers who buy from manufacturers and sell onto consumers. These companies sometimes operate normal retail outlets as well.

### ***Price structure***

It is common for distributor prices to be expressed at a percentage discount from the price to be paid by the consumer. Thus manufacturers give discounts to wholesalers as an incentive to stock their goods and to provide a profit margin for him. A similar system will be used by the wholesaler when dealing with the retailer. However, if a price to the final consumer is not envisaged or fixed, then the situation is less clear and each party must charge a price which his particular market will stand and

depending on what quantities he needs to sell and what his actual costs are going to be.

In addition to wholesale and trade discounts, quantity discounts may be offered to encourage distributors to buy in large quantities which may be more economical to supply and deliver, since there are economies of scale to be had, such as lower manufacturing, administrative, and transport costs.

Further incentives may be offered in giving cash discounts. Cash discounts reward immediate or early payment. This may enable the manufacturer or wholesaler to reduce his need for working capital, and reduce credit collection and control costs.

Discount structures are therefore used for many purposes. Firstly, to increase sales; secondly, to influence the pattern of sales; and thirdly, to reduce the costs of production and distribution. Sometimes discounts can produce more sales but have very little effect on profits since higher volumes and lower costs may not compensate for lower margins.

You should be aware that price is influenced by many factors:

- (a) actual cost of manufacture;
- (b) what the market will stand;
- (c) what others are charging for similar products;
- (d) consumers' perceptions of quality and value.

The interaction of supply and demand is complex and outside the scope of this course. However, if supply exceeds demand, in general this exerts a downward pressure on prices, as manufacturers and distributors seek to sell goods they have made or bought. The costs of storage and distribution may so high as to force sale at prices which might be below average cost. This is especially true of perishable goods and foodstuffs.

Alternatively, if demand exceeds supply that tends to bid up the price as consumers search for supplies. In some cases the increase in price then serves to limit demand since some potential buyers drop out when the price goes too high, this then acts to dampen demand again and tends to bring equilibrium between supply and demand.

As you can appreciate this can get very complicated when manufacturers, wholesalers, and retailers start to offer the different discounts to try to influence events in their favour. Sometimes, the competition is so cut throat that the only winner is the consumer. In some cases the weaker players go under, leaving the more efficient firms to operate in a less hostile environment. Sometimes the bigger and stronger firms deliberately cut prices so low as to force others out of business and then exploit the consumer when they can dominate the market. However, this can go in reverse again if, then prices go too high and this attracts new players into the market who will then increase supply which will then produce a downward pressure again on prices, and so it goes on.

### **Distribution: industrial markets**

---

For industrial market distribution the situation is more variable. Frequently the seller will have his own fleet of vehicles and may have warehousing facilities or geographically dispersed depots. An example might be a company which manufactures components for the motor industry. It may manufacture in one or more locations and have storage depots located near its customers. It may also deliver products directly to the motor manufacturer's plant either using its own transport or by using an independent haulier.

If the company makes products for international markets it may have to prepare and package products for sea or air freight. This could include using haulage contractors who will deliver direct into Europe using roll-on-roll-off ferries.

### **Price structure**

This is also very variable, but is usually based on the negotiation between the seller and buyer. It may be done through a process of enquiry and quotation or may simply be based on price lists and discounts separately negotiated.

New product design and development is often a crucial factor in the survival of a company. In an industry that is fast changing, firms must continually revise their design and range of products. This is necessary because of the relentless progress of technology as well as the actions of competitors and the changing preferences of customers.

A good example of this situation is the motor industry. The British motor industry has gone through turbulent times, caused by its relative inefficiency compared with Japan and Germany, and also because the quality of its products was below that of its competitors. This difficult position was then made worse by having products that lacked some of the innovative features of the competition.

## **Product development**

---

There are three basic ways of approaching product design and development:

- driven by marketing,
- driven by technology,
- co-ordinated approach.

A system driven by marketing is one that puts the customer needs first, and only produces goods which are known to sell. Market research is carried out which establishes what is needed. If the development is technology driven, then it is a matter of selling what it is possible to make. The product range is developed so that production processes are as efficient as possible and the products are technically superior, hence possessing a natural advantage in the marketplace. Marketing's job is therefore to create the market and sell the product.

Both approaches have their merits, but each of them omit important aspects, hence the idea that a co-ordinated approach would be better. With this approach the needs of the market are considered at the same time as the needs of the production operation, and of design and development. In many businesses this interfunctional system works best, since the functions of R&D, production, marketing, purchasing, quality control, and material control are all taken into account.

However, its success depends on how well the interface between these functions is managed and integrated. Sometimes committees are used as matrix structures or task forces (the latter being set up especially to see in new product developments). In some parts of the motor industry a function called *programme timing* co-ordinates the activities of the major functions by agreeing and setting target dates and events using network planning techniques.

## The development process

---

The basic process is outlined as follows:

- idea generation,
- selection of suitable products,
- preliminary design,
- prototype construction,
- testing,
- final design.

This is a complex process and involves co-operative work between the design and development engineers, marketing specialists, production engineers, and skilled craft engineers to name some of the major players.

Ideas can come from the identification of new customer needs, the invention of new materials or the successful modification of existing products. Selection from new ideas will be based on factors like:

- market potential,
- financial feasibility,
- operations' compatibility.

This means screening out ideas, which have little marketability, are too expensive to make at a profit and which do not fit easily alongside current production processes.

After this, preliminary designs will be made within which trade-offs between cost, quality, and functionality will be made. This can involve the processes of *value analysis* and *value engineering*. These processes look at both the product and the methods of production with a view to maintaining good product performance and durability whilst achieving low cost.

Prototypes are then produced, possibly by hand and certainly not by using mass production methods. This is followed by rigorous testing to verify the marketing and technical performance characteristics required. Sometimes this process will involve test marketing to check customer acceptance of the new product.

Final design will include the modifications made to the design as a result of prototype testing. The full specification and drawings will be prepared so that production can be planned and started.

### Questions 1.1.1

1. Explain why management style is important in an engineering company.
2. Outline two characteristics of effective management style.

### Questions 1.1.2

1. List four factors that affect the price of an engineered product.
2. Draw a flowchart to show the typical stages used in the product development process.

## Production process and facilities management

---

The production or manufacturing operation is at the heart of the business. It translates the designs for products, which are based on market analysis, into the goods wanted by customers.

Decisions have to be made in relation to location of the factor or plant, and the design and layout of production facilities. The design of production processes is interactive with product design, requiring close co-operation with R&D and marketing functions.

Selecting the process of production is important and is strategic in nature. This means that it has a wide impact on the operation of the entire business. Decisions in this area bind the company to particular kinds of equipment and workforce because the large capital investments that have to be made limit future options. For example, a motor manufacturer has to commit very large expenditures to lay down plant for production lines to mass produce cars. Once in production the company is committed to the technology and the capacity created for a long time into the future. There are three basic methods for production processes:

- line flow,
- intermittent flow,
- project.

*Line flow* is the type of system used in the motor industry for assembly lines for cars. It also includes continuous type production of the kind that exists in the chemicals and food industries. Both kinds of line flow are characterised by linear sequences of operations and continuous flows, and tend to be highly automated and highly standardised.

*Intermittent flow* is the typical batch production or job shop, which uses general-purpose equipment and highly skilled labour. This system is more flexible, but is much less efficient than line flow. It is most appropriate when a company is producing small numbers of non-standard products, perhaps to a customer's specification.

Finally *project-based production* is used for unique products which may be produced one at a time. Strictly speaking, there is not a flow of products, but instead there is sequence of operations on the product which have to be planned and controlled. This system of production is used for prototype production in R&D and is used in some engineering companies who produce major machine tool equipment for other companies to use in their factories.

## Capacity planning

---

Once facilities for production have been put in place the next step is to decide how to flex the capacity to meet the predicted demand. Production managers will use a variety of ways to achieve this from maintaining excess capacity to making customers queue or wait for goods to having stocks to deal with excess demand. The process is complex and may require the use of forecasting techniques, together with careful planning.

Scheduling activities are different for each process method and require the use of a variety of techniques. The objectives of good scheduling are:

- meeting customer delivery dates;
- correct loading of facilities;

- planning the starting times;
- ensuring jobs are completed on time.

With any manufacturing facility good *inventory control* is an absolute essential. It is estimated that it costs up to 25% of the cost value of stock items per year to maintain an item in stock. Proper control systems have to be used to ensure that there is sufficient stock for production while at the same time ensuring that too much stock is not held. If stock levels are high there are costs associated with damage, breakage, pilferage, and storage which can be avoided.

### Workforce management

---

This is related to the need to have a workforce trained to use the facilities installed. The important aspects here are:

- work and method study,
- work measurement,
- job design,
- health and safety.

The production manager has to establish standards of performance for work so that the capacity of the factory can be determined, and so that the labour costs of products can be calculated. Work study, method study, and work measurement activities enable this to be done, as well as helping to promote efficient and safe methods of working. The design of jobs is important in respect of workers' health as well as effective work. Good job design can also make the work more interesting and improves employee job satisfaction, which in turn can improve productivity.

### Quality control

---

Quality is a key objective for most engineering companies. It is especially important to the production function that is actually manufacturing the product for the customer.

What is meant by the word quality? It is generally defined as 'fitness for purpose'. In effect this means meeting the identified needs of customers. Thus it is really the customer that determines whether or not a company has produced a quality product, since it is the customer who judges value received and registers satisfaction or dissatisfaction.

This does bring problems for manufacturers since customer perceptions of quality vary, some customers will like a product more than other customers will. Hence a manufacturer has to use some more objective criteria for assessing fitness for purpose. It has been suggested that this must include:

- design quality,
- conformance quality,
- reliability,
- service.

*Design quality* is the primary responsibility of R&D and marketing. It relates to the development of a specification for the product that meets identified needs.

*Conformance quality* means producing a product that conforms to the design specification. A product that conforms is a quality product, even

if the design itself is for a cheap product. That may seem contradictory, but consider the following example. A design is drawn up for a 'budget' camera, which is made from inexpensive materials and has limited capability. This camera serves a particular market. If the manufacture conforms to the specification then the product is of high quality, even though the design is of 'low' quality compared with other more up-market cameras.

*Reliability* includes things like continuity of use measured by things like *mean time between failure* (MTBF). Thus a product will operate for a specified time, on average, before it fails. It should also be maintainable when it does fail, either because it can easily and cheaply be replaced or because repair is fast and easy. *Service* relates to after-sales service, guarantees, and warranties.

Quality control is therefore concerned with administering all of these aspects. In the UK, there are general standards for *quality systems*, the most relevant one here is BS 5750 and the international counterpart, ISO 9000. The activities of quality control include the following:

- inspection, testing, and checking of incoming materials and components;
- inspection, testing, and checking of the company's own products;
- administering any supplier quality assurance systems;
- dealing with complaints and warranty failures;
- building quality into the manufacturing process.

Some of these activities are done after the event to monitor quality, other activities may be carried out to prevent problems before they occur. Some activities may be carried out to determine causes of failure that relate to design rather than manufacturing faults.

### Questions 1.1.3

1. Describe three basic methods used for the production process.
2. Explain what is meant by inventory control.
3. Explain why good job design is important.
4. List five activities carried out by a quality control department.

## Purchasing and supply

In large businesses purchasing is done by professional buyers, and is therefore a centralised activity. When a company has a large purchasing budget this makes economic sense, since the large purchasing power gives advantages in negotiating for keen prices, better delivery times or increased quality. In small businesses the purchasing function is not centralised, usually because the operation is not large enough to support the employment of specialists. However, the basic principles of purchasing are the same, whatever the structure of the engineering company.

The main functions are:

- researching sources of supply;
- making enquiries and receiving quotations;
- negotiating terms and delivery times;



## 14 Higher National Engineering

- placing contracts and orders;
- expediting delivery;
- monitoring quality and delivery performance.

The basic documents used are as follows:

- requisitions from departments to buyer;
- enquiry forms or letters to suppliers;
- quotation document in reply to enquiry;
- order or contract to buy;
- advice note: sent in advance of goods;
- invoice: bill for goods sent to the buyer;
- debit note: additional/further charge to invoice;
- credit note: reduction in charge to invoice;
- consignment note: accompanies goods for international haulage, containing full details of goods, consignee, consignor, carrier, and other details;
- delivery note: to accompany goods.

Purchasing procedures involve raising a requisition to buy. This may then require obtaining quotations or estimates in order to choose the best supplier. Orders are sent to the chosen supplier. Goods are despatched by the supplier, together with a delivery or advice note. When goods are received they are checked to see that the details on the delivery note agree with the actual goods received and that the goods have in fact been ordered. Goods not ordered may be refused. Accepted deliveries are signed for on the supplier's copy and given back to the driver. A goods' received note is raised and sent to the purchasing department, so that the accounting function can be given confirmation of delivery before making payment against receipt of invoice.

These procedures are for the purpose of making sure that the goods are delivered on time, in the correct quantities, of the correct specifications and of the desired quality. Only if all are well, then only the payment is authorised.

## Organisational structures

The 1990s in the UK has seen a significant move towards having flat organisations. The process has been described as *de-layering* and it is regarded as an organisation structure that permits better communications both up and down the organisational levels. It is also seen as more cost effective and responsive in dealing with demands placed on modern businesses.

Note that flat organisations are also hierarchical. A hierarchy is an organisation with grades or classes ranked one above another. A flat organisation also meets that criterion. A typical flat organisational structure is shown in Figure 1.1.1.

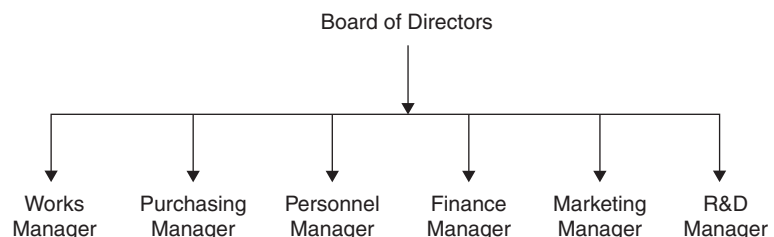


**Figure 1.1.1** A typical flat organisational structure

*Tall structures* are opposite to flat structures and contain many layers. Figure 1.1.2 shows an example. Tall structures usually exist only in large organisations because of the necessity of dividing the tasks into chunks of work that can be handled by individual managers, departments, and sections. Tall structures are also hierarchies, only they contain many more levels than flat structures.

These are structures which have many or few layers showing rank order from top to bottom. They show a chain of command, with the most senior posts at the top and the most junior posts at the bottom. Plain hierarchies are the most common representations of organisations, as some aspect of hierarchy exists in all organisations.

An *organisation chart* is a useful way of representing the overall structure, but it tells only part of the story. You should be aware that other documents are needed to fully understand how the organisation works, as we observed earlier.

Hierarchical organisations can take many forms. We have already examined flat and tall structures. There are several other forms with which you should be familiar. One of the most common is the functional design. Figure 1.1.3 shows a functional organisation. The main functions of a commercial business are marketing, finance, purchasing and supply, manufacturing, R&D, and personnel. Notice how each functional manager

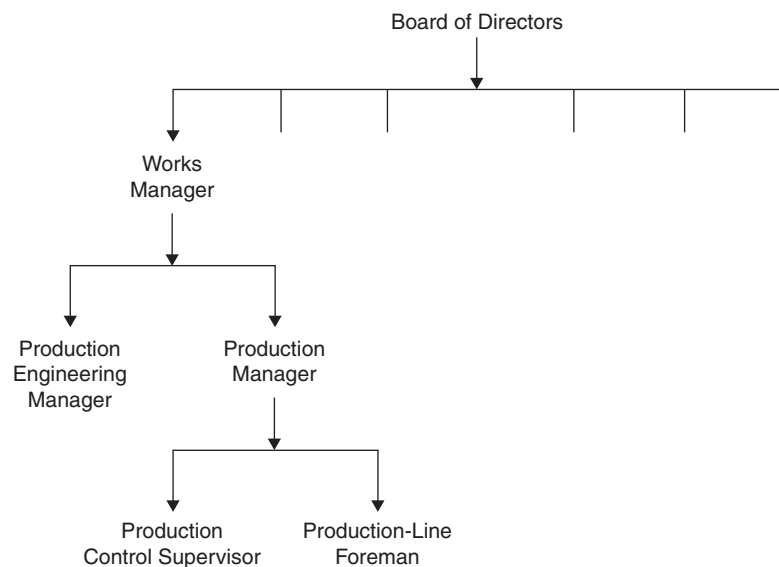


**Figure 1.1.2** A typical tall organisational structure

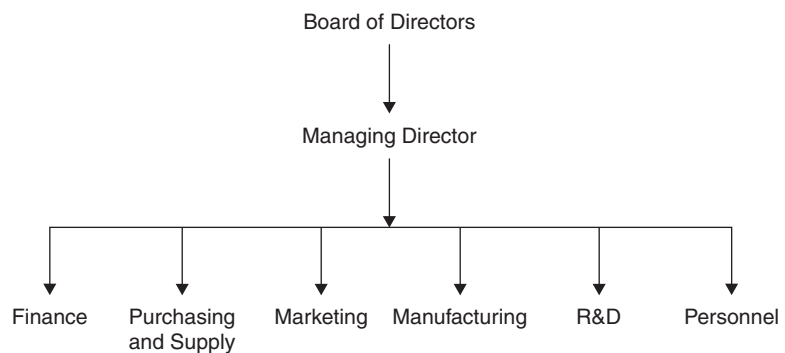


**Figure 1.1.3** A typical functional organisational structure

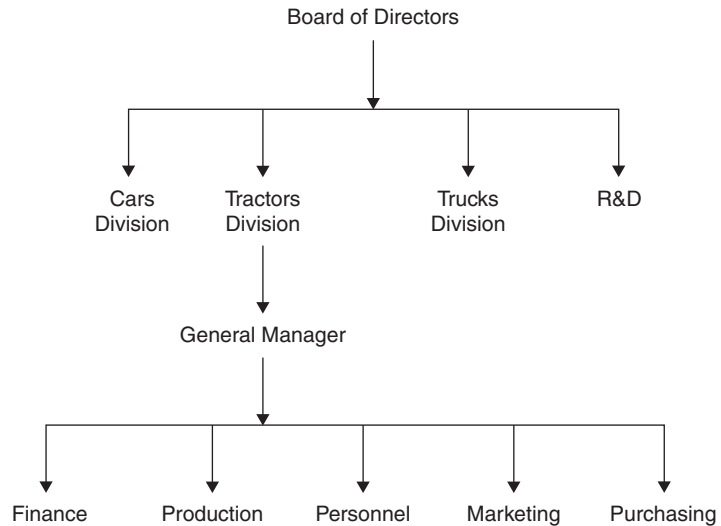


**Figure 1.1.4** A typical federal organisational structure

reports to the managing director who co-ordinates their activities. There are a number of advantages in functional structures:

- specialists can be grouped together;
- it appears logical and easy to understand;
- co-ordination is achieved via operating procedures;
- suits stable environments and few product lines;
- works well in small- to medium-size businesses.

As a business grows the functional structure becomes less and less useful. This is because there are many more products and these may be manufactured in separate divisions of a company, especially if economies of scale are introduced into the manufacturing process.

Figure 1.1.4 shows an organisation based on major product lines and is really a federal structure which still has functional activities, but at a lower level in the organisation, except for the R&D function which is centralised. The managers of these operating divisions will control most of the functions required to run the business. In many conglomerate businesses this federal arrangement is achieved by having a holding company, which may be a *public limited company* (plc), which wholly owns a number of subsidiary companies, which are in effect divisions of the main business. Figure 1.1.5 shows an example.

An alternative to the product-based divisional design is one based on geographical divisions. Figure 1.1.6 shows a geographical design which still has many of the functions located at a head office, but which has branches dispersed around the country. These branches or divisions handle sales and manufacture, but are supported by head office for the other functions.

A *matrix structure* as its name implies has a flow of authority in two dimensions. Departmental or functional authority flows vertically and project management authority flows horizontally. Such a structure is shown in Figure 1.1.7. This depicts a company that designs and makes machines and tools to customer specifications.

Some businesses operate as a series of projects. This is common, for example, in construction and in some types of engineering company, especially those that design one-off products or who design manufacturing equipment used by other manufacturers. As the chart shows, there are

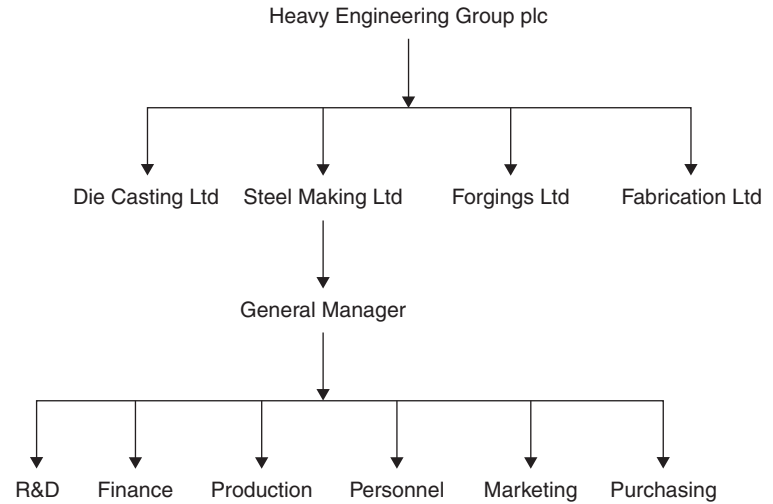


**Figure 1.1.5** A typical conglomerate organisational structure

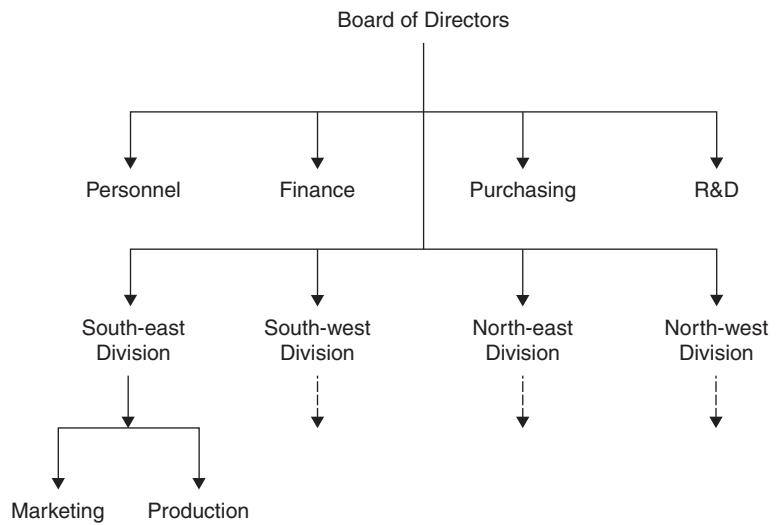


**Figure 1.1.6** A typical divisional organisational structure

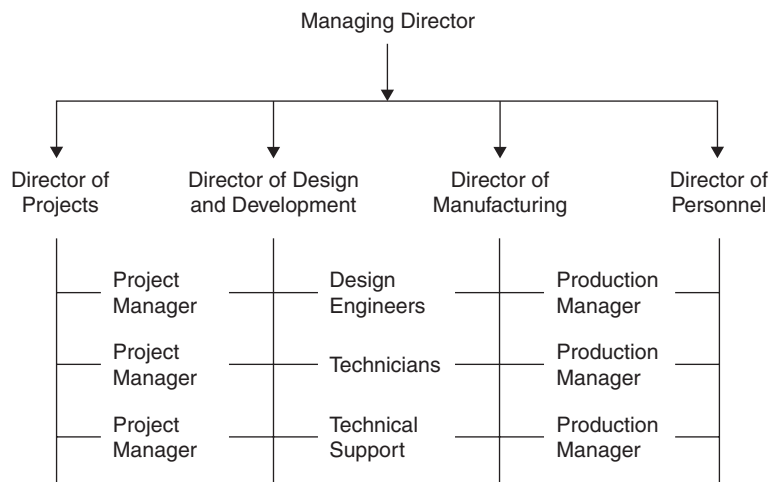


**Figure 1.1.7** A typical matrix organisational structure

familiar functional divisions, and members of functional departments still report to their functional line manager.

However, the *project manager* has the job of co-ordinating the work of functional specialists to ensure that the project is completed on time, to specification and within cost limits. In the main, project managers do not have direct line authority, but have to influence and persuade others to achieve *targets*. They may have formal authority over project budgets and can set time schedules and decide what work is to be done, but little else. They also work as an interface with clients and subcontractors and their influence is often critical to the success of the project. Although they do not have formal authority over individual staff or their line managers, they nevertheless operate with the full support of senior managers. This means that functional specialists are obliged to provide the fullest co-operation and help; otherwise, they become answerable for failure to their own senior line managers.

The matrix system works very well in project-based industries, and that is why the design is used. It still retains many of the ingredients of other structures, and still has substantial hierarchical elements.

The structures discussed above are just examples of the main design principles for organisations. There are numerous variations and rarely do we find 'pure' forms of organisation structure. We need to remember that organisations are created to serve the goals of their owners and that the precise structure will be designed to meet the needs of the business.

### Questions 1.1.4

1. Explain the justification for (a) flat structures and (b) tall structures.
2. Explain why federal or divisional structures are used.
3. Explain the reason for matrix structures.
4. Explain what is meant by a hierarchical structure and illustrate your answer with a diagram.

### Question 1.1.5

Draw a chart showing the organisational structure for your company or your college. Use a computer-aided design (CAD) or drawing package to produce the chart and include job titles and names where appropriate. Comment on the type of structure, including how and why it has involved.

## Information and interaction

The commercial and engineering functions within an engineering company, as well as individual teams within each, need to work together in order to function effectively and efficiently. This interaction might involve sharing information and making decisions about:

- processes and systems;
- working procedures;

- the people involved, including customers, suppliers, and other employees.

Understanding the information that is shared and exchanged between the various departments within an engineering company is extremely important. This can include:

(a) documents such as:

- design specifications,
- purchase orders,
- invoices,
- production schedules,
- quotations;

(b) information about:

- stock levels,
- work in progress,
- resource utilisation sales.

Note that all of the above may exist either as *hard copy* or may be in electronic form (such as a word processed document, a hypertext document, or a spreadsheet file). Many companies use an *intranet* to facilitate access to this information.

The finance and accounting function interfaces with all other functions within an engineering company. Its recording and monitoring activities are central to, and have a major impact on, the whole business. In conjunction with the manufacturing function, *sales forecasts* will be used to prepare *factory schedules*. Production may be sent to a warehouse and then put into the delivery and distribution system. From there the sales force will ensure that customers receive their orders when required.

Alternatively, delivery of specific customer orders may be made directly to customers.

Marketing will supply information on pricing structures. Prices may be determined primarily by the market rather than the cost of manufacture. Finance may provide cost information, but marketing may make the final pricing decisions.

Marketing will also identify customer needs and will liaise with product development activities on possible new products or modifications to existing products. R&D will initiate design studies and prototypes for new products and may supply some items for market testing. Engineers involved with design and development will be given information on customer needs and preferences and will be expected to produce designs which meet those requirements.

There will be a need to communicate details of the costs of new products or redesigned products. The processes required to produce new components or whole new products will also require costing. R&D may specify the manufacturing process, but manufacturing engineering departments located at the production facility will implement them, and may also share in the costing process.

New products will have different characteristics, and perhaps be made from different materials from previous products with similar functionality. This will require liaison between design engineers and manufacturing engineering on methods for production and in deciding what manufacturing equipment and machine tools are required. Detailed process

sheets may be required which show how products are to be assembled or made.

Whilst the particular methods of production are the province of production management, the designer has to be aware of the implications for his design of different methods of manufacture, whether this be batch production, assembly-lines or one-off projects. Detailed specifications of the new and changed product will be communicated and there may be liaison on temporary and permanent deviations to original specifications in order to facilitate production.

When quality problems appear and are related to faulty design there will be liaison on ways in which design modifications can be phased into production as soon as possible. There will be proposals for the replacement of machines and equipment used for manufacturing and production. This function may require quite sophisticated techniques for what is called *investment appraisal*, so that the company can choose the best methods of manufacture from several alternatives.

Also important is the control of raw materials and component stocks, especially the levels of 'work-in-process'. Finance manager will want to restrict stock levels to reduce the amount of capital tied up in stocks, whilst the production manager will be concerned with having sufficient stock to maintain production, but avoiding congestion of factory floor space.

Budgetary control of production cost centres will involve regular contact and advice from the finance function. Matters of interest will be costs of production, wastage rates, labour costs, obsolescent stock, pilferage, etc.

Specifications and drawings will be sent to the *buyer* for new products or machines for purchase. Problems of design and delivery will be discussed, modifications to designs will be sent to suppliers through the buying department. New product launches will be co-ordinated with R&D, the supplier, and of course, the manufacturing department. The buyer would be involved with supplies of new raw materials, new designs for components and will negotiate costs of tooling and long-term contracts.

Liaison between the buyers and production managers will be required to establish levels of supplies for new materials or components. Assistance may be given by the buyer to deal with quality and inspection problems and in dealing with return and replacement of defective materials. Sometimes buyers may be an interface with production and R&D in dealing with temporary or permanent deviations from the original engineering specifications. Chasing deliveries and ensuring supplies for factory use may be a major daily routine for some buying departments.

Systems for quality control may include some form of supplier quality assurance. The buying department will represent company interests to suppliers and may only use suppliers who have passed the company's quality assurance standards. The buyer will be involved with searches for new suppliers who can meet existing and new quality requirements.

*Stock control* systems used within the factory will affect the way purchasing is done. Economic order quantities may be established, in which the buyer has to take into account when arranging supplies. Deliveries may have to be phased according to minimum, maximum, and re-order stock levels. The buyer will need a clear understanding of the importance of deliveries which enable the company to control its inventory costs, while at the same time ensuring a reliable supply of materials and components for production.

The company may operate a *just-in-time* (JIT) system. JIT originated in Japan and is a way of delivering supplies at the point in time they are required by production. JIT avoids the costs of holding buffer stocks of raw materials and components. It works well when suppliers are dependable and when transport systems are good. The buyer will liaise with the factory on the establishment and operation of the JIT for given products. Other techniques employed by companies for manufacturing control include *value-added chains* (each stage in the manufacturing process ‘adds value’ to the output of the previous stage) and *statistical process control* (which uses statistical analysis to control production including the acceptable levels of defects in production).

In all organisations there will also be the routine matters of passing invoices for payment of goods or dealing with returns for credit so that accounts department can pay for goods received. Materials purchasing will be subject to budgetary constraints like most other company activities. The purchasing department will be involved, either directly or indirectly in budgets for inventory levels, and in setting up minimum, maximum, and re-order levels for stocks.

Monthly monitoring of *inventory levels* will be performed by the company’s accounting function and purchasing activities may be regulated in order to ensure that stock of components and raw materials stay within the agreed levels.

## Financial constraints

---

Ultimately, all engineering activities are bound by financial decisions. For example, a company will need to decide whether to make a component or to buy one from another company, or whether to invest in a new production facility. They will need to put a clear case for a new product or service a financial planning stage. Companies use financial information and data to help make these decisions. You need to be familiar with the sorts of data and information used to make decisions about:

- whether a company should *make or buy* components, assemblies or services for a product;
- the production volume needed to generate a required profit (using break-even analysis);
- investment and operational costs.

The usual techniques for costing, budgeting, inventory control, investment appraisal, make or buy decisions, and forecasting are examined and compared using examples drawn from a variety of situations. *Cost control* is concerned with collecting operating cost data, and then monitoring and controlling these costs. Budgeting is the process of forecasting the financial position and providing a plan against which to monitor profitability. *Inventory control* is the establishment of economic levels of stock whilst maintaining production. *Investment appraisal* enables managers to choose the best projects for investment using *discounted cash-flow* techniques. Other techniques are examined that enable choices to be made to *make or buy*, and to forecast sales and production.

*Cost accounting* is part of the management accounting function, indeed, without a system for cost accounting, effective management accounting could not exist. *Management accounting* exists to provide information for the internal control and management of a business. Thus cost accounting is necessary for a company to be able to identify responsibility for



costs and to exercise control over actual costs compared with planned expenditure.

*Budgetary control* is very important function in respect of R&D as well as manufacturing activities. Engineers are subject to the discipline of budgetary control as are other specialists and it is essential that costs can be monitored and controlled so that engineering projects meet target costs and profits.

### Questions 1.1.6

1. Explain why stock control is important.
2. Explain what is meant by 'JIT' manufacturing.
3. Describe three different types of information that is exchanged between the different departments within an engineering company. What form does this information take?

## 1.2 COSTING SYSTEMS AND TECHNIQUES

To meet the requirements of this unit, you need to be able to identify and describe appropriate costing systems and techniques for specific engineering business functions. You also need to be able to measure and evaluate the impact of changing activity levels on engineering business performance. We will start by introducing some common costing systems used in engineering.

### Costing systems

Any modern business enterprise needs to have in place an effective costing system that take into account the real cost of manufacturing the product or delivering the service that it provides. Without such a system in place it is impossible to control costs and determine the overall profitability of the business operation.

Cost accounting is necessary for a company to be able to exercise control over the actual costs incurred compared with planned expenditure. From the point of view of cost control, a costing system should not only be able to identify any costs that are running out of control but should also provide a tool that can assist in determining the action that is required to put things right.

#### **Job costing**

Job costing is a very simple costing technique. It usually applies to a unique operation, such as fitting a part or carrying out a modification to a product. Typical operations in which job costing is commonly used include:

- supplying a unique or 'one-off' item;
- painting and decorating a building;
- converting or adapting a product to meet a particular customer's requirements.

Job costing always has at least three elements: direct labour; direct materials and absorbed overheads. Sometimes there is an additional direct cost—direct expenses. Let us take an example.

A *jobbing builder*, John Smith, has been asked to supply a quotation for supplying a 'one-off' shipping container for a diesel generator. The parts and materials required to build the shipping container are as follows:

| <i>Item</i>                          | <i>Quantity</i> | <i>Price per unit</i>  | <i>Cost (£)</i> |
|--------------------------------------|-----------------|------------------------|-----------------|
| ½" chipboard                         | 4 m × 2 m       | @ £1.75/m <sup>2</sup> | 14.00           |
| 2" × 1" timber                       | 10 m            | @ £0.80/m              | 8.00            |
| Panel pins                           | 50              | @ £0.01/ea.            | 0.50            |
| Countersunk screws                   | 30              | @ £0.02/ea.            | 1.60            |
| Adhesive                             | 1               | @ £1.28/ea.            | 1.28            |
| Joints                               | 16              | @ £0.45/ea.            | 7.20            |
| <b>Total for parts and materials</b> |                 |                        | <b>£32.58</b>   |

ea., each article.

To this should be added the cost of labour. Let us assume that this amounts to 3 h at £20.00/hour (this figure includes the overheads associated with employment, such as National Insurance contributions). Hence the cost of labour is:

| <i>Item</i>             | <i>Quantity</i> | <i>Price per unit</i> | <i>Cost (£)</i> |
|-------------------------|-----------------|-----------------------|-----------------|
| Labour                  | 3 h             | @ £20.00/hour         | 60.00           |
| <b>Total for labour</b> |                 |                       | <b>£60.00</b>   |

We can add the cost of labour to the total *bill of materials* to arrive at the final cost for the job which amounts to £92.58. Note that parts and materials *may* be supplied 'at cost' or 'marked up' by a percentage which can often range from 10% to 50% (and sometimes more).

### Example 1.2.1

John Smith has decided to mark up the costs of timber by 25% and all other sundry items by 10%. Determine the amount that he will charge for the transit container.

The revised bill of materials is as follows:

| <i>Item</i>                 | <i>Quantity</i> | <i>Price per unit</i>  | <i>Cost (£)</i> | <i>Charge (£)</i> |
|-----------------------------|-----------------|------------------------|-----------------|-------------------|
| ½" chipboard                | 4 m × 2 m       | @ £1.75/m <sup>2</sup> | 14.00           | 17.50             |
| 2" × 1" timber              | 10 m            | @ £0.80/m              | 8.00            | 10.00             |
| Panel pins                  | 50              | @ £0.01/ea.            | 0.50            | 0.55              |
| Countersunk screws          | 30              | @ £0.02/ea.            | 1.60            | 1.76              |
| Adhesive                    | 1               | @ £1.28/ea.            | 1.28            | 1.41              |
| Joints                      | 16              | @ £0.45/ea.            | 7.20            | 7.92              |
| Labour                      | 3 hours         | @ £20.00/hour          | 60.00           | 60.00             |
| <b>Total amount charged</b> |                 |                        | <b>£99.14</b>   |                   |

ea., each article.

Large companies also use job costing when they produce a variety of different, and often unique, products. These products are often referred to as *custom built* and each is separately costed as a 'job' in its own right. This type of production is described as *intermittent* (and traditionally referred to as *job shop* production) to distinguish it from the *continuous* or *assembly-line* production associated with the manufacture of a large number of identical units.

In jobbing production, individual manufactured units are normally produced to meet an individual customer's requirements and production is not normally speculative. Costs are agreed before manufacturing starts and form the basis of a contract between the manufacturer and the customer.

### **Contract costing**

Contract costing relates to larger jobs (so is conceptually the same as job costing) and is longer lasting. Contract costing is usually used for things like civil engineering, shipbuilding, and defence. Contract costing is more complex than job costing.

### **Parts costing**

Parts costing is straightforward and is simply a question of determining the cost of all of the physical parts and components used in a manufactured or engineered product. Parts costing works from the 'bottom-up' – in other words, the cost of each individual component (i.e. the *per unit cost*) is determined on the basis of the given *standard supply multiple*. As an example of parts costing, consider the following example.

Centralux is a small engineering company that specialises in the manufacture of domestic central heating controllers. Their latest product uses the following parts:

| <i>Component</i>            | <i>Quantity</i> | <i>Cost ea.</i> | <i>Total (£)</i> |
|-----------------------------|-----------------|-----------------|------------------|
| Bridge rectifier            | 1               | 0.18            | 0.18             |
| Capacitor ceramic disk 20%  | 4               | 0.04            | 0.16             |
| Capacitor electrolytic 20%  | 1               | 0.21            | 0.21             |
| Capacitor polyester 10%     | 4               | 0.12            | 0.48             |
| Clips – plastic             | 2               | 0.04            | 0.08             |
| Connector – mains           | 1               | 0.25            | 0.25             |
| Connector – PCB type        | 2               | 0.08            | 0.16             |
| Connector – solder terminal | 2               | 0.02            | 0.04             |
| Display – LED               | 1               | 0.45            | 0.45             |
| Fascia trim                 | 1               | 0.15            | 0.15             |
| Fuse – 20 mm                | 1               | 0.12            | 0.12             |
| Fuse holder – 20 mm         | 1               | 0.13            | 0.13             |
| Keypad – membrane type      | 1               | 0.89            | 0.89             |
| Miniature PCB transformer   | 1               | 1.99            | 1.99             |
| Nuts – M3                   | 8               | 0.02            | 0.16             |
| Opto-isolator               | 2               | 0.22            | 0.44             |
| Pillars – plastic           | 4               | 0.03            | 0.12             |
| Plastic enclosure           | 1               | 0.89            | 0.89             |
| PCB                         | 1               | 1.45            | 1.45             |

(continued)

| <i>Component</i>           | <i>Quantity</i> | <i>Cost ea.</i> | <i>Total</i>  |
|----------------------------|-----------------|-----------------|---------------|
| Programmed controller chip | 1               | 1.05            | 1.05          |
| Resistor 0.25 W 5%         | 5               | 0.01            | 0.05          |
| Resistor 0.5 W 5%          | 2               | 0.02            | 0.04          |
| Screws – M3                | 8               | 0.03            | 0.24          |
| Switch – mains             | 1               | 0.55            | 0.55          |
| Temperature sensor         | 1               | 0.25            | 0.25          |
| Transducer – piezoelectric | 1               | 0.33            | 0.33          |
| Triac                      | 2               | 0.42            | 0.84          |
| Voltage regulator          | 1               | 0.21            | 0.21          |
| Washers – M3               | 8               | 0.01            | 0.08          |
| <b>Total cost</b>          |                 |                 | <b>£11.99</b> |

ea., each article; PCB: printed circuit board; LED: light emitting diode.

It is often useful to group together individual component parts under groupings of similar items. The reason for this is that such groupings tend to be the subject to the same fluctuation in cost. We can thus quickly determine the effect of market fluctuations by examining the effect of changes on particular groups of parts.

### Example 1.2.2

Group together the parts used in the Centralux domestic central heating controller under the following headings: *Hardware*, *Semiconductors*, *Passive components*, and *Miscellaneous*. Determine the proportion of the total cost by part category.

| <i>Item</i>                | <i>Quantity</i> | <i>Cost ea.</i> | <i>Total (£)</i> |
|----------------------------|-----------------|-----------------|------------------|
| <b>Hardware</b>            |                 |                 |                  |
| Clips – plastic            | 2               | 0.04            | 0.08             |
| Fascia trim                | 1               | 0.15            | 0.15             |
| Nuts – M3                  | 8               | 0.02            | 0.16             |
| Pillars – plastic          | 4               | 0.03            | 0.12             |
| Plastic enclosure          | 1               | 0.89            | 0.89             |
| PCB                        | 1               | 1.45            | 1.45             |
| Screws – M3                | 8               | 0.03            | 0.24             |
| Washers – M3               | 8               | 0.01            | 0.08             |
| <b>Subtotal</b>            |                 |                 | <b>£3.17</b>     |
| <b>Semiconductors</b>      |                 |                 |                  |
| Bridge rectifier           | 1               | 0.18            | 0.18             |
| Display – LED              | 1               | 0.45            | 0.45             |
| Opto-isolator              | 2               | 0.22            | 0.44             |
| Programmed controller chip | 1               | 1.05            | 1.05             |
| Triac                      | 2               | 0.42            | 0.84             |
| Voltage regulator          | 1               | 0.21            | 0.21             |
| <b>Subtotal</b>            |                 |                 | <b>£3.17</b>     |

(continued)

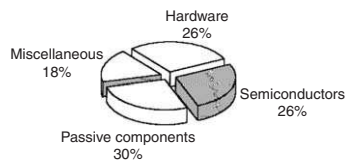


**Figure 1.2.1** Proportion of costs by part category for Centralux's domestic heating controller

| Item                        | Quantity | Cost ea. | Total (£)     |
|-----------------------------|----------|----------|---------------|
| <i>Passive components</i>   |          |          |               |
| Capacitor ceramic disk 20%  | 4        | 0.04     | 0.16          |
| Capacitor electrolytic 20%  | 1        | 0.21     | 0.21          |
| Capacitor polyester 10%     | 4        | 0.12     | 0.48          |
| Miniature PCB transformer   | 1        | 1.99     | 1.99          |
| Resistor 0.25 W 5%          | 5        | 0.01     | 0.05          |
| Resistor 0.5 W 5%           | 2        | 0.02     | 0.04          |
| Temperature sensor          | 1        | 0.25     | 0.25          |
| Transducer – piezoelectric  | 1        | 0.33     | 0.33          |
| <b>Subtotal</b>             |          |          | <b>£3.51</b>  |
| <i>Miscellaneous</i>        |          |          |               |
| Connector – mains           | 1        | 0.25     | 0.25          |
| Connector – PCB type        | 2        | 0.08     | 0.16          |
| Connector – solder terminal | 2        | 0.02     | 0.04          |
| Fuse – 20 mm                | 1        | 0.12     | 0.12          |
| Fuse holder – 20 mm         | 1        | 0.13     | 0.13          |
| Keypad – membrane type      | 1        | 0.89     | 0.89          |
| Switch – mains              | 1        | 0.55     | 0.55          |
| <b>Subtotal</b>             |          |          | <b>£2.14</b>  |
| <b>Total cost</b>           |          |          | <b>£11.99</b> |

ea., each article.

The proportion of costs by part category is shown in the pie chart of Figure 1.2.1.

### Question 1.2.1

Determine the effect on the total cost of the Centralux domestic central heating controller when the cost of semiconductors increases by 5% and the cost of hardware falls by 10%.

### Process costing

Process costing takes into account the cost of a continuous manufacturing process and apportions part of the cost of each process to an individual product. Typical processes might be:

- forming, bending, or machining of metal and plastic parts;
- flow soldering of PCBs;
- heat treatment of metal parts;
- paint spraying and finishing.

Process costing is used in industries that operate on a continuous basis, such as chemical plants, petroleum, or food production. In order to carry out process costing it is necessary to show how the flow of products is costed at each stage of the process; Process 1, Process 2, Process 3, and so on to the finished product.

The following example illustrates one stage of process costing. Note that, when determining the total cost of manufacturing a product, it is essential to take into account the notional cost of *all* of the processes involved.

### Example 1.2.3

Centralux has invested in a flow soldering plant in order to partly automate the manufacture of their domestic central heating controller. The flow soldering plant operates at a rate of 50 units per hour and its operating cost (including capital cost recovery calculated over a nominal 8-year asset life) amounts to £6000/week plus £10 material costs per hour. Determine the unit cost of the flow soldering process based on: (a) 70 h operation per week and (b) 84 h per week.

(a) Based on 70 h operation per week, the total cost of the flow soldering process will be given by:

$$\text{Total cost} = £6000 + (70 \times £10) = £6700.$$

At 50 units per hour, the total weekly production will be given by:

$$\text{Total production} = 70 \times 50 = 3500.$$

The cost, per unit, will thus be given by:

$$\text{Cost per unit} = £6700/3500 = \mathbf{£1.91}.$$

(b) Based on 84 h operation per week, the total cost of the flow soldering process will be given by:

$$\text{Total cost} = £6000 + (84 \times £10) = £6840.$$

At 50 units per hour, the total weekly production will be given by:

$$\text{Total production} = 84 \times 50 = 4200.$$

The cost, per unit, will thus be given by:

$$\text{Cost per unit} = £6840/4200 = \mathbf{£1.63}.$$

## Costing techniques

Any engineering business is liable to incur a variety of costs. These will typically include:

- rent for factory and office premises;
- rates;
- energy costs (including heating and lighting);
- material costs;
- costs associated with production equipment (purchase and maintenance);
- salaries and National Insurance;
- transport costs;

- postage and telephone charges;
- insurance premiums.

Given the wide range of costs above, it is often useful to classify costs under various headings, including fixed and variable costs, overhead and direct costs, average and marginal costs, and so on.

In order to be able to control costs, it is, of course, vital to ensure that all of the costs incurred are known. Indeed, the consequences of not being fully aware of the costs of a business operation can be dire!

This section examines a number of different methods used by businesses to determine the total cost of the product or service that they deliver. The prime objective of these techniques is that of informing commercial decisions such as:

- How many units have to be produced in order to make a profit?
- Is it cheaper to make or buy an item?
- What happens to our profits if the cost of production changes?
- What happens to our profits if the cost of parts changes?

### **Absorption costing**

One method of determining the total cost of a given product or service is that of adding the costs of overheads to the direct costs by a process of *allocation*, *apportionment*, and *absorption*. Since *overheads* (or *indirect costs*) can be allocated as whole items to production departments, it is possible to arrive at a notional amount that must be added to the cost of each product in order to cover the production overheads.

$$\text{Mark-up} = \frac{\text{total of fixed and variable costs attributable to the product}}{\text{total number of units produced}}.$$

### **Marginal costing**

Marginal costing provides us with an alternative way of looking at costs that provides an insight into the way costs behave by allowing us to observe the interaction between costs, volumes, and profits. The marginal cost of a product is equal to the cost of producing one more unit of output (we shall return to this later).

There are a number of advantages of using marginal costing, notably:

- Marginal costing systems are simpler to operate than absorption costing systems because they avoid the problems associated with overhead apportionment and recovery.
- It is easier to make decisions on the basis of marginal cost calculations. Where several products are being produced, marginal costing can show which products are making a contribution and which are failing to cover their variable costs.

The disadvantages of marginal costing include:

- The effect of time (and its effect on *true cost*) tends to be overlooked in marginal costing.
- There is a temptation to spread fixed costs (or to neglect these in favour of more easily quantified variable costs).

Marginal costing is more useful for management decision-making than absorption costing because it avoids using estimation to determine

#### **Another view**

In absorption costing, each product manufactured is made (at least in theory) to cover all of its costs. This is achieved by adding a notional amount to the total unit cost of each product. We sometimes refer to this as *cost-plus*.

overheads. The choice of whether to use absorption costing or marginal costing is usually governed by factors such as:

- the system of financial control used within a company (e.g. *responsibility accounting* is consistent with absorption costing);
- the production methods used (e.g. marginal costing is easier to operate in simple processing applications, whereas absorption costing is usually preferred when several different products require different plant and processing techniques);
- the significance of the prevailing level of overhead costs.

### Activity-based costing

Activity-based costing is an attempt to assess the 'true' cost of providing a product or service. Knowledge of the 'true' cost is not only important in helping us to identify opportunities for cost improvement but it also helps us to make strategic decisions that are better informed.

Activity-based costing focuses on indirect costs (*overheads*). It does this by making costs that would traditionally be considered indirect into direct costs. In effect, it traces costs and overhead expenses to an individual *cost object*. The basic principles of activity-based costing are shown in Figure 1.2.2.

Activity-based costing is particularly useful when the overhead costs associated with a particular product are significant and where a number of products are manufactured in different volumes. Activity-based costing is particularly applicable where competition is severe and the margin of selling price over manufacturing cost has to be precisely determined.

The steps required to carry out activity-based costing are as follows:

- (1) Identify the activities.
- (2) Determine the cost of each activity.
- (3) Determine the factors that drive costs.
- (4) Collect the activity data.
- (5) Calculate the product cost.

The use of activity-based costing is best illustrated by taking an example. A small manufacturing company, EzBild, has decided to carry out activity-based costing of its two products: an aluminium folding ladder and a modular work platform. The following table summarises the activity required for these two products:

| Activity     | Ladders<br>(per unit) | Cost<br>(£)   | Platforms<br>(per unit) | Cost<br>(£)   | Total<br>(£)  |
|--------------|-----------------------|---------------|-------------------------|---------------|---------------|
| Set-up       | 1 @ £25 000           | 25 000        | 1 @ £35 000             | 35 000        | 60 000        |
| Manufacture  | 1500 @ £6             | 9000          | 500 @ £30               | 15 000        | 24 000        |
| Assembly     | 1500 @ £2             | 3000          | 500 @ £10               | 5000          | 8000          |
| Inspection   | 1500 @ £1             | 1500          | 500 @ £2                | 1000          | 2500          |
| Packaging    | 1500 @ £1             | 1500          | 500 @ £6                | 3000          | 4500          |
| <b>Total</b> |                       | <b>40 000</b> |                         | <b>59 000</b> | <b>99 000</b> |

The activity-based product cost for each ladder thus amounts to  $\text{£}40\,000/1500 = \text{£}26.67$ , whilst the activity-based product cost for each

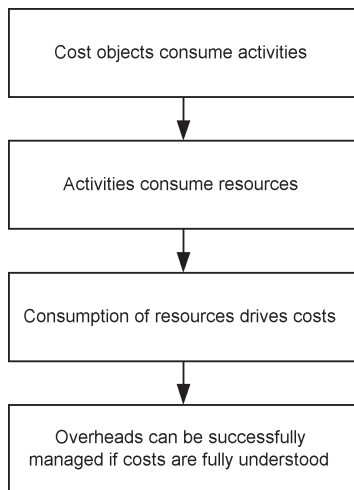


**Figure 1.2.2** Principles of activity-based costing



platform amounts to  $\text{£}59\,000/500 = \text{£}118$ . To this should be added the direct (material) costs of each product. Assuming that this amounts to  $\text{£}20$  for the ladder and  $\text{£}80$  for the platform, we would arrive at a cost of  $\text{£}46.67$  for the ladder and  $\text{£}198$  for the platform.

Traditional cost accounting would have arrived at two rather different figures. Let us assume that 3300 h of direct labour are used in the manufacturing plant. Dividing the total overhead cost of  $\text{£}99\,000$  by this figure will give us the hourly direct labour cost of  $\text{£}30/\text{hour}$ . If ladders require 1 hour of direct labour and platforms require 3.6 h of direct labour the allocation of costs would be  $\text{£}30$  per ladder and  $\text{£}108$  per platform. Adding the same direct (material) costs to this yields a cost of  $\text{£}50$  for the ladder and  $\text{£}188$  for the platform!

### Question 1.2.2

DataSwitch Inc. specialises in the production of switches that can be used to link several personal computers (PCs) to a shared printer. The company currently manufactures a low-cost manually operated data switch and a more expensive automatic data switch. Both types of switch are packed in multiples of five before they are despatched to retail outlets. An analysis of the company's production reveals the following:

|                                           | <i>Manual data switch</i>      | <i>Automatic data switch</i>   |
|-------------------------------------------|--------------------------------|--------------------------------|
| Production volume                         | 1000                           | 250                            |
| Direct materials cost                     | $\text{£}15/\text{unit}$       | $\text{£}25/\text{unit}$       |
| <i>Activity-based production analysis</i> |                                |                                |
| Set-up costs                              | 1 @ $\text{£}10\,000$          | 1 @ $\text{£}10\,000$          |
| Manufacturing costs                       | 1000 @ $\text{£}4/\text{unit}$ | 250 @ $\text{£}10/\text{unit}$ |
| Assembly costs                            | 1000 @ $\text{£}2/\text{unit}$ | 250 @ $\text{£}4/\text{unit}$  |
| Packaging                                 | 200 @ $\text{£}10/\text{unit}$ | 50 @ $\text{£}10/\text{unit}$  |
| Despatch/delivery                         | 200 @ $\text{£}20/\text{unit}$ | 50 @ $\text{£}20/\text{unit}$  |

Determine the cost of each product using activity-based costing.

## Engineering business functions

Within an engineering company there are a number of discrete business functions. These include design, manufacturing, and engineering services. We shall briefly examine each of these essential functions and the effect that they have on costs.

### Design

By definition, an engineered product cannot be manufactured until it has been designed. Design is thus an essential engineering function. Design is itself a complex activity requiring inputs from a team of people with differing, but complementary, skills. To be effective, this team needs to undertake a variety of activities as part of the design process. These activities include liaison with clients and customers, concept design, specification, layout and detail design, and liaison with those responsible

for manufacturing, sales, service, and customer support. With most engineering projects, design costs may be significant; furthermore, these costs are normally incurred *before* manufacturing starts and income (attributable to the product or service being designed) is received.

## Manufacturing

Manufacturing involves having the right components and materials available and being able to apply appropriate processes to them in order to produce the end product. In this context, 'right' must not only be taken to mean appropriate in terms of the design specification but also the most cost-effective solution in every case. Costs of manufacturing are appreciable. These costs can be attributed to a number of sources including material and component costs, and the *added value* inherent in the manufacturing process resulting from labour, energy, and other overheads. Later we shall examine this in greater detail.

## Engineering services

Engineering services can be described as any engineering activity that is not directly concerned with manufacturing. Thus maintenance, sales, and customer support can all be described as *engineering services*. These functions may also represent significant costs which normally have to be recovered from manufacturing income.

## Measures and evaluation

An engineering company will normally employ a number of different control methods to ensure that its operation is profitable. These control methods include making forecasts of overall profitability, determining the contribution made by each individual activity towards overheads and fixed costs, and performing 'what-if' analysis to determine the effects of variations in cost and selling price. We shall start by describing the most simple method, *break-even analysis*.

### Break-even charts

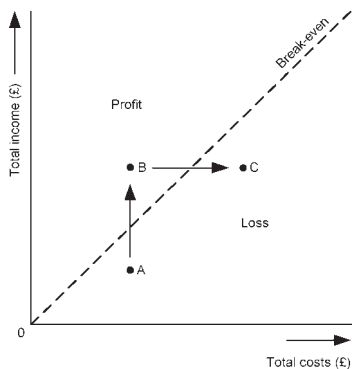


**Figure 1.2.3** Total income plotted against total costs showing profit and loss regions

Break-even charts provide a simple (and relatively unsophisticated) method for determining the minimum level of sales that a company must achieve in order for the business to be profitable. Consider the simple relationship illustrated in Figure 1.2.3. Here total income has been plotted against total costs using the same scale for each axis. At point A, total costs exceed total income and the operation is not profitable, that is, it makes a *loss*. If we charge more for the product, whilst keeping the costs fixed, we would move from point A to B. At a certain point, total income exceeds total costs and we move into *profit*. Finally, let us assume that our total costs increase whilst the total income from sales remains unchanged. We would then move from profit (point B) to loss (point C).

The *break-even point* is the volume of sales at which the operation becomes profitable and it marks the transition from loss into profit. A *break-even chart* takes the form of a graph of costs plotted against volume of product sold. At this point, it is important to recall that the total

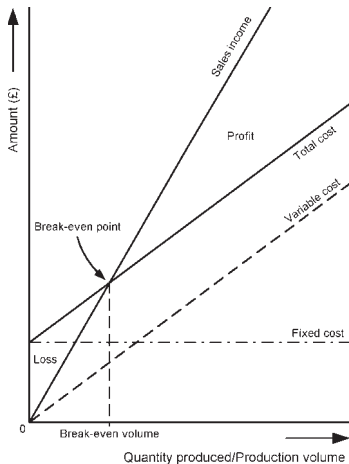


**Figure 1.2.4** Fixed and variable costs plotted against production volume showing break-even

costs of the business operation are the sum of the fixed and overhead costs with the variable costs of production. Thus:

$$\text{Total cost} = \text{fixed cost} + \text{overhead cost} + \text{variable cost}.$$

The income derived from the sale of the product (assuming a constant pricing structure) will simply be the product of the quantity sold (i.e. the volume of product sold) and the price at which it is sold (i.e. the per unit selling price). This relationship (a straight line) can be superimposed on the break-even chart and the point of intersection with the total cost line can be identified. This is the *break-even point* and the corresponding production volume can be determined from the horizontal axis (see Figure 1.2.4).

The *break-even quantity* can be determined from:

$$\text{Break-even quantity} = \frac{\text{fixed cost}}{\text{selling price} - \text{variable cost}}.$$

(Note that, in the above formula, *selling price* and *variable cost* are per unit.)

It is also possible to use the break-even chart to determine the profit that would result from a particular production quantity.

*Profit* can be determined from:

$$\text{Profit} = (\text{selling price} \times \text{quantity sold}) - (\text{fixed cost} + (\text{variable cost} \times \text{quantity sold})).$$

(Note that, in the above formula, *selling price* and *variable cost* are per unit.)

### Example 1.2.4

Centralux has analysed its fixed and overhead costs which together amount to £250 000 whilst the variable costs of its domestic central heating controller amount to £15 per unit manufactured.

Construct a break-even chart and use this to determine:

- (a) The break-even production volume when the controller is sold at:
  - (i) £25 per unit;
  - (ii) £30 per unit;
  - (iii) £35 per unit.
- (b) The profit for a production quantity of 25 000 units if the selling price is £35 per unit.

The break-even chart for the Centralux business operation is shown in Figure 1.2.5.

- (a) From Figure 1.2.5 the break-even points are:
  - (i) 25 000 units;
  - (ii) 16 667 units;
  - (iii) 12 500 units.
- (b) Also from Figure 1.2.5 the profit based on 25 000 units at a selling price of £35 per unit is £250 000.

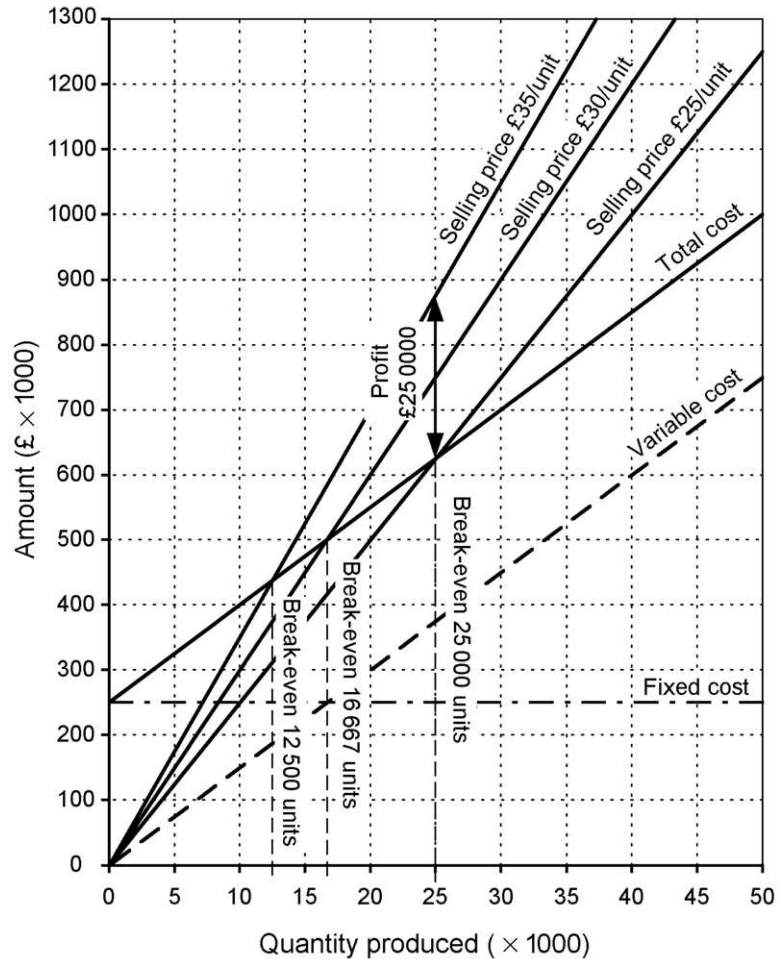


**Figure 1.2.5** Break-even chart for Centralux's domestic central heating controller

### Example 1.2.5

A small manufacturing company, KarKare, manufactures a trolley jack for the DIY motor enthusiast. The fixed cost of the company's jack manufacturing operation is £175 000 and the variable costs of producing the trolley jack amount to £21 per unit.

If the jack is to be sold for £45, determine the break-even quantity and the profit that would be returned from sales of 5000 units.

$$\text{Now Break-even quantity} = \frac{\text{fixed cost}}{\text{selling price} - \text{variable cost}}$$

$$\text{Thus break-even quantity} = £175\,000 / (£45 - £21) = 7292.$$

The profit based on sales of 10 000 units will be given by:

$$\text{Profit} = (\text{selling price} \times \text{quantity sold}) - (\text{fixed cost} + (\text{variable cost} \times \text{quantity sold}))$$

$$\begin{aligned} \text{Profit} &= (£45 \times 10\,000) - (£175\,000 + (£21 \times 10\,000)) \\ &= £450\,000 - £385\,000. \end{aligned}$$

Thus, the profit on 10 000 units will be £65 000.

It is important to realise that simple break-even analysis has a number of serious shortcomings. These may be summarised as follows:

- The sales income line (i.e. the product of the volume produced and its selling price) takes no account of the effect of price on the volume of sales. This is important as it is likely that the demand for the product will fall progressively as the selling price increases and the product becomes less competitive in the open market.
- The assumption that fixed costs remain fixed and variable costs increase linearly with production are somewhat dangerous. The reality is that both of these will change!

For the foregoing reasons it is important to regard break-even analysis as a 'rule-of-thumb' method for evaluating product pricing. Before making any business decisions relating to pricing and targets for production volume it is important to undertake further research into effect of pricing on potential sales as well as the pricing of competitive products.

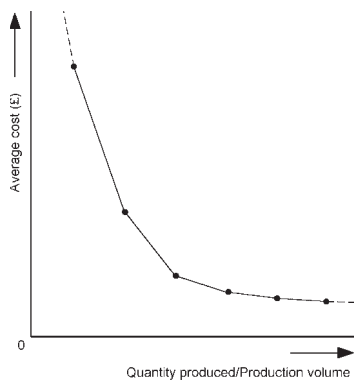


**Figure 1.2.6** Average cost plotted against production volume

### Average cost

The fixed costs associated with production have to be shared between the entire volume produced. Hence, a proportion of the final cost of a product will be attributable to the fixed costs of manufacture. The larger the quantity produced, the smaller this proportion will be. In other words, the *average cost* of the product will fall as the volume increases. We can illustrate this in the form of a graph (see Figure 1.2.6). Note that:

$$\text{Average cost} = \frac{\text{total cost}}{\text{quantity produced}}$$

### Example 1.2.6

Determine the average costs of KarKare's trolley jack based on fixed costs of £175 000 and variable costs of £21 per unit for production levels of 5000, 10 000, 15 000, and 20 000 units.

$$\text{Now Average cost} = \frac{\text{total cost}}{\text{quantity produced}}$$

and

$$\text{Total cost} = \text{fixed cost} + \text{variable cost}.$$

Thus:

| Quantity produced       | 5000      | 10 000      | 15 000       | 20 000       |
|-------------------------|-----------|-------------|--------------|--------------|
| Fixed cost (£)          | 175 000   | 175 000     | 175 000      | 175 000      |
| Variable cost (£)       | 105 000   | 210 000     | 315 000      | 420 000      |
| Total cost (£)          | 280 000   | 385 000     | 490 000      | 595 000      |
| <b>Average cost (£)</b> | <b>56</b> | <b>38.5</b> | <b>32.67</b> | <b>29.75</b> |

### Marginal cost

Once we have established a particular volume of production, the cost of producing one more unit is referred to as the *marginal cost*. The marginal cost of a product is that cost of the unit that results only from changes in

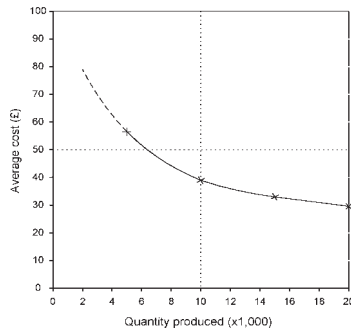


**Figure 1.2.7** Average cost plotted against production volume for KarKare's trolley jack

those costs that do not vary with the amount produced. Marginal cost is not the same as average cost – the reason for knowing the marginal cost of a product is that it can help us decide whether or not to increase production from an existing level. This is best illustrated with the use of an example (Figure 1.2.7).

### Example 1.2.7

KarKare has established a production level of 12 000 units for its trolley jack. As before, the fixed cost of the company's jack manufacturing operation is £175 000 and the variable costs of producing the trolley jack amount to £21 per unit. The jack is normally sold for £45 but a large high-street chain store has offered to take an additional 2000 units at a non-negotiable price of £30 per unit. Use marginal costing to determine whether this proposition is financially sound.

The total cost associated with a production volume of 12 000 units is found from:

Total cost = fixed cost + variable cost

$$\begin{aligned}\text{Total cost} &= £175\,000 + (12\,000 \times £21) \\ &= £175\,000 + £252\,000 = £427\,000.\end{aligned}$$

Now the average cost (based on 12 000 units) will be given by:

$$\begin{aligned}\text{Average cost} &= \frac{\text{total cost}}{\text{quantity produced}} \\ &= £427\,000 / 12\,000 = £35.58.\end{aligned}$$

Based on an average cost of £35.58, a selling price of £30 per unit does not appear to be sound business sense. However, if we consider the marginal cost of the trolley jack based on an existing production level of 12 000 units, we arrive at a different view. The rationale is as follows:

Let us assume a scenario in which we sell 12 000 trolley jacks at £45 and 2000 trolley jacks at £30. The total income produced will be given by:

$$\begin{aligned}\text{Total income} &= (12\,000 \times £45) + (2000 \times £30) \\ &= £540\,000 + £60\,000 = £600\,000.\end{aligned}$$

The total cost associated with producing 14 000 trolley jacks will be:

$$\begin{aligned}\text{Total cost} &= £175\,000 + (14\,000 \times £21) \\ &= £175\,000 + £294\,000 = £469\,000.\end{aligned}$$

Thus the resulting profit will be given by:

$$\begin{aligned}\text{Profit} &= \text{total income} - \text{total cost} \\ &= £600\,000 - £469\,000 = £131\,000.\end{aligned}$$

Had we decided not to accept the order for the extra 2000 units, we would have generated a profit given by:

$$\begin{aligned}\text{Profit} &= \text{total income} - \text{total cost} \\ &= £540\,000 - £427\,000 = £113\,000.\end{aligned}$$

### Another view

In marginal costing, we consider the cost of a product when all of the variable costs are removed. We arrive at this figure by calculating the cost of producing just one more unit – the difference in the cost of this unit and the previously manufactured one is the variable cost attributable to just one unit (i.e. the variable cost *per unit*). This assumes that the variable cost per unit is the same for all volumes of production output. This will usually be true for significant production volumes (note how the average cost tends towards a fixed value as the quantity increases in Figure 1.2.6).

Thus, meeting the order for an additional 2000 units at £30 has helped to increase our profits by £18 000. The important thing to note here is that, although the selling price of £30 per unit is less than the average cost per unit of £35.58, it is actually greater than the marginal cost of £21!

### **Profitability**

Profit, or *return on capital employed* (ROCE), is the expressed or implied goal for every business. Being able to make a realistic forecast of profits is an essential prerequisite to making a financial case for investment. It is also an essential ingredient in any business plan.

The need to maximise profits should be an important factor in decision-making. Traditional theory assumes that a company will invest in the most profitable projects first, and then choose projects of descending profitability, until the return on the last project just covers the funding of that project (this occurs when the *marginal revenue* is equal to the *marginal cost*).

The process of choosing projects is, however, much more complex. It may, for example, involve strategic issues (such as the need to maintain a presence in a particular market or the need to developing expertise in a particular technology with the aim of improving profits at some later date). Furthermore, many companies do not have sufficient funds available to reach the marginal position. Instead, they will rely on one or two 'hurdle' rates of return for projects. Projects that do not reach these rates of return will be abandoned in favour of those that are considered 'profitable'.

### **Contribution analysis**

In marginal costing, the excess of sales revenue over variable cost is known as the *contribution margin*. This margin represents the contribution made by the item in question to the fixed costs and profit. In Figure 1.2.4 the contribution margin is equivalent to the distance from the variable cost curve and the sales income line. We can thus say that:

$$\text{Contribution} = \text{sales revenue} - \text{variable costs}.$$

But since:

$$\text{Sales revenue} = \text{fixed costs} + \text{variable costs} + \text{profit}.$$

Thus:

$$\text{Contribution} = \text{fixed costs} + \text{profit}.$$

Contribution margin can be easily calculated. For example, if the selling price of a Centralux central heating controller (see Example 1.2.4) is £25 and its variable costs amount to £15 then the contribution margin is  $£(25 - 15) = £10$ .

## **1.3 FINANCIAL PLANNING AND CONTROL**

By now, you should have begun to understand how costing systems and techniques are applied in a typical engineering company. This next topic introduces you to the financial planning process. To satisfy the requirements of this unit you need to be able to describe the factors influencing

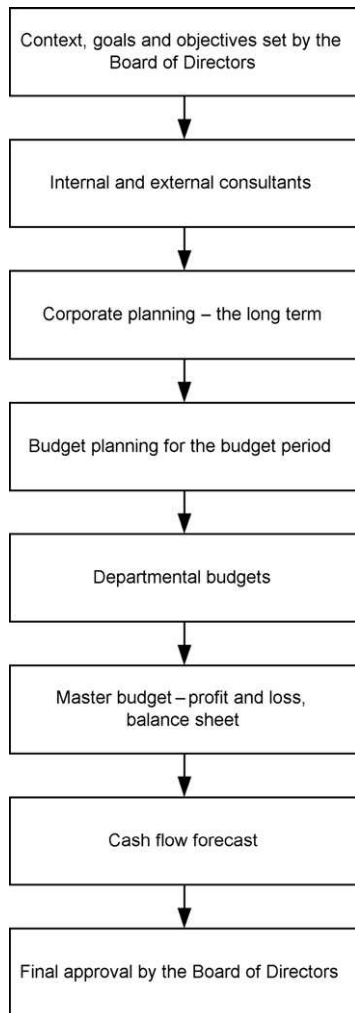


**Figure 1.3.1** *Process of financial planning and control*

the decision-making process during financial planning, examine the budgetary planning process and its application to financial planning decisions. You also need to be able to apply the standard costing techniques that we met in the previous section, analysing deviations from the planned outcome.

## Financial planning and control

Adequate financial planning is essential if a business is to achieve its objectives and profit targets. The basic procedure required to formulate a financial plan is as follows:

- (a) Formulate company policy, profit targets, and long-term plans.
- (b) Prepare forecasts for sales, production, stocks, costs, capital expenditure, and cash.
- (c) Compile these separate forecasts into a master forecast.
- (d) Consider all the alternatives available and select the plan which gives the best results, for example, in terms of profit and long-term financial stability.
- (e) Review limiting factors and the principal budget factor. This process takes place concurrently with (d) and enables work to begin on the framing of the budgets in (f).
- (f) Prepare individual budgets and finally the master budget which includes a forecasted profit and loss account, and balance sheet.

This process is illustrated in Figure 1.3.1.

## Budgetary control

The starting point of budgetary control is with the Board of Directors who determine the scale and nature of the activities of the company. This policy and objective setting is done within the constraints which exist at the time. For example, plans may have to be made within the current capacities and capabilities of the company, since making changes to the location of operations, the size and composition of the workforce and the product range are usually long-term matters. The budget is essentially for the short term, usually for 1 year, and created within the framework of long-term corporate planning. Successive budgets will be influenced by the preparation of long-term plans, but will always relate to the current period.

Some organisations prepare outline budgets over much longer periods perhaps for a 5–10-year horizon, but such budgets are really part of the long-term corporate planning activity and subject to major revision before being used as a basis for current period budgetary planning.

External factors will exercise considerable effects on the company in preparing its forecasts and budgets. Government policy, the proximity of a general election, taxation, inflation, world economic conditions, and technological development will all combine to constrain or influence the budget planning process. Once the Board of Directors has settled on a policy within the prevailing situation, then the process of turning the policy into detailed quantitative statements can begin.

We normally assume a budget period of 1 year, which is usual for most industries. It is therefore recognised that the budget period is fixed in



relation to the needs of the organisation concerned and could be any period ranging from 3 months to 5 years. The shorter the period the more accurate the forecasts, and that is why most companies find that an annual budgeting procedure is a satisfactory compromise.

### ***Preparing a business plan***

On occasions, it is necessary to provide a detailed business plan in order to make a case for a particular business venture or project. Before a business plan is written, it is necessary to:

- clearly define the target audience for the business plan;
- determine the plan's requirements in relation to the contents and levels of detail;
- map out the plan's structure (contents);
- decide on the likely length of the plan;
- identify all the main issues to be addressed (including the financial aspects).

Shortcomings in the concept and gaps in supporting evidence and proposals need to be identified. This will facilitate an assessment of research to be undertaken before any drafting commences. It is also important to bear in mind that a business plan should be the end result of a careful and extensive R&D project which must be completed before any serious writing is started.

A typical business plan comprises the following main elements:

- An *introduction* which sets out the background and structure of the plan.
- A *summary* consisting of a few pages which highlight the main issues and proposals.
- A *main body* containing sections or chapters divided into numbered sections and subsections. The main body should include financial information (including a *profitability forecast*).
- *Market and sales projections* should be supported by valid market research. It is particularly important to ensure that there is a direct relationship between market analysis, sales forecasts, and financial projections. It may also be important to make an assessment of competitors' positions and their possible response to the appearance of a rival product.
- *Appendices* should be used for additional information, tabulated data, and other background material. These (and their sources) should be clearly referenced in the text.

The financial section of the plan is of crucial importance and, since it is likely to be read in some detail, it needs to be realistic about sales expectations, profit margins, and funding requirements ensuring that financial ratios are in line with industry norms. It is also essential to make realistic estimates of the cost and time required for product development, market entry, and the need to secure external sources of funding.

When preparing a plan it is often useful to include a number of 'what-if' scenarios. These can help you to plan for the effects of escalating costs, reduction in sales, or essential resources becoming scarce. During a *what-if analysis*, you may also wish to consider the *halve-double*

scenario in which you examine the financial viability of the project in the event that sales projections are halved, and costs and time are doubled. The results can be sobering!

When writing a business plan it is necessary to:

- avoid unnecessary jargon;
- economise on words;
- use short crisp sentences and bullet points;
- check spelling, punctuation, and grammar;
- concentrate on relevant and significant issues;
- break the text into numbered paragraphs, sections, etc.;
- relegate detail to appendices;
- provide a contents page and number the pages;
- write the summary last.

Finally, it can be useful to ask a consultant or other qualified outsider to review your plan in draft form and be prepared to adjust the plan in the light of comments obtained and experiences gained.

## **Budgetary planning**

---

Budgets are used as a means of achieving planning and control objectives in most businesses and in many non-commercial organisations. A budget has been defined as:

*A financial or quantitative statement prepared and approved, prior to a defined period of time, of the policy to be pursued during that period for the purpose of attaining given objectives.*

The benefits that derive from budgetary control arise from the ability to co-ordinate policy, plans, and action and to be able to monitor the financial consequences of carrying out the plans.

An engineering company will prepare a number of budgets, each corresponding with a particular functional area. Each budget will normally be controlled by a specified manager, although some managers may control several budgets according to the particular management organisation employed within the company. In a typical engineering business you will find the following:

- marketing budget;
- manufacturing budget;
- R&D budget;
- administration budget;
- capital expenditure budget;
- cash budget.

Each of these budgets may be subdivided into further budgets. For example, the manufacturing budget may be subdivided into a budget for direct materials, a budget for direct labour, and a budget for factory overheads (heating, lighting, and other energy costs).

Each functional manager will forecast his/her own budget however, there is a need for managers and departments to co-ordinate their budget activities. For example, the capital expenditure budget may reflect the purchase of major items of capital equipment (such as a fork-lift truck or an overhead crane) that will be shared by several departments.

### Budget centres

The concept of *budget centres* (we have used this term in preference to *cost centre* or *profit centre*) is central to the process of budgetary control because it provides the means by which it is possible to identify and control costs at the point at which they are incurred. The number and selection of budget centres varies according to the size and complexity of the company. The following are typical:

- sales and marketing,
- production,
- personnel,
- R&D,
- finance and administration.

#### Example 1.3.1

The Sales Manager at KarKare has been asked to produce a quarterly budget forecast for next year's sales. In order to set about preparing the budget, he has discussed this with the Production Manager and they have agreed the following production volumes and likely levels of sales in each quarter of the year (note that Battery Charger production is not due to start until the third quarter).

| <i>Product</i>        | <i>Units</i> | <i>Selling price (£)</i> |
|-----------------------|--------------|--------------------------|
| <i>First quarter</i>  |              |                          |
| Battery charger       | Nil          | 20                       |
| Jack                  | 4000         | 30                       |
| Warning triangle      | 2500         | 8                        |
| <i>Second quarter</i> |              |                          |
| Battery charger       | Nil          | 20                       |
| Jack                  | 4000         | 30                       |
| Warning triangle      | 3000         | 8                        |
| <i>Third quarter</i>  |              |                          |
| Battery charger       | 1000         | 20                       |
| Jack                  | 3500         | 30                       |
| Warning triangle      | 2500         | 8                        |
| <i>Fourth quarter</i> |              |                          |
| Battery charger       | 2000         | 20                       |
| Jack                  | 3000         | 30                       |
| Warning triangle      | 2250         | 8                        |

Complete the quarterly sales budget showing the revenue by product line and total income on sales for each quarter. Include summary columns for the 12-month period. (Note that this task is greatly simplified by making use of a spreadsheet package!) Table 1.3.1 shows how the quarterly sales budget can be presented.

**Table 1.3.1** *KarKare sales budget 1999*

| Product line     | Quarter 1 |           |                | Quarter 2 |           |                | Quarter 3 |           |                | Quarter 4 |           |                | 12 months |                |
|------------------|-----------|-----------|----------------|-----------|-----------|----------------|-----------|-----------|----------------|-----------|-----------|----------------|-----------|----------------|
|                  | Units     | Price (£) | Revenue        | Units     | Price (£) | Revenue        | Units     | Price (£) | Revenue        | Units     | Price (£) | Revenue        | Units     | Revenue        |
| Battery charger  | 0         | 20        | 0              | 0         | 20        | 0              | 1000      | 20        | 20 000         | 2000      | 20        | 40 000         | 3000      | 60 000         |
| Trolley jack     | 4000      | 30        | 120 000        | 4000      | 30        | 120 000        | 3500      | 30        | 105 000        | 3000      | 30        | 90 000         | 14 500    | 435 000        |
| Warning triangle | 2500      | 8         | 20 000         | 3000      | 8         | 24 000         | 2500      | 8         | 20 000         | 2250      | 8         | 18 000         | 10 250    | 82 000         |
| <b>Total</b>     |           |           | <b>140 000</b> |           |           | <b>144 000</b> |           |           | <b>145 000</b> |           |           | <b>148 000</b> |           | <b>577 000</b> |

### Questions 1.3.1

Using the data and spreadsheet model of Table 1.3.1, determine the effect of each of the following 'what-if' scenarios:

- A 10% increase in all prices with effect from the start of the second quarter.
- Price increases of 15% with effect from the start of the third quarter for the battery charger and the jack (no price increase for the warning triangle).
- Production of all three items achieving only 50% of forecast in the fourth quarter.
- A 50p reduction in selling price of all three items with effect from the start of the third quarter.

### Budgets for materials

The direct materials costs of production depends upon knowing the quantity of materials and component parts purchased for use in the manufacture of an individual item. Once again, the process is most easily explained by the use of an example:

#### Example 1.3.2

Having agreed KarKare's quarterly sales budget, the Managing Director has asked the Production Manager to produce a direct materials budget for the year. The *bill of materials* for each item is as follows:

| Material/part          | Units | Price per unit |
|------------------------|-------|----------------|
| <i>Battery charger</i> |       |                |
| Transformer            | 6000  | 2.75           |
| Fuse panel             | 6000  | 0.95           |
| Indicator              | 6000  | 0.25           |
| Ammeter                | 6000  | 0.75           |
| Wire                   | 90    | 15.95          |
| Solder                 | 30    | 39.00          |
| Mains cable            | 60    | 25.50          |
| Mains plug             | 6000  | 0.45           |
| Plastic sheet          | 1500  | 15.50          |
| Crocodile clips        | 6000  | 0.20           |

(continued)

| <i>Material/part</i>    | <i>Units</i> | <i>Price per unit</i> |
|-------------------------|--------------|-----------------------|
| <i>Trolley jack</i>     |              |                       |
| Steel bar               | 1500         | 15.75                 |
| Steel rod               | 2500         | 21.00                 |
| Steel sheet             | 1750         | 19.75                 |
| Spray paint             | 1500         | 5.00                  |
| <i>Warning triangle</i> |              |                       |
| Coated aluminium strip  | 10 250       | 0.75                  |
| Plastic mouldings       | 30 750       | 0.95                  |
| Reflector panels        | 21 500       | 1.50                  |
| Fixings                 | 102 500      | 0.05                  |

Complete the direct materials budget showing the production value of each material/component part and the total material costs for each product. Also determine the KarKare's total direct materials cost. (Once again, this task is greatly simplified by making use of a spreadsheet package.) Table 1.3.2 shows how the direct materials budget can be presented.

**Table 1.3.2** KarKare direct materials budget 1999

| <i>Item</i>             | <i>Units</i> | <i>Price per unit</i> | <i>Value</i> | <i>Total</i>       |
|-------------------------|--------------|-----------------------|--------------|--------------------|
|                         | <i>(£)</i>   | <i>(£)</i>            |              |                    |
| <i>Battery charger</i>  |              |                       |              |                    |
| Transformer             | 3000         | 1.95                  | 5850.00      |                    |
| Fuse panel              | 3000         | 0.55                  | 1650.00      |                    |
| Indicator               | 3000         | 0.25                  | 750.00       |                    |
| Ammeter                 | 3000         | 0.65                  | 1950.00      |                    |
| Wire                    | 90           | 9.50                  | 855.00       |                    |
| Solder                  | 30           | 22.50                 | 675.00       |                    |
| Mains cable             | 60           | 19.90                 | 1194.00      |                    |
| Mains plug              | 3000         | 0.39                  | 1170.00      |                    |
| Plastic sheet           | 750          | 9.95                  | 7462.50      |                    |
| Crocodile clips         | 6000         | 0.15                  | 900.00       |                    |
|                         |              |                       |              | £22 456.50         |
| <i>Trolley jack</i>     |              |                       |              |                    |
| Steel bar               | 1 500        | 15.75                 | 23 625.00    |                    |
| Steel rod               | 2 500        | 21.00                 | 52 500.00    |                    |
| Steel sheet             | 900          | 19.75                 | 17 775.00    |                    |
| Spray paint             | 11 500       | 5.00                  | 7 500.00     |                    |
| Bearing                 | 29 000       | 1.75                  | 50 750.00    |                    |
|                         |              |                       |              | £152 150.00        |
| <i>Warning triangle</i> |              |                       |              |                    |
| Coated aluminium        | 10 250       | 0.65                  | 6662.50      |                    |
| Plastic mouldings       | 30 750       | 0.42                  | 12 915.00    |                    |
| Reflector panels        | 21 500       | 0.35                  | 7525.00      |                    |
| Fixings                 | 102 500      | 0.05                  | 5125.00      |                    |
|                         |              |                       |              | £32 227.50         |
| <b>Total</b>            |              |                       |              | <b>£206 834.00</b> |

### Questions 1.3.2

Using the data and spreadsheet model of Table 1.3.2, determine the effect of each of the following 'what-if' scenarios:

- The cost of steel increases by 10% over the entire period.
- The design of the battery charger is improved. The ammeter is eliminated and the fuse panel is replaced by a thermal trip costing 35p when purchased in quantities of 1000 or more.
- A cheaper source of plastic material is located. The new supplier offers a 15% discount on plastic sheet and 50% discount on moulded plastic parts.
- Finally, determine the effect of all three scenarios, (a), (b), and (c) applying at the same time.

### Direct labour budgets

The direct labour budget can be produced from a knowledge of the work time required to produce an item even though that time might be divided between several workers, each responsible for a different manufacturing operation or process. For each product we must determine an average time (i.e. the *standard time*) to manufacture a single unit. We can then multiply this figure by the mean wage rate for workers in that area.

If the manufacturing operation requires a variety of different skills and competency, wage rates may vary significantly (less skilled workers will command lower pay rates). In such cases, a more accurate determination of labour costs will require a more detailed analysis based on the time contribution made by workers on each different pay rate. Once again, the process is most easily explained by the use of an example:

### Example 1.3.3

KarKare's Production Manager has been asked to produce a labour budget for the next 12-month budget period. He has consulted the Personnel Manager on current wage rates and determined the standard work time required for each product. The results of these calculations are as follows:

|                         |                                                       |
|-------------------------|-------------------------------------------------------|
| <i>Battery charger</i>  | <i>Units</i>                                          |
| Standard hours/item     | 1.10                                                  |
| Wage rate in this area  | £4.75 (this product requires soldering/wiring skills) |
| <i>Trolley jack</i>     |                                                       |
| Standard hours/item     | 1.25                                                  |
| Wage rate in this area  | £6.25 (this product requires skilled operatives)      |
| <i>Warning triangle</i> |                                                       |
| Standard hours/item     | 0.25                                                  |
| Wage rate in this area  | £4.25 (this product uses unskilled labour)            |

**Table 1.3.3** *KarKare direct labour budget 1999*

| <i>Item</i>              | <i>Battery charger</i> | <i>Trolley jack</i> | <i>Warning triangle</i> |
|--------------------------|------------------------|---------------------|-------------------------|
| Total units              | 3000                   | 14 500              | 10 250                  |
| Standard hours/item      | 1.10                   | 1.25                | 0.25                    |
| Total hours              | 3300                   | 18 125              | 2562.5                  |
| Wage rate                | £4.75                  | £6.25               | £4.25                   |
| <b>Total labour cost</b> | <b>£15 675.00</b>      | <b>£113 281.25</b>  | <b>£10 890.63</b>       |

Using the production volumes agreed with the Sales Manager (see Table 1.3.1), determine the labour costs per product over the next 12 months. Also determine KarKare's total direct labour costs for the next 12 months.

Once again, a spreadsheet model is recommended, and Table 1.3.3 shows one way of presenting this.

### Questions 1.3.3

Using the data and spreadsheet model of Table 1.3.3, determine the effect of each of the following 'what-if' scenarios:

- Minimum wage rates for production workers are set at £4.50/hour for the entire period.
- The improvements to the design of the battery charger are instrumental in reducing its standard assembly time to 0.95 h.
- The design of the trolley jack is to be improved in order to increase the load that can be placed on it. This requires an additional 6 min of welding time.
- Finally, determine the effect of all three scenarios, (a), (b), and (c) applying at the same time.

### **Manufacturing overhead budgets**

We mentioned earlier that manufacturing overheads are incurred regardless of the volume of production. This is a rather simplistic view. Some overhead costs can truly be regarded as *fixed*, others can more correctly be referred to as *semi-variable*.

Semi-variable costs are those that have both a fixed element (this cost will be incurred regardless of production volume) and a variable element (this cost will increase as the production volume increases). A good example of a semi-variable cost is the supply of electrical energy. Some electrical energy will be required to support non-production activities (e.g. heating, lighting, office equipment, security equipment, etc.). Electrical energy will also be required to supply machine tools, soldering plants, and many other manufacturing processes (including standard factory plants such as conveyors, lifts, hoists, etc.). This use of electrical energy will clearly increase according to the level of production (note

that activity-based costing would actually apportion these costs to individual products). Let us take a further example to show how this works:

### Example 1.3.4

KarKare's Production Manager has been asked to produce a manufacturing overhead budget for the next 12-month budget period. He has established the fixed costs attributable to business rates, rent, and salaries (including his own):

|                            |                                                       |
|----------------------------|-------------------------------------------------------|
| <i>Fixed costs</i>         | £                                                     |
| Business rates             | 6500                                                  |
| Rent                       | 12 000                                                |
| Salaries                   | 68 000 (including costs of employment)                |
| Electricity                | 500                                                   |
| Water                      | 450                                                   |
| Gas                        | 650                                                   |
| <i>Semi-variable costs</i> |                                                       |
| Electricity                | 15p/hour for all three products                       |
| Water                      | 5p/hour for the battery charger and trolley jack only |
| Gas                        | 20p/hour for the trolley jack only                    |

Using the production volumes agreed with the Sales Manager (see Table 1.3.1) and the estimated production hours (see Table 1.3.3), determine the total fixed overhead and the semi-variable overhead for each product. Finally, determine the total manufacturing overhead for the next 12 months.

A spreadsheet model is again recommended and Table 1.3.4 shows one way of presenting this.

**Table 1.3.4** KarKare manufacturing overhead budget 1999

|                        | <i>Fixed</i> | <i>Var. per<br/>hour</i> | <i>Battery<br/>charger</i> | <i>Trolley<br/>jack</i> | <i>Warning<br/>triangle</i> |
|------------------------|--------------|--------------------------|----------------------------|-------------------------|-----------------------------|
| Total units            |              | 3000                     | 14 500                     | 10 250                  |                             |
| Standard<br>hours/item |              | 1.10                     | 1.25                       | 0.25                    |                             |
| Production<br>hours    |              | 3300                     | 18 125                     | 2562.5                  |                             |
| Semi-variable costs    |              |                          |                            |                         |                             |
| Electricity            | £500.00      | 0.15                     | £495.00                    | £2718.75                | £384.38                     |
| Water                  | £450.00      | 0.05                     | £165.00                    | £906.25                 | £0.00                       |
| Gas                    | £650.00      | 0.20                     | £0.00                      | £3625.00                | £0.00                       |
| Fixed costs            |              |                          |                            |                         |                             |
| Business rates         | £6500.00     |                          |                            |                         |                             |
| Rent                   | £12 000.00   |                          |                            |                         |                             |
| Salaries               | £68 000.00   |                          |                            |                         |                             |
| Total                  | £88 100.00   |                          | £660.00                    | £7250.00                | £384.38                     |
| <b>Total overhead</b>  |              |                          |                            |                         | <b>£96 394.38</b>           |



### Questions 1.3.4

Using the data and spreadsheet model of Table 1.3.4, determine the effect of each of the following 'what-if' scenarios:

- (a) A 5% increase in salaries for 6 months of the 12-month period.
- (b) A £2500 increase in business rates over the entire 12-month period.
- (c) A 30% increase in the cost of electricity over the entire 12-month period.
- (d) Finally, determine the effect of all three scenarios, (a), (b), and (c) applying at the same time.

### Question 1.3.5

Unfortunately, Karkare's Production Manager has not taken into account all of the fixed costs in his model (see Table 1.3.4). Suggest *one* additional semi-variable cost and *one* additional fixed cost that he should have taken into account.

### Cash-flow budgets

Cash-flow budgets are important to all businesses. Even a business that is highly profitable can get into serious difficulty if it has a shortage of cash. This may seem surprising but you must not forget that cash flows and profits may occur at quite different times. In the final analysis, if a company cannot pay employees and creditors when payments are due, it may fail, regardless of whether profits may be significant at some point in the future.

Cash budgets normally start from the opening cash balance and the flow of expected revenues from sales. However, as cash flows occur later than the sales which generate them, the budget has to account this time phasing. We will illustrate this with a simple example.

EzBild's cash-flow budget is shown in Table 1.3.5. This budget shows income from sales over the 6-month period, from January to June.

Take a good look at Table 1.3.5. There are a number of important things to notice:

- (1) The volume of sales each month is shown in the top row. The highest volume of sales is achieved in March (for EzBild this is the start of the highest season for sales). December is a quiet month for sales (perhaps nobody wants a ladder or a platform as a Christmas present!).
- (2) The balance *carried forward* (c/f) in the bottom line is *brought forward* (b/f) into the income stream for the next month.
- (3) Negative sums are shown enclosed in brackets. Note how March is a particularly poor month – production is high (thus outgoings are significant) but income from sales has not yet been received!

Temporary shortfalls of cash can be met in several ways. The company might liquidate some investments such as bank deposits, stocks and shares,

**Table 1.3.5** EzBild cash-flow budget (January to June 1999)

|                    | Dec    | Jan    | Feb    | Mar      | Apr    | May    | June   |
|--------------------|--------|--------|--------|----------|--------|--------|--------|
| Monthly sales      | 177    | 399    | 502    | 779      | 499    | 381    | 401    |
| <i>Income</i>      |        |        |        |          |        |        |        |
| Balance (b/f)      | 13 400 | 7915   | 2403   | (17 202) | 8786   | 19 139 |        |
| Receipts           | 17 700 | 39 900 | 50 205 | 77 012   | 49 933 | 38 381 |        |
|                    | 31 100 | 47 815 | 52 608 | 59 810   | 58 719 | 57 520 |        |
| <i>Expenditure</i> |        |        |        |          |        |        |        |
| Payments           | 23 185 | 45 412 | 69 810 | 51 204   | 39 580 | 31 224 |        |
| Balance (c/f)      | 13 400 | 7915   | 2403   | (17 202) | 8786   | 19 139 | 26 296 |

or it might arrange a bank overdraft. The production of a cash-flow forecast enables the company to do proper financial planning. Banks are happy to provide short-term overdrafts if they are planned, but are much less happy with businesses which fail to plan for these eventualities, for obvious reasons.

EzBild's cash-flow budget has been greatly simplified. In practice, much more detail is required. Receipts have been assumed to be for sales only. However, receipts may include interest on investments, disposals of fixed assets, etc. Similarly, payments will represent purchase of supplies, wages, purchase of equipment, overdraft interest, auditors' fees, payment of loan interest, or dividend paid to shareholders.

### Master budgets

The overall budget planning process shown in Figure 1.3.1 indicates the production of a master budget. This comprises a forecasted *profit and loss statement*, and a *balance sheet*. Budgets are often redrafted if the initial master budget they generate is unsatisfactory. It may not contain sufficient profit or some costs may be considered too high. This is an iterative process, whereby the information is refined and operating problems solved.

EzBild's master budget is shown in Table 1.3.6. This balance sheet summarises the financial position as at the end of the budget year. In effect, it is a snapshot of the position at 31 December 1999 (after the events of the year have taken place) whereas the profit and loss account is a period statement that provides us with a summary of the year's activities.

EzBild's balance sheet provides us with some important information about the company. In particular, it shows us where the money came from to run the business and what was done with the money obtained.

As before, there are a number of important things to be aware of. In particular, you should be able to see that:

- EzBild sold goods worth approximately £600 000 during the year. The *net profit* from these sales amounted to approximately £84 000.
- EzBild has a number of *fixed assets*. These include their office premises and various items of production equipment.
- EzBild's current assets amount to around £200 000. The company is financed by some share capital (valued at £140 000) and the accumulated profit and loss position.
- A loss of approximately £22 000 was made on the previous year's trading. This has been carried forward into the current year. Fortunately, this year's operations show a profit which adds to the available funds.

**Table 1.3.6** *EzBild master budget*  
*Profit and loss account – January to December 1999*

|                                             | £       | £       |
|---------------------------------------------|---------|---------|
| <i>Sale of goods</i>                        |         | 601 293 |
| <i>Less cost of goods sold:</i>             |         |         |
| Opening stock of finished goods             | 19 501  |         |
| Cost of finished goods                      | 315 249 |         |
|                                             | 334 750 |         |
| <i>Less closing stock of goods</i>          | 27 866  |         |
|                                             |         | 306 884 |
| <i>Gross profit</i>                         |         | 294 409 |
| Selling and distribution costs              | 88 730  |         |
| Administration and finance costs            | 76 211  |         |
| Design and development costs                | 45 629  |         |
|                                             |         | 210 570 |
| <i>Net profit</i>                           |         | 83 839  |
| <i>Balance sheet as at 31 December 1999</i> |         |         |
| <i>Assets employed:</i>                     |         |         |
| <i>Net fixed assets</i>                     |         |         |
| Premises                                    | 95 800  |         |
| Equipment                                   | 75 222  |         |
|                                             | 171 022 |         |
| <i>Current assets</i>                       |         |         |
| Stock                                       | 6950    |         |
| Debtors                                     | 28 565  |         |
|                                             | 12 771  |         |
| Cash                                        | 48 286  |         |
| <i>Less current liabilities</i>             |         |         |
| Creditors                                   | (9545)  |         |
| Overdraft                                   | (8000)  |         |
| <i>Net current assets</i>                   |         | 30 741  |
|                                             |         | 201 763 |

- The funds from shareholders and profits are the sole source of long-term funding for EzBild's business. The use of funds (or capital employed) is shown in the fixed assets and current assets totals.

Fixed assets are the long-term property of the business and current assets are the circulating assets of the business. The latter are so called because they constantly change through time, reflecting day-to-day business operations. The net current asset figure is an important figure because it shows what is left of current assets after current liabilities are met. Since current liabilities are sums of money owed to creditors (usually the company's suppliers) or to employees for wages, they have to be met out of current assets. Current assets are usually stock, debtors, and cash.

The net current assets figure is also called *working capital*. It must be enough to support the company's day-to-day business operations. If it is too small, the company may have difficulty in meeting its commitments.

## 1.4 PROJECT PLANNING AND SCHEDULING

A project can be defined as a series of activities with a definite beginning and ending, and with a series of actions that will lead to the achievement of a clearly defined goal. You need to be able to apply basic project planning and scheduling methods, including establishing timescales and resource requirements, and the relationships that exist between the various activities that make up a project.

Project planning is different from other forms of planning and scheduling simply because the set of activities that constitute a project are unique and occur only once. Production planning and scheduling relate to a set of activities that may be performed a large number of times.

Project planning involves considering the full set of activities necessary to achieve the project goal. Typical of these activities are:

- appointing consultants;
- appointing suppliers;
- forming a team to be responsible for carrying out the project;
- preparing a budget for the project;
- preparing a detailed costing;
- producing drawings;
- producing itemised parts lists;
- producing specifications;
- obtaining management approval;
- obtaining planning permission;
- scheduling the phases of the project.

*Project goals* must be clearly defined at the outset. If these are not clear, not properly understood or agreed by all members of the project team, then the chances of a successful outcome (or any outcome at all, for that matter) can be significantly reduced. In addition, you need to be able to establish the project resources and requirements, produce a plan with an appropriate timescale for completing the project, identify human resource needs, and identify approximate costs associated with each stage of the project.

## Programme evaluation and review technique

---

Programme evaluation and review technique (PERT) was developed for the US Navy in 1958 for planning and control of the Polaris nuclear submarine project. This project involved around 3000 contractors and the use of PERT was instrumental in reducing the project completion time by 2 years. PERT is widely used today as a fundamental project management tool both by governments and in industry.

PERT requires that project activities should be discrete and have definite start and end points. The technique provides most benefit when projects have a very large number of interrelated activities where it can be very effective in helping to identify the most effective sequence of activities from a variety of possibilities.

One important aspect of PERT is that it allows us to identify the path through the network for which the total activity times are the greatest. This is the *critical path*.

## Critical path method

---

Critical path method (CPM) is also widely used by both government departments and industry. CPM and PERT are very similar and the critical path is important for several reasons:

- (1) Since it represents the most time-critical set of activities, the total time to reach the project goal can *only* be reduced by reducing time spent in one or more of the activities along the critical path. In other

words, the critical path highlights those activities that should be critically reviewed to see whether they can be shortened in any way. Putting extra resources into one or more of the critical path activities can usually be instrumental in reducing the overall project time.

- (2) The critical path is unique for a particular set of activities and timings. If any of these are changed (e.g. by directing extra resources into them) a new critical path will be revealed. We can then apply PERT evaluation to this new critical path, critically reviewing the activities that it points us to. This process is iterative – in a large project we can continue to reduce the overall project time making changes as the project develops.
- (3) The critical path shows us where the most risky and potentially time-threatening activities occur. Since any problems or delays with activities on the critical path may jeopardise the entire project it is in our interests to focus particular attention on these tasks.

### Applying PERT

PERT is straightforward to apply. It comprises:

- (1) identifying all of the activities that make up the project;
- (2) identifying the sequence of the activities in step 1 (and, in particular, the order of precedence of these activities);
- (3) estimating the timing of the activities;
- (4) constructing a diagram that illustrates steps 1, 2, and 3;
- (5) evaluating the network and, in particular, identifying (and clearly marking) the *critical path*;
- (6) monitoring actual performance as the project is carried out against the schedule produced in 5, revising and re-evaluating the network as appropriate.

### Network diagrams

The network diagram used in PERT is comprised of a series of events which form the nodes in the network. Events are linked together by arrows which denote the activities, as shown in Figure 1.4.1.

Figure 1.4.1(a) shows two events, 0 and 1, linked by a single activity A. The expected time for the activity is 3 units of time (e.g. days, weeks, months, etc.).

Figure 1.4.1(b) shows three events, 0, 1, and 2 linked by two activities A and B. Activity A precedes activity B. The expected times for activities A and B are, respectively, 3 and 2 units of time.

Figure 1.4.1(c) shows four events, 0, 1, 2, and 3, linked by four activities, A, B, C, and D. In this network, activity A precedes activity C, whilst activity B precedes activity D. Event 3 is not reached until activities C and D have both been completed. The expected times for activities A, B, C, and D are, respectively, 2, 3, 3, and 2 units of time. There are several other things to note about this network:

- the events occur in the following order: 0, 1, 2, 3;
- activity B is performed at the same time as activity A;
- activity D is performed at the same time as activity C;
- the total time to reach event 3, whichever route is chosen, amounts to 5 units of time.

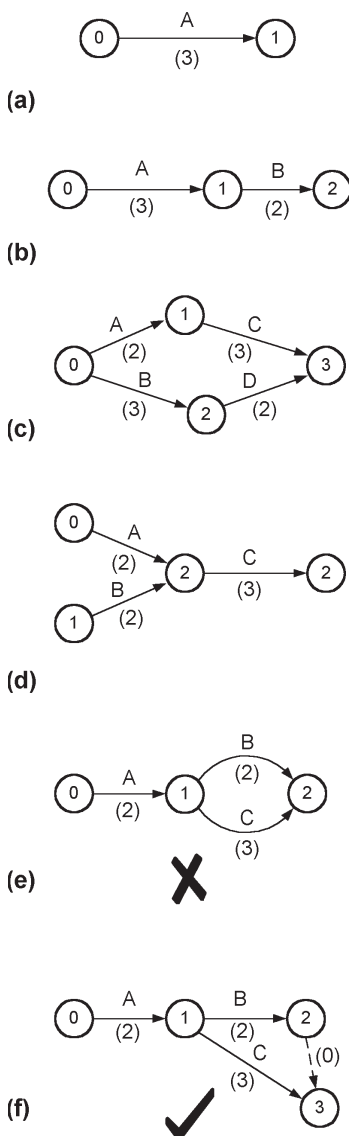


**Figure 1.4.1** Some simple network diagrams

Figure 1.4.1(d) also shows four events, 0, 1, 2, and 3, linked by three activities, A, B, and C. In this network, activities A and B both precede activity C (in other words, activity C cannot start until both activities A and B have been completed). The total time to reach event 3, whichever route is chosen, amounts to 5 units of time.

Within a network diagram, the activities that link two events must be unique. Consider Figure 1.4.1(e). This shows that event 2 can be reached via activities B and C (where the expected time for activity B is 2 units whilst the expected time for activity C is 3 units). To avoid potential confusion, we introduce a *dummy activity* between event 2 and event 3. This activity requires no time for completion and thus its expected time is 0 (note that we have adopted the convention that dummy activities are shown as a dashed line). Finally, you should see that event 2 is reached before event 3 and that the total expected time through the network amounts to 5 units and that there is *slack time* associated with activity B amounting to 1 unit of time (in other words, activity B can be performed up to 1 unit of time late without affecting the expected time through the network).

### Critical path

Within a network diagram, the *critical path* is the path that links the activities which have the greatest expected time. Consider Figure 1.4.2. This diagram shows six events linked by six real activities plus one dummy activity. The relationship between the activities and their expected times can be illustrated in the form of a table:

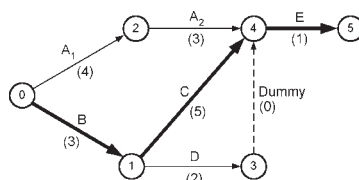
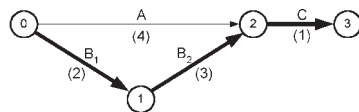
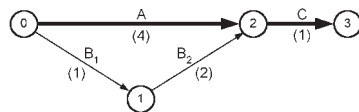


**Figure 1.4.2** Simple network diagram showing the critical path

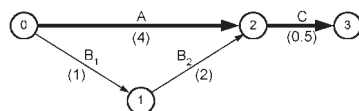
| Activity       | Preceding activity    | Expected time |
|----------------|-----------------------|---------------|
| A <sub>1</sub> | None                  | 4             |
| A <sub>2</sub> | A <sub>1</sub>        | 3             |
| B              | None                  | 3             |
| C              | B                     | 5             |
| D              | B                     | 2             |
| E              | A <sub>2</sub> , C, D | 1             |



(a)



(b)



(c)

**Figure 1.4.3** Using the network diagram to reduce overall project time

The critical path constitutes activities B, C, and E which produce a total expected time, between event 0 (the start of the project) and event 5 (the completion of the project) of 9 units.

The critical path allows us to identify those activities that are critical. By reducing the time spent on activities along the critical path we can reduce the expected time for the complete project. Next consider the network diagram shown in Figure 1.4.3.

Figure 1.4.3(a) shows a network diagram in which there are four events, 0–3, and four activities, A, B<sub>1</sub>, B<sub>2</sub>, and C. The critical path links together activities B<sub>1</sub>, B<sub>2</sub>, and C resulting in an expected completion time of 6 units. Let us assume that we can put some additional resources into activities B<sub>1</sub> and B<sub>2</sub>, and that this reduces the time spent on these tasks to 1 and 2 units, respectively. The critical path will move and it will now link activities A and C. The expected completion time will be reduced to 5 units. Our next task is to review tasks A and C to see if there is any way of reducing the time spent on these activities. Let us assume that we cannot reduce the time spent on activity A but we are able to reduce activity C by 0.5 units. This leaves the critical path unchanged but the expected time for completion will now be 4.5 units. This process is typical of that used on

larger much more complex projects. It is, however, important to note that there is a trade-off between time and cost. We shall examine this next.

### Project costs

Projects involve two types of cost: *indirect project costs* and *direct activity costs*. Indirect costs include items such as administrative overheads and facilities costs (heating, lighting, etc.). Direct costs are concerned with additional labour costs, equipment leasing, etc. We can spend extra money to reduce the time taken on the project; however, this only makes sense up to the point where further direct cost expenditure (such as the cost of employing additional contract staff) becomes equal to the savings in indirect project costs (heating, lighting, and other overheads).

To examine the trade-off between project time and costs, we need to have the following information:

- (1) A network diagram for the project showing expected times and indicating the initial critical path. We also need to know the minimum time for each activity when there are no resource constraints (this is known as the *crash time*).
- (2) Cost estimates for each project activity expressed in terms of indirect expenditure per unit time.
- (3) The costs of providing additional resources for each project activity and the consequent time saving expressed in terms of expenditure per unit time reduction.

With the above information we can reduce the critical path activity times, beginning with the activity that offers the least expenditure per unit time reduction. We can then continue with the second least costly, and continue until we are left with the most costly until either we reach the target minimum time for the project or the additional direct cost expenditure becomes equal to the savings in indirect costs.

Next we shall look at time analysis in a little more detail.

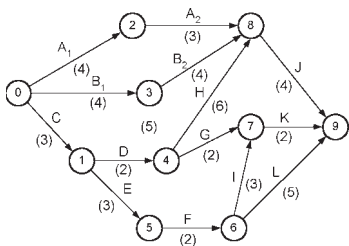


Figure 1.4.4 See Question 1.4.1

### Questions 1.4.1

- (1) The following data refers to the activities that make up a project. Use this information to construct a network diagram and identify the critical path.

| Activity       | Preceding activity              | Expected time |
|----------------|---------------------------------|---------------|
| A <sub>1</sub> | None                            | 1             |
| A <sub>2</sub> | A <sub>1</sub>                  | 2             |
| B <sub>1</sub> | None                            | 2             |
| B <sub>2</sub> | B <sub>1</sub>                  | 3             |
| C              | A <sub>2</sub> , B <sub>2</sub> | 1             |
| D              | C                               | 5             |
| E              | C                               | 3             |
| F              | D, E                            | 1             |

- (2) Determine the critical path through the network shown in Figure 1.4.4.
- (3) Determine the expected time to complete the project.

## Project time analysis using PERT

PERT defines a number of important times in the project life cycle. These are as follows:

|                        |       |                                                                                                                                                                                                                                                                                                                                                                                               |
|------------------------|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Expected time          | $t_e$ | The expected time for an activity is simply the average time for the activity.                                                                                                                                                                                                                                                                                                                |
| Optimistic time        | $t_o$ | The fastest time for the completion of the activity. This time will rarely be achieved and will only be bettered under exceptionally favourable circumstances.                                                                                                                                                                                                                                |
| Pessimistic time       | $t_p$ | The slowest time for the completion of the activity. This time will nearly always be bettered and will only be exceeded under exceptionally unfavourable circumstances.                                                                                                                                                                                                                       |
| Most likely time       | $t_m$ | This time represents the 'best guess' time for the completion of the activity. This time is the statistical mode of the distribution of the times for the activity.                                                                                                                                                                                                                           |
| Earliest expected time | $T_E$ | The earliest expected time for a particular event is the sum of all of the expected times ( $t_e$ ) that lead up to the event in question.                                                                                                                                                                                                                                                    |
| Latest allowable time  | $T_L$ | The latest allowable time for a particular event is the latest time that the event can take place yet still allow the project to be completed on time.                                                                                                                                                                                                                                        |
| Slack time             | $T_S$ | The slack time is the difference between the earliest expected time ( $T_E$ ) and the latest allowable time ( $T_L$ ). In effect, it is the amount of time that can elapse after completing one activity and starting another whilst still allowing the project to be completed on time. Note that, by definition, there is no slack time when following the critical path through a network. |

Estimates of project times are often based on previous experience of performing similar tasks and activities. The expected time,  $t_e$ , is often calculated from the formula:

$$t_e = \frac{t_o + 4t_m + t_p}{6}$$

whilst the variance of an activity time can be determined from:

$$\sigma^2 = \left( \frac{t_p - t_o}{6} \right)^2$$



Detailed time analysis using PERT can best be illustrated by taking an example:

**Example 1.4.1**

KarKare has engaged a consultant to advise on the construction of a new production facility for its trolley jack. The consultant has identified the following sequence of activities:

| Reference | Activity               | Preceding activity | Time estimate (weeks) |       |       |
|-----------|------------------------|--------------------|-----------------------|-------|-------|
|           |                        |                    | $t_o$                 | $t_m$ | $t_p$ |
| A         | Outline design         | None               | 4                     | 5     | 6     |
| B         | Prepare cost estimates | A                  | 2                     | 2     | 3     |
| C         | Client review          | B                  | 2                     | 3     | 4     |
| D         | Planning application   | C                  | 8                     | 12    | 14    |
| E         | Site investigation     | C                  | 2                     | 3     | 4     |
| F         | Detailed plans         | C                  | 4                     | 5     | 6     |
| G         | Building regulations   | F                  | 6                     | 7     | 8     |
| H         | Tendering              | F                  | 4                     | 5     | 6     |
| I         | Client approval        | H                  | 1                     | 2     | 4     |
| J         | Construction           | I                  | 30                    | 34    | 44    |
| K         | Fitting out            | J                  | 1                     | 2     | 3     |
| L         | Client handover        | K                  | 1                     | 2     | 3     |

- (i) Draw the network diagram for the project.
- (ii) Determine the expected time for each activity and mark this on the network diagram.
- (iii) Determine the critical path and mark this on the network diagram.
- (iv) Determine the earliest expected time and latest permissible time at each node of the network and use this to determine the slack time for each event.

Figure 1.4.5 shows the network diagram and the critical path for KarKare's new facility.

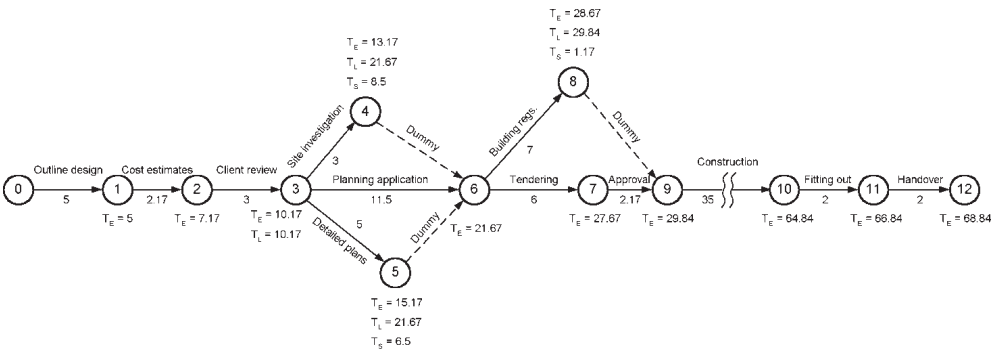


**Figure 1.4.5** See Example 1.4.1

### Questions 1.4.2

The Board of Directors at Centralux has decided to replace the company's ageing telephone system. The following events and activities have been identified as forming part of this project:

| Reference | Activity                           | Preceding activity | Time estimate (weeks) |       |       |
|-----------|------------------------------------|--------------------|-----------------------|-------|-------|
|           |                                    |                    | $t_o$                 | $t_m$ | $t_p$ |
| A         | Prepare a specification            | None               | 1                     | 1     | 2     |
| B         | Identify potential suppliers       | A                  | 1                     | 1     | 2     |
| C         | Plan new system location           | A                  | 2                     | 3     | 4     |
| D         | Tendering                          | B                  | 4                     | 5     | 6     |
| E         | Board approval                     | C, D               | 1                     | 1     | 2     |
| F         | New system built by supplier       | E                  | 5                     | 6     | 8     |
| G         | New system delivered and installed | F                  | 1                     | 2     | 3     |
| H         | New system commissioned            | G                  | 1                     | 2     | 3     |
| I         | Training                           | G                  | 1                     | 1     | 2     |
| J         | Handover to Centralux              | H, I               | 1                     | 1     | 2     |
| K         | Allocate new numbers               | E                  | 1                     | 2     | 4     |
| L         | Prepare new telephone directory    | K                  | 1                     | 2     | 4     |

- Draw the network diagram for the project.
- Determine the expected time for each activity and mark this on the network diagram.
- Determine the critical path and mark this in the network diagram.
- Determine the earliest expected time and latest permissible time at each node in the network and use this to determine the slack time for each event.

### Gantt charts

A Gantt chart is simply a bar chart that shows the relationship of activities over a period of time. When constructing a Gantt chart, activities are listed down the page whilst time runs along the horizontal axis. The standard symbols used to denote the start and end of activities, and the progress towards their completion, are shown in Figure 1.4.6.

A simple Gantt chart is shown in Figure 1.4.7. This chart depicts the relationship between activities A and F. The horizontal scale is marked off in intervals of 1 day, with the whole project completed by day 18. At the start of the 8th day (*see time now*) the following situation is evident:

- activity A has been completed;
- activity B has been partly completed but is running behind schedule by 2 days;
- activity C has been partly completed and is running ahead of schedule by 1 day;

| Symbol                                                                              | Meaning                        |
|-------------------------------------------------------------------------------------|--------------------------------|
| [                                                                                   | Start of an activity           |
| ]                                                                                   | End of an activity             |
| [ — ]                                                                               | Actual progress of an activity |
|  | (alternative representation)   |
| ∇                                                                                   | Time now                       |

**Figure 1.4.6** Symbols used in Gantt charts

- activity D is yet to start;
- activity E has started and is on schedule;
- activity F is yet to start.

Let us move onto a more realistic example. The Board of Directors at Centralux have decided to relocate their main office. The Gantt chart for this project is shown in Figure 1.4.8. This chart has been produced by a specialist project management software package. Note the following:

- the entire project is expected to be completed in a time period of 8 months;
- the chart identifies key personnel involved in the project;
- arrows have been added to indicate the flow of activities;

Figure 1.4.7 A simple Gantt chart

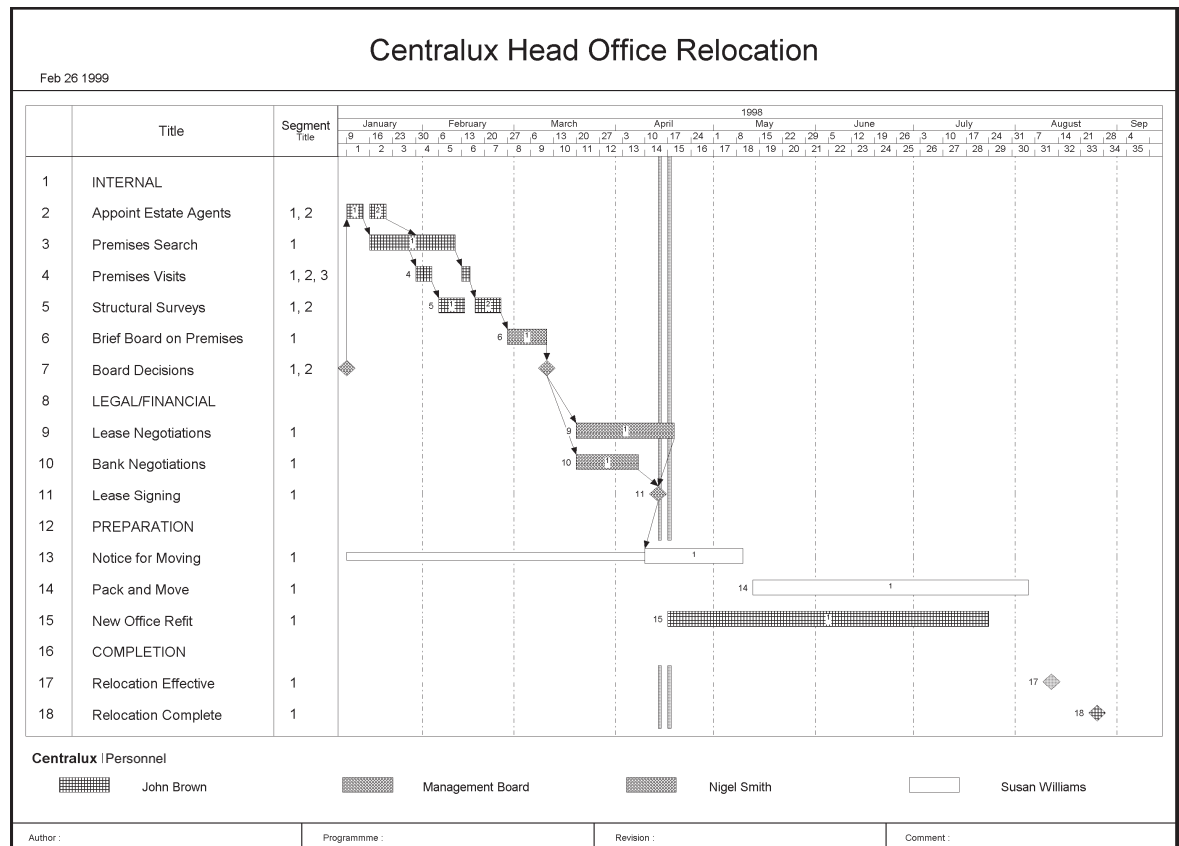
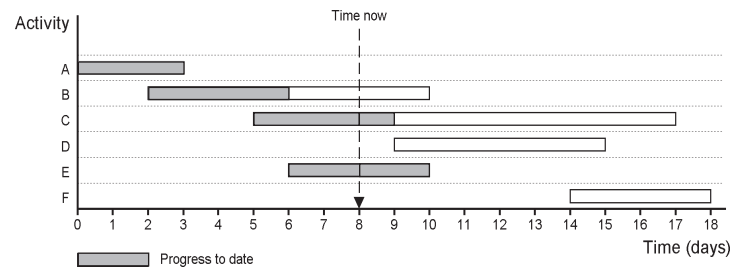


Figure 1.4.8 Gantt chart for Centralux's head office relocation

- the first part of project involved searches, visits, and surveys (equivalent to the design phase of a manufacturing project);
- the Board of Directors take the initial decision to engage in the project and later authorise the next (legal/financial) phase of the project which has major financial implications.

Questions 1.4.3

Centralux has decided to install a site radio communication system. The company has selected a supplier, Bizcom, who will carry out the work of building, installing, and commissioning the radio equipment. At the beginning of the 8th week of the project, Centralux’s Managing Director has asked Bizcom to supply a Gantt chart showing progress to date (see Figure 1.4.9). Use the Gantt chart to identify:

- (a) the number of weeks required to complete the project;
- (b) activities that have been completed;
- (c) activities that have been completed ahead of schedule;
- (d) activities that are behind schedule (and how far each is behind).

Figure 1.4.9 See Question 1.4.3

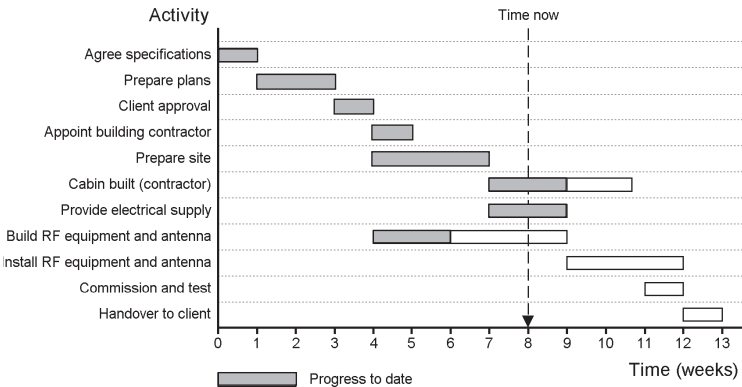
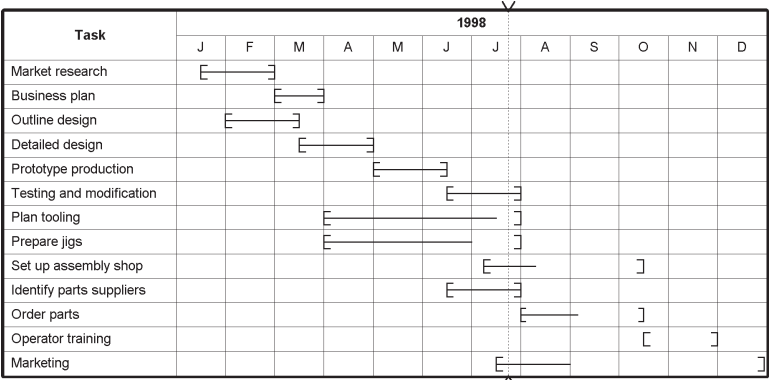


Figure 1.4.10 See Question 1.4.4



### Questions 1.4.4

KarKare is developing a car roof box mounted on a metal rack. This new product will go into production at the end of the current year (1998). The Gantt chart for this project is shown in Figure 1.4.10. Use the Gantt chart to identify:

- (a) the number of weeks required to complete the project;
- (b) activities that are due to be completed at the end of the current month (July);
- (c) activities that are ahead of schedule but not yet completed;
- (d) activities that are more than 2 weeks behind the schedule.

# 2

# Engineering design

## Summary

The aim of this unit is to give students an opportunity to experience the various phases of carrying out a design project, including the preparation of a detailed design specification and the production of a comprehensive design report. Computer technology is used extensively in engineering design and students are expected to demonstrate an ability to use appropriate hardware and software as part of the engineering design process.

## INTRODUCTION

Design itself is a complicated activity requiring individual inputs from a whole group of people, who go to make up the *design team*. The total design process requires the design team to consider activities that include: market forces, design management, design specification, concept design, layout design and detail design, materials technology, manufacturing methods, financial and legal aspects, publicity, and sales. The total design process requires us to first analyse the design requirements, and then, to synthesise the many diverse design parameters so that, apart from the simplest of design tasks, the whole design process requires the combined skills and knowledge of a design team.

In this chapter we will be concentrating on those aspects of engineering design, which enable us not only to produce engineering components, but also to consider some of the problems associated with the design of engineering systems. We will be very much concerned with looking at the overall design process and the production of a design report. The use of computers in engineering design will also be considered.

We start by examining the essential components of a *design specification*, since it is this document that forms the essential reference point for all the activities concerned with the engineering design process.

Having established the nature of the design specification, another essential task for engineers is to present their ideas to company management. It is no use achieving design perfection, if we are unable to persuade those in authority that our ideas are viable! To help us achieve this aim, we look at how to prepare and present the *design report*.

Computer technology has been an integral part of engineering design for many years. So our final topic is concerned with the use of computers in the overall design process and in particular, the use of computer packages for drafting, project scheduling, and mathematical analysis.

## 2.1 THE DESIGN SPECIFICATION

### Introduction

---

As an introduction to specification writing consider the following modified paragraphs taken from BS 7373 *Guide to the preparation of specifications*. Note that we are primarily concerned here with the product design specification (PDS), that is, the specification which conveys the designers description of the product.

*A specification is essentially a means of communicating the needs or intentions of one party to another. It may be a user's description, to a designer (a brief), detailing requirements for purpose or duty; or it may be a designer's detailed description to the operator, indicating manufacturing detail, materials, and manufacturing tolerances; or it may be a statement, by a sales person, describing fitness for purpose to fulfil the need of a user or possible user. It may, of course, be some or all of these in one.*

*The contents of the specification will, therefore, vary according to whether it is primarily from the using, designing, manufacturing, or selling aspect. Specifications will also vary according to the type of material, or component being considered, ranging from a brief specification for a simple component to a comprehensive specification for a complex assembly or engineering system. We are primarily concerned here, in preparing a design specification from the point of view of the designer attempting to meet customer needs.*

*The requirements of the specification should be written in terms of describing the optimum quality for the job, not necessarily the highest quality. It is usually unwise to over specify requirements beyond those for a known purpose. It is costly and restrictive to seek more refinements than those necessary for the function required. The aim, should therefore be, to produce a minimum statement of optimum quality in order not to increase costs unnecessarily; not to restrict processes of manufacture; and not to limit the use of possible materials.*

As already mentioned the design specification has to take into account parameters such as function, performance, cost, aesthetics, and production problems, all of these issues are concerned with *customer requirements* and this has to be the major consideration when producing a design specification. The major *design parameters* (such as layout, materials, erection methods, transportation, safety, manufacture, fabrication, and legal implications) must also be considered when producing the design specification. Finally, all *design information* must be extracted from appropriate sources, such as British Standards (BS) and International Standards, and all legislative requirements concerning processes, quality assurance, and the use of new technologies must be applied.

### Customer requirements

---

In order to provide a successful winning design, it is essential that the requirements of the customer are met whenever possible. Because we wish to produce a specification which expresses our customer requirements it is most important that we spend time in consultation with them, to

ensure that customer needs are well understood and, if necessary, to agree amendments or reach a compromise dependent on circumstances.

Remember that a specification is essentially a listing of all the parameters essential to the design. Then, generally, a customer will list essential values as part of his/her requirements, but each value must be examined before transfer to the specification. Thus the design specification must always be formulated by the designer.

To illustrate the customer requirements for a particular engineering design, and the role of the designer in interpreting such requirements, consider the following example:

### Example 2.1.1

A potential customer approaches you, as a designer, with a *brief* for the design of an electric drill. In order to secure the job, you need to produce a comprehensive specification, which takes into account all of your customer's requirements which are listed below:

#### **A.B. Brown Engineering** **Outline specification for electric drill**

- Performance:** Capable of taking drill bits up to 0.75 inch diameter.  
Operate from 240 V, 50 Hz power supply.  
Capable of two speed operation.  
Have hammer action.  
Operate continuously for long periods of time.  
Suitable for soft and hard drilling.  
Eccentricity of drilling action must be limited to 1.01 inches, for drill bits up to 12 inches in length.  
Have a minimum cable reach of 5 m.
- Environment:** Able to operate internally and externally, within a temperature range of  $-20$  to  $+40^{\circ}\text{C}$ .  
Have no adverse effects from dirt, dust, or ingress of oil or grease.  
Capable of operation in wet conditions.  
Capable of operation where combustible dusts are present.
- Maintenance:** Capable of being dismantled into component parts, for ease of maintenance.  
Require no special tools, for dismantling/assembly operations.  
Component parts to last a minimum of 2 years, before requiring replacement or rectification.
- Costs:** To cost a maximum of £60.00.
- Quantity:** 2000 required from first production run.
- Aesthetics/ergonomics:** Polymer body shell with two colour finish.  
Pistol grip lower body and upper body steady handle.  
Metal chuck assembly, with chrome finish.



Size/weight: Maximum weight of 3 kg.  
Overall length not to exceed 30 cm.

Safety: Complies with all relevant BS.

As designers we must ensure that the customer's requirements are unambiguous, complete, and attainable. This is where the dialogue with the customer begins! For the purpose of this example, let us re-visit each of the requirements.

The title for the specification needs clarification: is electric drill a suitable title? We know that A.B. Brown Engineering have specified information about type of grip and body design, so perhaps a more accurate description might be: *electric hand drill (mains operated)*, since cable length and supply conditions have been mentioned.

### Performance

If we consider the *performance* requirements, we note that in certain areas they are ambiguous and generally incomplete. Is the electric hand drill to be made available for export? The metric equivalent of 0.75 inches may be necessary, in any event it is a good marketing ploy to ensure that all dimensions are available in Imperial and SI units, this appeals to both the European and USA markets. Thus to ensure ease of production of the 'chuck' a metric equivalent should be given.

Are the supply details and power requirements for the drill sufficient? There is, for example, no indication of the power requirements for the drill motor, this must be given together with details of the supply, that is, alternating current. The rpm of the two speed operation must also be given, this will depend on service loading, time in use, and materials to be drilled.

Statements such as 'operate continuously for long periods of time' and 'suitable for soft and hard drilling' should be avoided. What periods of time? What types of materials are required to be drilled?

Although not directly obvious, the eccentricity requirements will have an effect on the quality of the gearing and the type of bearing required for the drive spindle. High-quality bearings are expensive and this will need to be taken into account when designing to a maximum cost.

### Environment

The criteria for the operating environment are quite clear, however, they do have quite serious consequences for the design. To be able to operate the drill in quite harsh external conditions, the insulation for the plug and cable assembly will need to meet stringent standards. Motor insulation and protection will be required to insure that sparking and arcing does not occur when the drill is being used in a combustible gas atmosphere

(refer to BS 4999, BS 5000, BS 5501, and BS 6467). Suitable motor caging will be required to ensure that the ingress of dirt, dust, oil, and grease do not adversely affect the performance of the drill motor.

### **Maintenance**

The number and nature of component parts need to be established, prior to any detail design being carried out. There are obvious cost implications if all component parts are to last a minimum of 2 years. This needs clarification to ascertain whether we are only talking about major mechanical parts or genuinely all parts. The likelihood of failure dependent on service use also needs to be carefully established to determine the feasibility of this requirement.

### **Costs**

The viability of this figure needs to be determined by taking into account the costs of component parts, tooling requirements, fabrication costs, machining costs based on required tolerances and materials finishing, and environmental protection costs.

### **Quantity**

This will determine the type of manufacturing process, and whether or not it is necessary to lay-on additional tooling, or buy-in standard parts, and concentrate only on assembly and test facilities, for the production run. Future component numbers will need to be established in order to make predictions about the most cost-effective production process.

### **Aesthetics/ergonomics**

Is the two colour finish absolutely necessary? This will depend on target market and results of consumer research, which will need to be known by the designer, prior to determining the unit cost. Chrome finishing is an expensive process and not altogether suitable for a drill chuck, which will be subject to harsh treatment in a hostile environment. Knocks, dents, scratches, and pollution in the work environment would quickly affect the chrome protective coating. Consideration needs to be given to alternative materials, which provide good corrosion protection and durability, as well as looking aesthetically pleasing.

### **Size/weight**

The weight and size criteria are not overly restrictive and allow the designer some room for manoeuvre. Light alloys and polymers may be used to help keep weight down, provided the performance criteria are not compromised.

**Safety**

There is a need to establish whether or not the product is intended for the European and/or World market. European legislation already has a major influence on safety standards, particularly relating to electrical goods. So ISO standards, European legislation, and other relevant quality standards would need to be followed. Is it, for instance, the intention of A.B. Brown Engineering to provide an electric hand drill which is capable of operation from the continental 220 V supply?

Having considered the customer's requirements and obtained answers to the questions posed, we need to ensure that all design parameters have been covered, prior to the production of the specification. For our example there is a need to include more design requirements on materials specification, component function, testing, prototype production, and timescales.

**Question 2.1.1**

Based on the information provided in Example 2.1.1, re-write the specification for A.B. Brown Engineering, avoiding ambiguous information and adding design requirements for materials specification, component function, and timescales. For the purpose of this exercise, avoid reference to BS or other sources of information. Consider only the argument offered in the example and information supplied in textbooks. You should assume that the electric hand drill has to be designed for the home and European market.

Since the design requirements, their associated design parameters, and the requirements of the customer are inextricably linked, we next look in more detail at the design requirements and parameters necessary to produce a comprehensive PDS. At this point, it is worth remembering that it may not always be necessary to include *all* requirements since these will be dependent on the complexity (or otherwise) of the engineering component or system.

**The PDS and design requirements****PDS**

The designer's description of the product is presented in the form of a PDS. The structure of the PDS is again mentioned in PD 6112 and this should be referred to for full details. A brief summary of what might be included, based on the information provided in PD 6112, is given below, this list is neither exclusive nor exhaustive, but it does act as a useful guide for those new to PDS writing:

- Title: This should provide an informative unambiguous description of the product.

- **Contents:** A list of contents, which acts as a useful introduction to the document for the reader.
- **Foreword:** This sets the scene and provides the reader with useful background information concerned with the project and the customer's brief.
- **Scope of specification:** This section provides the reader with details on the extent of the coverage and the limitations imposed on the information provided. It also gives details on the function of the product or system under consideration.
- **Consultation:** Information on any authorities who must be consulted concerning the product's design and use. These might include the health and safety executive (HSE), fire service, patent office, or other interested parties.
- **Design:** Main body of text detailing the design requirements/parameters for the product or system, such as performance, ergonomics, materials, manufacture, maintenance, safety, packaging, transportation, etc.
- **Appendices:** Containing definitions that might include: complex terminology, abbreviations, symbols, and units. Related information and references, such as statutory regulations, BS, ISO standards, design journals, codes of practice, etc.

Figure 2.1.1 shows a typical layout for the 'design' element of a PDS, detailing the customer requirements and the parameters which might affect these requirements.

Product / System : Hand held electric drill Sheet No: 2 of 8

Date : 1.1.98 Issue number : 03.1.98

| Requirements                                                              | Parameters                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Costs :</b></p> <p>Component selling price not to exceed £60.00</p> | <p><b>Competition :</b></p> <p>Provide details of existing prices for similar specification models that currently exist. Give best and worst case examples of competing models</p> <p><b>Target Market :</b></p> <p>Obtain details from marketing department on target audience for component sales. What are they willing to pay? What performance, quality and reliability do they expect?</p> <p><b>Costing :</b></p> <p>Explore ways for the company to minimise the following costs :</p> <ul style="list-style-type: none"> <li>- materials</li> <li>- Labour</li> <li>- overheads</li> <li>- manufacture</li> <li>- tooling</li> <li>- packaging and documentation</li> <li>- freight</li> </ul> |
| <p><b>Maintenance :</b></p> <p>etc</p>                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

Note that flexibility of selling price may be determined by considering parameters

Check you are working to latest issue

Note that parameters may not be exclusive to one requirement, but need to be considered in conjunction with others for example; costs, target market and performance.

One or more requirements may be entered on page, dependent on detail

Figure 2.1.1

### ***Design requirements***

Information on some of the major design requirements, in the form of a list, are detailed below. You should ensure that you understand the significance of each of these and, in particular, you are aware of the important parameters associated with each.

### ***Performance***

You will already be aware of some typical performance parameters, when studying Example 2.1.1. The performance specified by customers, must be attainable and clearly defined. The performance required by A.B. Brown Engineering in our electric hand drill example, shows how easy it is to produce ambiguous performance criteria and leave out essential information.

A balance needs to be established between performance and costs, ultimate performance demanded by a customer is likely to be prohibitively expensive. Designers must be aware of the economic viability of meeting performance requirements and, where these are not feasible, dialogue must be entered into to seek a more cost-effective solution.

Over-specification of performance is more likely to occur in specialist 'one-off' products where operational information is limited. Designers should draw on the knowledge and experience gained from similar designs, or seek the advice of more experienced colleagues. In any case, the urge to over design should be avoided, computer simulation, scale models, or the use of prototypes may be a way of establishing appropriate performance data, for very large one-off products, when trying to avoid over-specification.

### ***Ergonomics (or human factors)***

The word ergonomics originates from the Greek, *ergos* – work, and *nomics* – natural laws, thus ergonomics literally means '*the natural laws of work*'. It first came into prevalence during the Second World War, when aircraft pilots confused, for example, the landing gear lever with the flap lever, which on occasions, resulted in disaster. Prior to the Second World War, little interest had been paid, by engineers, to ergonomic design. We only have to look at the arrangement of controls and displays used in old steam engines, power stations, and heavy process equipment, to realise that little consideration had been given to the needs of the operator.

Ergonomic design has risen in prominence over the past decade and the needs of the user have taken on much more importance. Consideration should now be given to one or more of the following ergonomic design parameters: controls and displays; instruments and tools; workspace arrangement; safety aspects; anthropometrics (measurement of physical characteristics of humans, in particular human dimensions); environment – visual, acoustic, and thermal.

No matter how complex, sophisticated or ingenious the product, if human operation is involved, the ergonomics of the situation need careful consideration by the designer.

### ***Environment***

All aspects of the product or system's environment need to be considered. In our electric drill example, the operating environment was to include: dirt, dust, oil, and grease, as well as combustible dust. This has a significant

effect on the robustness of the design and the environmental protection required for safe and efficient use.

So factors such as dirt, dust, oil, grease, chemical spillage, temperature, pressure, humidity, corrosives and other pollutants, animal infestation, vibration, and noise should be considered. Apart from the effects on the designed product or system we, as engineers, should also take into account the likelihood of our design polluting the natural environment. For example, when considering the design of nuclear reactors, the vast majority of the design effort is focused on ensuring that fail-safe systems and back-up facilities exist. Thus minimising the possibility of a nuclear accident, which might result in an ecological disaster. Filtration systems to prevent the leakage of dangerous substances from plant, machinery, and vehicles should also be introduced into the design, as a matter of routine.

### ***Maintenance***

When purchasing a domestic appliance like our electric drill, the ease with which it can be maintained and serviced is of importance to the average DIY enthusiast. Therefore, the ease with which parts can be obtained and the drill can be assembled/dismantled are important when considering design for maintenance. When deciding whether or not to design-in a significant amount of maintainability, there are several factors that need to be determined. For example, we will need to know the likely market and establish their philosophy on servicing, repair, and rectification. Estimates of component life will need to be found in order to assess the economic viability of repair, reconditioning, or repair by replacement.

So when designing for ease of maintenance, we need to consider the extra costs involved in using more sophisticated manufacturing processes and component parts, and balance these against customer satisfaction with the finished product or system. The service life of the product and the life of component parts also needs to be established in order to make informed decisions about the need for maintenance.

### ***Costs***

A realistic product cost needs to be established as early as possible in the negotiating process. Estimates for products and systems are often set lower than reality dictates, because of the need to gain the competitive edge. However, it is no use accepting or giving unrealistic estimates which are likely to put a business into debt. Costing has become a science, and design engineers need to be aware of how to accurately estimate costs and financially evaluate the viability of new ventures. Because of the importance of finance, more in-depth information on costs and costing is given later in this chapter.

### ***Transportation***

Think of the consequences of designing and assembling a very large 100-ton transformer in Newcastle which is required in Penzance! The cost and practicalities of transporting such a monster, would make the task prohibitively expensive, if not impossible. Transportation becomes an issue when products that are very large and very heavy need to be moved. Thus when faced with these problems as a design engineer due consideration needs to be given to ease of transportation and packaging. Small items, such as our electric drill, can be packaged and transported by rail, road, sea, or air, with relative ease. Size and weight restrictions

must always be pre-determined, to ensure that the product will fit into the space allocated by the customer.

If, for example, we are designing, installing, and commissioning an air-conditioning system for a large hotel in Cairo, then due consideration must be given, at the design stage, to ease of transportation and assembly. In fact, large structures and systems are designed in kit form and dry-assembled, this ensures that all component parts are available, that they fit together, and that the installation does not exceed required dimensions, prior to shipment. Thus enabling the product or system to be easily installed and commissioned on sight and preventing any unnecessary transportation costs from being incurred.

### ***Manufacture***

Due consideration needs to be given at the design stage, to the ease of manufacture of products and their associated parts. The cost implications of 'over-engineering' must be remembered when designing, particularly with respect to design detail. Component parts should be designed with the over-riding thought of saving costs. All non-functional features and trimmings from part should be omitted. For example, do not design-in radii if a square corner will do and do not waste money on extra machining operations, if stock-size material is available and acceptable.

The cost and complexity of production methods are also need to be considered, at the design stage. One-off items are likely to require 'jobbing production', which is expensive in time and requires a high degree of skill. Batch production and mass flow production may need to be considered, always remembering the facilities available on customer premises.

Should items be bought-in, assembled using standard parts, or manufactured on-site? This will again depend on the philosophy adopted by the customer and the manufacturing facilities and tooling available.

### ***Aesthetics***

Once all the primary requirements regarding function, safety, use, and economy have been fulfilled, the aim of designers is to create products that appeal to customers, thus industrial design lies somewhere between engineering and art. Consideration needs to be given to shape or form, the use of colour, and surface texture.

The aesthetic design of an engineering product or system, may be outside the remit of the engineering designer. Specialists, such as graphic artists, could very easily be seconded to the design team to assist with product aesthetics, packaging, labelling, and graphics, for a particular market.

### ***Legal implications***

Designers will need to be aware of the legal constraints involved with any particular product or system. With the advent of more and more EEC legislation concerning product liability, disposal of toxic waste, control of substance hazardous to health (COSHH), and general Health and Safety, all legal aspects must be considered during the early stages of the design process.

The fundamentals of the law regarding patents and copyright should be understood by the engineering designer. A patent, for example, does not stop anybody using your idea, it merely provides a channel for redress. The incentive gained by patenting a product normally results in the item

being made and marketed, knowing that legal protection is offered. New industrial designs are protected under the provisions of the Copyright Act (1968) in the UK. This protection is offered without any form of registration and is valid for 15 years from the date of manufacture (in quantity) or when first marketed.

Some knowledge of the law of contract is also very useful for the engineering designer. A contract is a formal written agreement between two parties. Both parties agree to abide by the conditions that are laid down, in all respects. However, certain 'let-out' clauses may be included in order to accommodate unforeseen difficulties.

### **Safety**

Safety requirements vary according to the product or system being designed and the use to which they will be put. Many areas of engineering exist within highly regulated industries, where safety is of paramount importance, for example, nuclear power, petro-chemical, aircraft operation, hospitals, and the emergency services. For these industries and many others, adherence to HASAWA, COSHH regulations, and BS for product liability will form an essential part of the design process.

### **Quality**

To ensure quality in design, the design methods adopted, technical documentation, review processes, testing, and close co-operation with the customer must be such that the performance, reliability, maintainability, safety, produceability, standardisation, interchangeability, and cost are those required.

Many companies have achieved the Total Quality (TQ) Management System kite mark, in that they have gained BS 5750 or the ISO 9000 series equivalent. To obtain such a standard requires a company to establish documented methods of control of quality-affecting activities, training personnel in these methods, implementing these methods, verifying implementation, measuring effectiveness, and identifying and correcting problems to prevent reoccurrence, in short good business practice.

Quality control (QC) is the final part of the TQ process, if the TQ process is effective QC will be less and less needed, until eventually the ultimate *right first time* is achieved.

We have spent some time explaining the nature of the design parameters, although rather tedious, this should be treated as essential learning.

### **Question 2.1.2**

Your company receives the undermentioned customer's description (design brief) for a 'hydraulic hose connector'. From the customer's brief, using appropriate sources of reference, produce a PDS which meets all essential requirements and constraints. Your specification should be written following the format given earlier. Under the 'design' heading you should include all design requirements you think necessary. The style of the PDS should enable it to be used for communication between your company (via you, the design engineer) and your customer who wishes to produce the new hose connectors.



At this time, and for the purpose of this question, you do not need to include estimates of costs, timescales, or sketches of alternative design solutions. You will be asked to complete these exercises later!

#### **Customer's brief**

We currently produce portable hydraulic power packs used on a range of agricultural machinery. Our current design of flexible hose fittings used with the pack have caused difficulties when removing/fitting the power pack from/to the machines during maintenance.

Hose end connectors are required that will allow the attachment and removal of hoses to be carried out quickly and safely, without undue leakage of hydraulic fluid. The fittings will need to accommodate 10, 14, 20, and 30 mm flexible wire reinforced rubber hoses, capable of withstanding pressures up to 50 bar. The fittings must be robust enough to withstand harsh agricultural environmental conditions.

## **2.2 THE DESIGN REPORT**

### **Introduction**

When a professional engineer carries out a design project, a note book is kept in which are recorded the initial specification, design requirements, design parameters, alternative solutions, ideas, test requirements and results, calculations, general schemes, references, contacts, and a host of other related information. The contents of the note book are primarily for the design engineers use and serve as an essential reference during the lifetime of the project.

When the project is completed, the designer will generally convey his ideas by producing a report, which contains the all important 'description of the final design' together with layout drawings. The information required for the report being obtained and transferred from the designer's notebook. The reason for writing a report is to present findings in a readily understood form which would allow any reader to appreciate the nature of a particular design, and to allow people the opportunity to comment on and make suggestions for possible design improvements, if the specification changes or the need arises.

The production of the design report is the culmination of the design process, and it forms an important part of the final design documentation. It requires a high degree of intellectual ability in order to analyse all the major design parameters, draw conclusions, make judgements, and synthesise all the parts into a coherent, logical, and effective design solution, which is to be presented in the form of a design report. Some of the techniques, which have been designed to assist these thought processes, are described in the following section.

### **The design process**

Before considering the design report itself, it will be useful to look at the overall design process and determine some logical order and structure, in which to proceed from the design problem to a possible design solution.

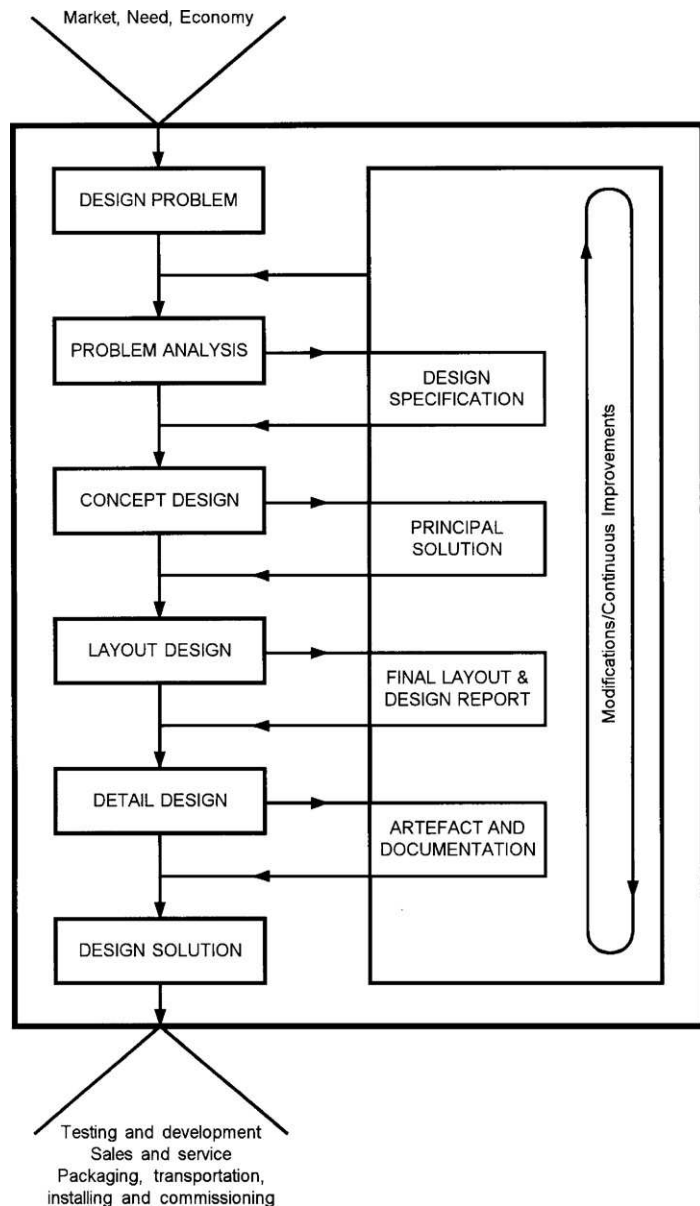


Figure 2.2.1

Figure 2.2.1 provides an overview of the design process showing one possible way in which to reach a satisfactory solution. We start by considering *the design problem*, this may be presented in the form of a customer's brief, or it may come as an idea for improving an existing in-house artefact. Tackling the design problem requires us to establish exact customer requirements, determine the major design parameters, research and obtain design information from appropriate sources, and prepare and produce *the design specification*.

Using the design specification as our guide, we next need to prepare an analysis of possible design solutions, produce concept designs, evaluate alternative concepts, and select an *optimum design solution*. From the *conceptual design* phase, where engineering principles are established, we move to the *layout design* or general arrangement design. Here we are concerned with the selection of appropriate materials, determining the design of the preliminary form, checking for errors, disturbances and

minimum costs, and producing the *definitive layout and design report*. The final stage of the design process is the *detail design*, where we are concerned with the arrangement, dimensions, tolerances, surface finish, materials specification, detail drawings, assembly drawings, and production costs of the individual parts of the product or system being designed.

### ***Determining possible design solutions***

In our search for an optimum solution to an engineering design problem, there are many methods available, which will help us to find such a solution. Which method should be used in which situation, will depend upon the nature of the problem, the magnitude of the task, the information available, and the skill, knowledge, and experience of the designers.

All solution finding methods are designed to encourage lateral thinking and foster an open-minded approach to problem solving. These methods may be conveniently divided into *general methods* and *problem specific methods*. The former are not linked to a specific part of the design process or to a particular product or system. They do, however, enable us to search for solutions to general problems that arise throughout the design process. Problem specific methods, as their name suggests, can only be used for specific tasks, for example to estimate costs or determine buckling capability of specified materials. In this section we will consider two examples of general methods: *brainstorming* and the *systematic search method*.

### ***Brainstorming***

The object of brainstorming is to generate a flood of new ideas, it is often used when there is a feeling that matters are becoming desperate. It involves using a multi-disciplinary team of people, with diverse backgrounds, who are brought together to offer differing perspectives to the generation of ideas. This method is particularly useful where no feasible solution principle has been found; where a radical departure from the conventional approach is considered necessary or, where deadlock has been reached.

The group will normally consist of between 6 and 15 people drawn from a diverse variety of backgrounds, and must not be limited to specialists. They are formally asked to focus their attention on a specific problem, or group of problems, in order to rapidly generate ideas for a solution. This technique was originally suggested by Alex Osborn, during his time in the advertising industry. He devised the set of criteria given below, in order to ensure that all participants were given equal opportunity to express themselves freely without inhibition:

- The leader of the group should be responsible for dealing with organisational issues and outlining the problem, prior to the start of the brainstorming session. The leader should also ensure that all new ideas are encouraged and that no one criticises the ideas of other group members.
- All ideas are to be accepted by all participants, no matter how absurd, frivolous, or bizarre they may seem.
- All ideas should be written down, sketched out, or recorded for future reference.
- Building on the ideas of others to create a group chain reaction should be encouraged.
- The practicality of any suggestion should be ignored at first, and judged later.
- Sessions should be limited to less than 1 hour. Longer sessions tend to cause participant fatigue and the repetition of ideas.

- The results should be reviewed and evaluated by experts, to find potential solutions to design problem(s).
- After classification and grading of the ideas, one or more suggested final solutions should be presented again to the group for interpretation, comment, and feedback.

### **Systematic search method**

This method relies on a mechanistic approach to the generation of ideas, through the systematic presentation of data. Data is often presented in the form of a *classification scheme*, which enables the designer to identify and combine criteria to aid the design solution. The choice of classifying criteria and their associated parameters requires careful thought, since they are of crucial importance. In order to illustrate this method, consider the following example:

#### **Example 2.2.1**

An engineering system is required to operate the ailerons of a light aircraft. The ailerons must be capable of being operated by the pilot from the cockpit. Use the systematic search method to produce design ideas for possible motion converters for the required aileron system. Figure 2.2.2 illustrates the required output motion for the system.

Possible motion parameters might include: linear to linear, linear to rotary, rotary to linear, rotary to rotary, and oscillatory, all of these forms of motion are sub-sets of translational (linear) and rotational motion. In order to provide design ideas for possible motion converters, a classification scheme needs to be produced. The first attempt is shown below in Figure 2.2.3 the column headings indicate the motion parameters and the solution proposals are entered in the rows.

In my classification scheme illustrated above, you will note that gears are included as a possible motion converter. In fact improvements to our scheme could be made by subdividing the solution proposals. For example, gears could be subdivided into spur and bevel gears, epicyclic gears, harmonic drives, worm gears, and helical gears. Note that our motion parameters have already been subdivided from translational and rotary motion. This process of layering our classification scheme enables us to produce many varied design proposals. No mention has been made of energy sources for the system, this has been left for you as an exercise.

#### **Question 2.2.1**

For Example 2.2.1, subdivide cables/belts/pulleys, actuators/rams, and linkages/rods/levers into as many different variants as you can. Reference to standard engineering design textbooks and specialist texts on mechanisms, systems, and machine design should prove useful.

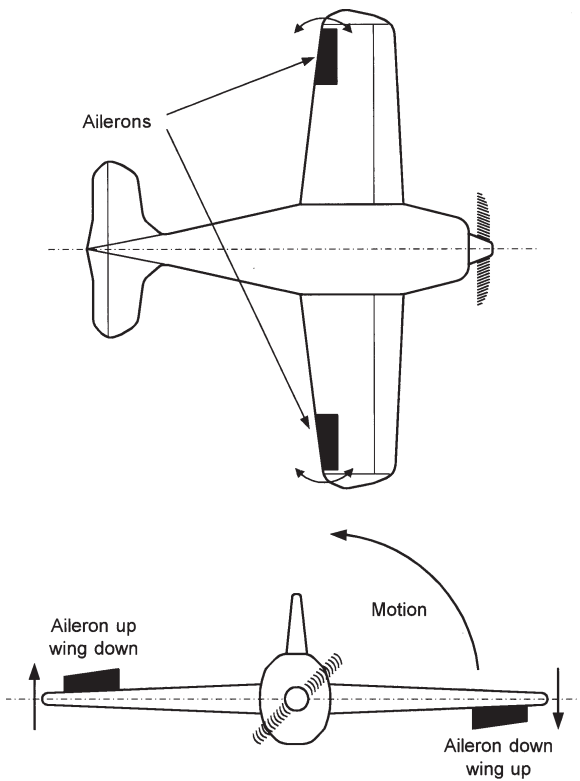


Figure 2.2.2

| Motion parameters<br>Solution proposals | Translational and Rotary |                                |                    |                                              |
|-----------------------------------------|--------------------------|--------------------------------|--------------------|----------------------------------------------|
|                                         | Linear to Linear         | Linear/Rotary<br>Rotary/Linear | Rotary to Rotary   | Rotary/Linear to Oscillatory or Intermittent |
| 1. Gears                                |                          | Worm and wheel                 |                    |                                              |
| 2. Cables/Belts/Pulleys                 |                          |                                |                    |                                              |
| 3. Cams/Followers                       |                          |                                |                    |                                              |
| 4. Chains/Sprockets                     |                          |                                |                    |                                              |
| 5. Actuators/Rams                       |                          |                                |                    |                                              |
| 6. Screws                               |                          |                                |                    |                                              |
| 7. Linkages/Rods/Levers                 |                          |                                | Universal coupling |                                              |

Figure 2.2.3

Question 2.2.2

List the possible energy sources that might be used to provide the input power for the aileron system described earlier. Also give solution proposals for the possible types of energy provider, that might be used to power the system.

There are many other solution finding methods available, which should be considered in addition to those mentioned. These include: literature search, analysis of natural and existing technical systems, model testing, galley method, and many more. As design engineers, all these methods should be familiar to you, time does not permit a full study of them all in this chapter, but further information is available from the reference material given at the end of this book.

Selecting and evaluating possible design solutions

Until now, we have been concerned with generating ideas for possible design proposals and we have looked in detail at one or two methods which help us to produce design concepts. It is now time to consider ways in which we can evaluate and so select the best of these solution variants.

| Design proposal<br>Selection criteria | Sketch of design proposal<br>with written description |    |   |   |    |   |    |   |   |    |
|---------------------------------------|-------------------------------------------------------|----|---|---|----|---|----|---|---|----|
|                                       | 1                                                     | 2  | 3 | 4 | 5  | 6 | 7  | 8 | 9 | 10 |
| i                                     | +                                                     | -  | + | - | -  | + | -  | + | + | +  |
| ii                                    | -                                                     | -  | + | - | -  | + | -  | + | E | +  |
| iii                                   | -                                                     | +  | + | - | +  | - | +  | + | - | -  |
| iv                                    | +                                                     | -  | - | + | +  | + | -  | - | - | -  |
| v                                     | +                                                     | -  | + | + | -  | + | +  | + | + | +  |
| vi                                    | +                                                     | E  | + | + | -  | + | -  | + | - | +  |
| vii                                   | E                                                     | -  | + | E | -  | + | +  | + | + | +  |
| viii                                  | +                                                     | +  | + | E | -  | + | -  | - | + | E  |
| Score<br>$\Sigma$ +, -, E             | 3                                                     | -3 | 6 | 0 | -4 | 6 | -2 | 4 | 1 | 3  |
| Rank order                            | 3                                                     | 7  | 1 | 5 | 8  | 1 | 6  | 2 | 4 | 3  |

Score value  
 + = +1, - = -1,  
 E = 0

Equal first

Figure 2.2.4

### The evaluation matrix

One of the most common methods is to produce an *evaluation matrix*, where each solution concept is set against a list of selection criteria. For each criteria some kind of scoring system is used to indicate how the individual design concept compares with an agreed norm. This process is illustrated in Figure 2.2.4.

Prior to inclusion in the evaluation matrix, if there are a large number of solution proposals, the design engineer should produce some form of pre-selection procedure, in order to reduce the proposals to a manageable size. The pre-selection process being based on fundamental criteria which the design proposal must meet. Such criteria might include: compatibility with required task, meets the demands of the design specification, and feasibility in respect of performance, etc. meet mandatory safety requirements and expected to be within agreed costs.

The concept proposals should be in the form of sketches together with a short written explanation. Equality with the agreed norm is shown in the skeleton matrix by the letter 'E', if the design solution is considered better than the norm, in some way, then a + sign is used, conversely a - sign is used, if the design solution is worse than the norm, in some way. A score may be obtained by allocating a +1 to the positives, -1 to the negatives, and 0 to the Es. More sophisticated scoring systems may be used involving 'weightings', when the selection criteria are not considered to be of equal importance.

### Example 2.2.2

Assume that you are to manufacture a large diameter flywheel, for a heavy pressing machine, at minimum cost:

- (i) Write a short list of selection criteria against which the given design solutions can be evaluated.

- (ii) Produce an evaluation matrix using the ‘casting method’ of manufacture as your norm, and rank all the remaining design solutions.

Figure 2.2.5 shows the given design solutions, we now need to produce our evaluation criteria. We could use one or more of the previous methods, to generate ideas. However, for the purpose of this example, since cost is of paramount importance, we will just look closely at the manufacturing methods which minimise cost. We will assume that the requirements of the specification have been met and that all design alternatives are compatible for use with the pressing machine under all operating conditions.

Then for each of the design options, we need to consider:

- materials costs;
- skill and amount of labour required;

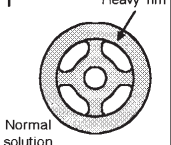
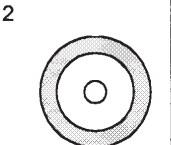
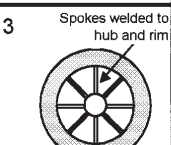
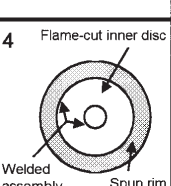
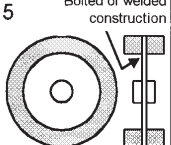
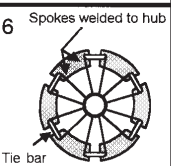
|   |                                                                                                                                                        |                                                                                                                                                                                                                                 |
|---|--------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 |  <p>Heavy rim</p> <p>Normal solution</p>                              | <p>CAST</p> <p>Hubs, spokes and rim sand cast in one operation. Wooden platen required to produce mould. Cast Iron (C1) or a steel may be used. Cast wheel requires limited final machining. Suitable finish to be applied.</p> |
| 2 |                                                                     | <p>MACHINED</p> <p>Flywheel machined from solid. Good machinability properties required of candidate material. C1 or a steel suitable.</p>                                                                                      |
| 3 |  <p>Spokes welded to hub and rim</p>                                | <p>FABRICATED</p> <p>Solid ring spun from standard section steel. Spokes welded to hub and outer rim. Hub machined from solid.</p>                                                                                              |
| 4 |  <p>Flame-cut inner disc</p> <p>Welded assembly</p> <p>Spun rim</p> | <p>FABRICATED</p> <p>Solid ring spun from standard section. Inner disc flame-cut to accommodate hub and rim. Hub machined from standard round section bar. Welding used for assembly.</p>                                       |
| 5 |  <p>Bolted or welded construction</p>                               | <p>LAMINATED CONSTRUCTION</p> <p>Outer rim spun from standard section. Centre plate flame-cut from standard section. Bolted or welded assembly. Simple machining operation for assembly.</p>                                    |
| 6 |  <p>Spokes welded to hub</p> <p>Tie bar</p>                         | <p>SEGMENTED CONSTRUCTION</p> <p>Number of segments varied according to stock-size materials and size of flywheel. Spokes welded to hub assembly and bolted to outer segments with the use of tie-bars.</p>                     |

Figure 2.2.5

- complexity of construction;
- tooling costs;
- machining and finishing costs;
- safety (this will be related to the integrity of the design solution assuming it is chosen);
- amount of waste generated;
- company preference – knowledge, skills, and equipment.

The above list of criteria is not exhaustive, but should enable us to select one or two preferred design alternatives. Further refinements/criteria may be necessary if two or more concepts are closely ranked.

### Scoring

Figure 2.2.6 shows the completed evaluation matrix for this problem. You will note that the company preference, immediately skews the scores. The company does not have or does not wish to use foundry facilities; in any case, the production of the mould would be prohibitively expensive for what appears to be a 'one-off' job. Obviously you, as the design engineer, would be aware of these facts before evaluating the options. Note that options 2–6 all involve some form of fabrication, assembly or machining, which we will assume is the company preference.

Proposal 2: Machining parts for a heavy flywheel will require several machining operations and the use of elaborate fixtures, not to mention operator skill

| Design proposal<br>Selection criteria | See Figure 2.2.5 for design proposal sketch and written description |    |   |   |   |   |
|---------------------------------------|---------------------------------------------------------------------|----|---|---|---|---|
|                                       | 1                                                                   | 2  | 3 | 4 | 5 | 6 |
| Materials                             | +                                                                   | E  | – | – | E | + |
| Labour                                | –                                                                   | –  | + | + | + | – |
| Complexity of constraint              | +                                                                   | +  | E | E | E | – |
| Tooling                               | –                                                                   | –  | E | + | + | + |
| Machining/finish                      | +                                                                   | –  | E | – | + | + |
| Safety                                | +                                                                   | +  | E | E | E | – |
| Waste generation                      | +                                                                   | –  | + | – | E | – |
| Company preference                    | –                                                                   | +  | + | + | + | + |
| Score<br>$\Sigma +, -, E$             | 2                                                                   | –1 | 2 | 0 | 4 | 3 |
| Rank order                            | NORM                                                                | 5  | 3 | 3 | 1 | 2 |

Figure 2.2.6



- for an object of such size; so labour, tooling, and machining costs are relatively high. This process also involves a large amount of material waste.
- Proposal 3: The major advantage of this method is that standard stock materials can be used. Difficulties include: the use of jigs and fixtures, weld decay, and possibility of complicated heat treatments.
- Proposal 4: Similar advantages and disadvantages to option 3.
- Proposal 5: Advantages include use of standard stock materials, little machining required after assembly, relatively easy to assemble. Disadvantages include necessity for positive locking of bolts after assembly and outer rim would require skimming after spinning.
- Proposal 6: Labour intensive fabrication and assembly, complex assembly, and integrity of construction would raise safety issue. No specialist tooling required, finishing relatively easy and cheap, minimal waste from each machining operation, company preference.

Note that if company preference had been for casting, then option 1 would probably have been preferable, provided it met the cost requirements. Options 5 and 6 appear next to favourite, although option 3 might also be worth looking at again, dependent on the skills of the labour force.

If there was insufficient evidence on which to make a decision, then more selection criteria would need to be considered. For example, do the options just meet or exceed the design specification, bursting speeds, and other safety criteria might have to be further investigated. This process would need to be adopted no matter what the artefact, an *iterative approach* being adopted, in an attempt to get ever closer to the optimum design solution.

### **Costing**

The cost of an engineering component or system is of paramount importance. Engineering designs require the specification to be met, the artefact to be produced on-time and *at the right cost*, if the design solution is to be successful. Thus an understanding of costs and costing procedures is something that every design engineer needs to achieve. A detailed exposition on costs and costing methods is given in Chapter 1, Business Management Techniques. Set out below are one or two important points concerning costing, directly related to the production of an engineering artefact.

### **Importance of costing and pricing**

The importance of producing a successful tender cannot be overemphasised, in fact the future of jobs within the company may depend upon it.

To ensure that a commercial contract to design, manufacture, and supply on time is won; there must be an effective costing and pricing policy.

Not only must the contract be won, against competition, but a profit margin needs to be shown. Price fixing needs careful planning, clearly too high a price may not result in a successful tender and too low a price may cause financial loss to the company, particularly if there are unforeseen difficulties.

For profit, and as a 'useful rule of thumb', we should know our costs to within  $\pm 2\%$ . The tolerable margin between maximum and minimum prices is small, thus the necessity for *design* and *costing accuracy*.

Some important general reasons for costing are to:

- determine the viability of a proposed business venture;
- monitor company performance;
- forecast future prospects of a business deal;
- price, products, and/or services;
- meet legal requirements to produce records of company viability, for public scrutiny as required.

Below are some useful definitions concerned with cost and price:

*Price:* money paid for products or services.

*Value:* the amount of money someone is prepared to pay for products or services.

*Cost:* all money spent by a supplier to produce goods and services.

*Material cost:* (volume  $\times$  density  $\times$  cost/kg) plus an amount for wastage.

*Labour cost:* (operational time  $\times$  labour rate) plus wasted labour time which is not directly related to the task.

### **Standard costing sheets**

These are used to ensure that all parameters are considered when costing a product or service. Some standard costing sheet headings together with their definition, are given here:

- A: Direct material cost: raw material and bought-in costs.
- B: Direct material scrap: materials subsequently scrapped (typically 3–5% of A).
- C: Direct labour cost: wages of production operations, including all incentive payments.
- D: Direct labour scrap: time spent and paid for on artefacts, which are subsequently scrapped, this would include the costs of machine breakdown or other reasons for stoppages to production (typically 3–5% of C).
- E: Prime cost: the sum of all material and labour costs that is  $A + B + C + D$ .
- F: Variable overheads: cost of overheads which vary with rate of production, these might include: fuels costs, cost of power supplied, consumables, etc. (typically 75–80% of C).
- G: Manufacturing cost: this is the sum of prime costs and variable overheads ( $E + F$ ).
- H, I, J: These are packaging, tooling, and freight costs, respectively.
- K: Variable cost (VC) this is the sum of the previous costs,  $G + H + I + J$ .

- L: Fixed overheads (FO): overheads which do not vary with production output, these include all indirect personnel not involved with production, marketing costs, research and development costs, equipment depreciation, premises costs (typically 30–40% of K).
- M: Total cost (TC): the sum of all direct VCs (K) plus indirect costs (L).

Thus,  $TC = \text{direct VCs} + \text{indirect costs (FO)}$ .

### Example 2.2.3

A company has been commissioned to produce 2000 high-quality metal braided shower hoses, complete with fixtures and fittings. Assuming that:

- (i) direct material cost per item is £1.50 and material scrap is estimated to be 3% of material costs;
- (ii) direct labour costs total £8000 and the labour scrap rate is 4% of direct labour costs;
- (iii) variable overheads are 75% of direct labour costs;
- (iv) FO are 30% of VCs;
- (v) packaging, tooling, and freight costs are 10% of manufacturing costs.

Estimate the selling price of the shower hose, if the company wish to make a 25% profit.

This problem is easily solved by laying out the costing sheet as shown below, and totalling the amounts.

| Cost                                 |                                                   | Amount (£) |
|--------------------------------------|---------------------------------------------------|------------|
| Direct material cost                 | $(1.5 \times 2000)$                               | 3000       |
| Direct material scrap                | (3% of 3000)                                      | 90         |
| Direct labour cost                   |                                                   | 8000       |
| Direct labour scrap                  | (4% of 8000)                                      | 320        |
| Prime cost                           |                                                   | 11 410     |
| Variable overheads                   | (75% of 8000)                                     | 6000       |
| Manufacturing cost                   | (prime + variable)                                | 17 410     |
| Packaging, tooling, and freight cost | (10% of 17 410)                                   | 1741       |
| VC                                   | (manufacturing + packaging, tooling, and freight) | 19 151     |
| FO                                   | (30% of 19 151)                                   | 5745       |
| TC                                   | (variable + FO)                                   | £24 896    |

Now company is required to make 30% profit.

So selling price per item is  $\frac{24\,896 + 7469}{2000} = \text{£}16.18$ .

Here is something to remember when considering costs: *always design parts with the over-riding thought of saving money and do not forget that omitting all non-functional features and trimmings from parts saves production time!*

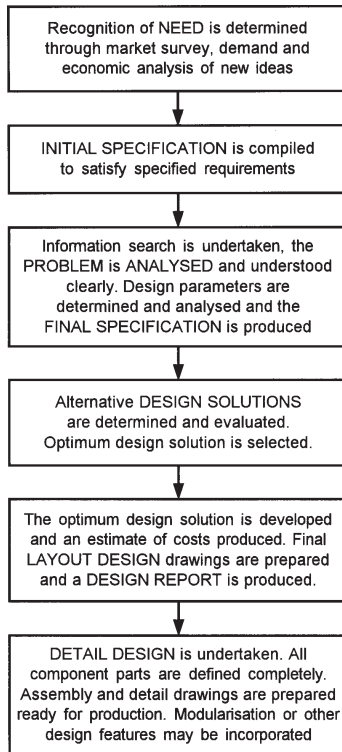


Figure 2.2.7

## Summary of the design process

We leave this section with a summary of the design process (Figure 2.2.7). If the process is followed it will ensure that an orderly and logical approach to engineering design is achieved.

## Format of the design report

In the last section we looked at the design process, including a number of ways in which to identify and evaluate design solutions. These evaluation methods are equally suitable for determining solutions during the concept, general arrangement (layout), and detail phase of the design process. Information on the layout design, as well as conceptual design should appear in the report. At the end of this chapter you will find several questions, which have been designed to help you improve your ability to think laterally, throughout all phases of the design process. Here, we are concerned with report writing, the layout of the design report, and the detail expected within each section.

The following general information is given for guidance only. The report content and layout may differ slightly from that given, depending on the nature and requirements of the design task. More specific information on report writing may be found in *BS 4811. The presentation of research and development reports* British Standard Institution (BSI) (1972).

**Title page:** This should include a clear and precise title for the design and contain the designer's name and company details as appropriate.

**Acknowledgements:** These should always appear at the front of the report. They should include individuals, companies, or any associated body who has provided the design engineer with help and advice. This may include assistance with regard to literature, materials, information, finance, or any form of resource.

**Summary:** This should provide a brief statement of the design problem, its solution, and any further recommendations with respect to development and testing. References may be made to other areas of the report, in order to clarify the design description.

**List of contents:** This should contain a list, which provides the page number of all the main headings as they appear in the report. A separate list of all diagrams, sketches, drawings, illustrations, and photographs should be provided, indicating figure numbers, page numbers, plate numbers, and drawing numbers, as appropriate.

**Introduction:** This should provide all background detail to the project and, give an indication to the reader as to why the design was undertaken.

**Specification:** This section should include the design requirements in the form of a statement of the initial design specification.

**Design parameters:** A description of all the design parameters related specifically to the design in question should be given. The design parameters will include those concerned with the engineering aspects of the product or system being considered, as well as organisational factors. Any modifications to the original design specification should be given, stating all assumptions made, and giving reasons for such decisions.

**Description of design:** This is the most important section within the report. It should contain an explicit, succinct description of the final design solution, indicating clearly its function and operation. Sketches should be

provided to clarify specific areas of the design solution and references to formal drawings; in particular, the general arrangement drawing should be made.

*Design evaluation:* This section should contain a critical appraisal and appreciation of the final design solution. Recommendations for further development and testing should also be given, to enable improvements to be made to specific features of the design, as required.

*References:* The reference list should contain only those references that are mentioned in the text. They are normally numbered in the same order in which they appear in the text.

*Appendices:* These contain all supporting material necessary for the report which is not essential for inclusion or appropriate for inclusion into the main body of the report. The material contained in the appendices should be referred to in the text of the main report.

Appendices are often identified using a Roman numeral. The following list gives a typical selection of appendix material for a design report:

- Evaluation of alternative design solutions, including sketches and description of alternatives.
- Details of decision-making processes, such as evaluation matrices, decision trees, etc.
- Theoretical calculations, mathematical derivations, formulae, and repetitive calculations.
- Evaluation of materials selection, for all phases of the design.
- Evaluation of appropriate manufacturing processes.
- Consideration of human factors.
- Costing considerations and pricing policy.
- Details of correspondence, associated with the design.
- Description of specialist test and development equipment.
- Details of experimentation and record of associated data.
- Computer programs and evaluation of computer printouts.

*General layout:* The design report should follow a recognised hierarchy for headings. There are several ways in which the relative headings can be laid out these include: numbering, indenting, use of capital and lower case letters, boldening, and underlining. Below is an example of the decimal numbering system, which has the advantage of easy and accurate cross-referencing, when required.

## 2 FIRST LEVEL HEADING

### 2.1 SECOND LEVEL HEADING

#### 2.1.1 Third level heading

##### 2.1.1.1 Fourth level heading

The report should be typed or word processed on one side of the paper, with appropriate margins. The text is normally double-spaced or similar, to provide the opportunity for specific comments/suggestions, for amendment.

Pages before the table of contents are numbered using lower case Roman numerals, pages following the table of contents should be numbered using Arabic numerals. The style and placement of page numbers should be consistent.

In order to assist with the design process and the design report, the design engineer needs to be familiar with some important aspects of computer technology, this is detailed in Section 2.3.

## 2.3 COMPUTER TECHNOLOGY AND THE DESIGN PROCESS

### Introduction

---

The computer has become a very important and powerful tool in engineering design and manufacture. Computer-aided design (CAD) is a design process which uses sophisticated user friendly computer graphic techniques together with computer software packages, which assist in solving the visualisation, analytical, development, economic, and management problems associated with engineering design work.

The key features of a CAD system include: two-dimensional (2D) and three-dimensional (3D) drafting and modelling, parts and materials storage and retrieval, provision for engineering calculations, engineering circuit design and layout, and circuit and logic simulation and analysis.

Computers may be applied directly to the design process in a number of areas, these include: *geometric modelling*, where structured mathematics is used to describe the form or geometry of an object. The *analysis* of engineering situations where forces, motion parameters, endurance, and other variables may be investigated for individual design situations. *Reviewing and evaluating* the design is made easier using computer graphics, size dimensions and tolerances may easily be checked for accuracy, and minute detail can be closely scrutinised using the magnification facility of the graphics system. *Automated drafting* has greatly improved the efficiency of producing hard copy drawings, which can be easily amended, as the design evolves (Groover and Zimmers, 1984).

There is a plethora of software packages associated with the engineering design process. 2D drafting packages enable the production of single part, layout, general arrangement, assembly, subassembly, installation, schematic, and system drawings. 3D drafting and modelling packages, in addition to engineering drawing, enable us to visualise and model, product aesthetics, packaging, ergonomics, and the differing effects of colour change, textures, and surface finish.

Design analysis packages enable us to perform calculations involving area, volume, mass, fluid flow, pressure, and heat loss as well as the analysis of stress, and the modelling and analysis of static, kinematic, and dynamic engineering problems. Circuit and system modelling packages enable us to simulate and modify the layout, design, and operation of electronic, fluid, control, and other engineering systems.

Many other packages exist which enable design engineers to:

- program computer numerical controlled (CNC) machinery;
- design jig and fixtures;
- project plan;
- store, retrieve, and modify technical information, publications, and company literature;
- prepare tenders and estimates;
- generate and maintain materials stock lists;
- perform other engineering management functions.

The computer and its associated software, has thus become a very important part of a design engineer's armoury.

### Computer-aided design

---

The use of a CAD system has many advantages over the more traditional design systems it has replaced. Improvements in productivity have obvious benefits for the company. Lead times from conception to embodiment,

detail design, and manufacture may be significantly reduced, dependent on the type and sophistication of the CAD system adopted.

Drafting packages range from simple 2D systems to semi-automated highly sophisticated 3D drafting and modelling systems, which contain large libraries of commonly used utilities. *Macros* may be provided which enable the user to define a sequence of commands to be executed in one instruction, this is a useful aid to speeding up repetitive routines (Lock, 1992). Drawing visualisation is easier with 3D modelling systems, orthographic projection may unknowingly be misinterpreted. However, being able to produce the equivalent isometric or perspective projection, which may be further enhanced with shading or colour, makes the CAD drawing much easier to interpret.

Design analysis using a CAD system, has already been mentioned, but it is worth emphasising here the versatility, power, and accuracy of these systems. Greater accuracy may be achieved in design calculations, with correspondingly fewer errors. Since errors cost time and money to rectify, the employment of a CAD analysis system, at the right scale for company needs, is advantageous.

The classification of engineering information into specific categories is necessary in order to be able to establish and implement document handling procedures. A wealth of historical data which may be needed by the company for future projects, needs to be stored in such a manner that it is easily retrieved for use, the modern CAD system is ideal for this purpose. Drawings, planning documents, correspondence, tenders, and other useful information is easily stored on magnetic tape, floppy disk, CD ROM, or some other electronic form, which is easily catalogued for quick access.

The remainder of this section is concerned with the computer hardware needed for a typical CAD workstation, and a more in-depth look at software packages for drafting, modelling, and analysis.

## Computer hardware

A typical engineering design computer workstation will contain the following items of hardware:

- input devices;
- microprocessor and interface unit;
- visual display unit (VDU);
- output devices;
- storage devices.

Figure 2.3.1 shows the components of a typical CAD workstation.

### Input devices

Typical input devices include the *alphanumeric* keyboard which is used to input numbers, text, or instructions. The *electronic digitiser* which is used for entering information on co-ordinates from existing graphical images with the help of a hand-held cursor. A *menu tablet* may be used as a low-resolution digitiser for accessing standard large scale details. The *mouse* which is a hand-held device with a roller in its base which controls the screen cursor by quickly following the input movements made by the mouse. The *joystick* is another cursor control device which functions in the same way as a mouse. The *light pen*, which looks similar

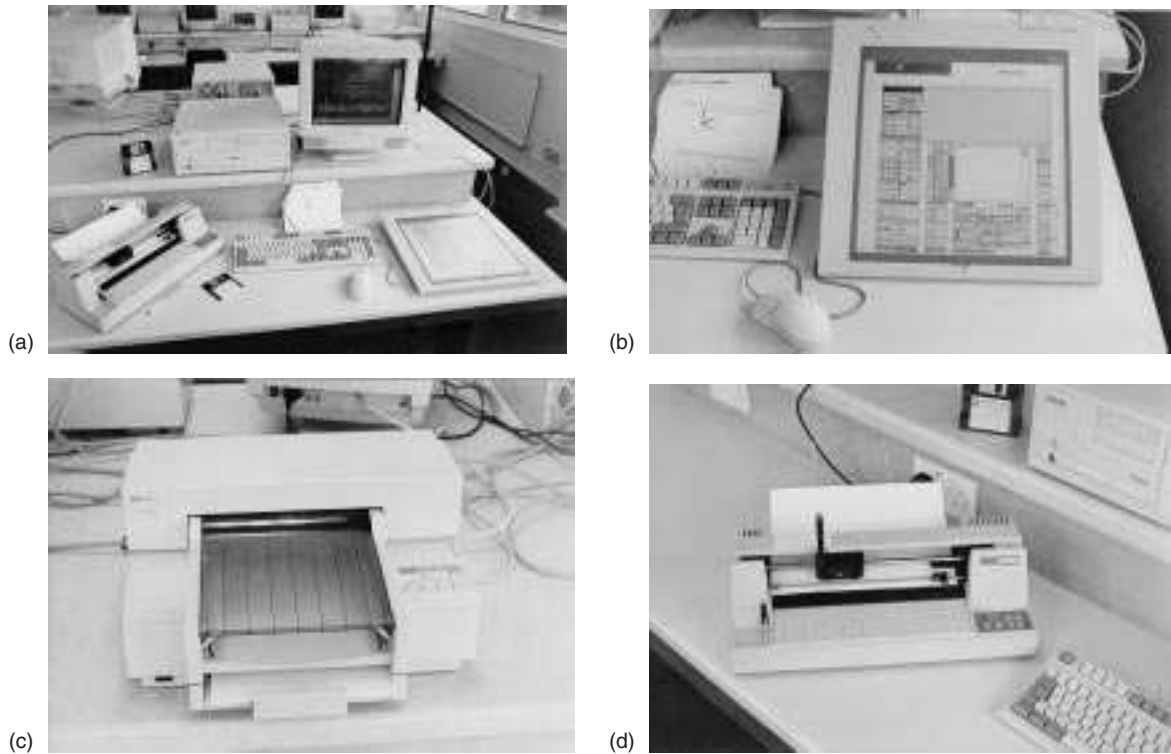


Figure 2.3.1a–d

to an ordinary writing pen except that it has a stylus tip, it is operated by placing the tip on the screen and moving as required.

Although not strictly hardware information for use in the computer can be stored, loaded, and downloaded to or from the machine using floppy disks, CD ROMs, scanners, and now the electronically modelled information from a digital camera. The latter with the ability to alter detail and other features of the digital photograph may prove useful for designers in the future.

### Microprocessor

This is often thought of as the 'brain' of the computer, the microprocessor temporarily stores and executes program instructions on data. A microprocessor is a miniature version of the central processing unit (CPU) of a digital computer. Its versatility is due to the fact that it is program controlled, simply by changing the program it can be used as the 'brain' of not only a microcomputer but a calculator, washing machine, industrial robot, for sequencing traffic signals, and many other applications.

The operations within the microprocessor dictate its architecture, which essentially consists of seven pieces of hardware (Figure 2.3.2). The *arithmetic and logic unit* (ALU) which handles the arithmetic calculations and logic decision-making processes. *Registers* which temporarily store data, address codes, and operating instructions. One of them, called the *accumulator*, contains the data actually being processed at any one time. Some type of *control unit* must be provided for the control and timing of operations within the entire microprocessor chip. The *clock* circuit generates timing pulses at a frequency of several megahertz to synchronise and control operations. The *bus controller* connects and controls the



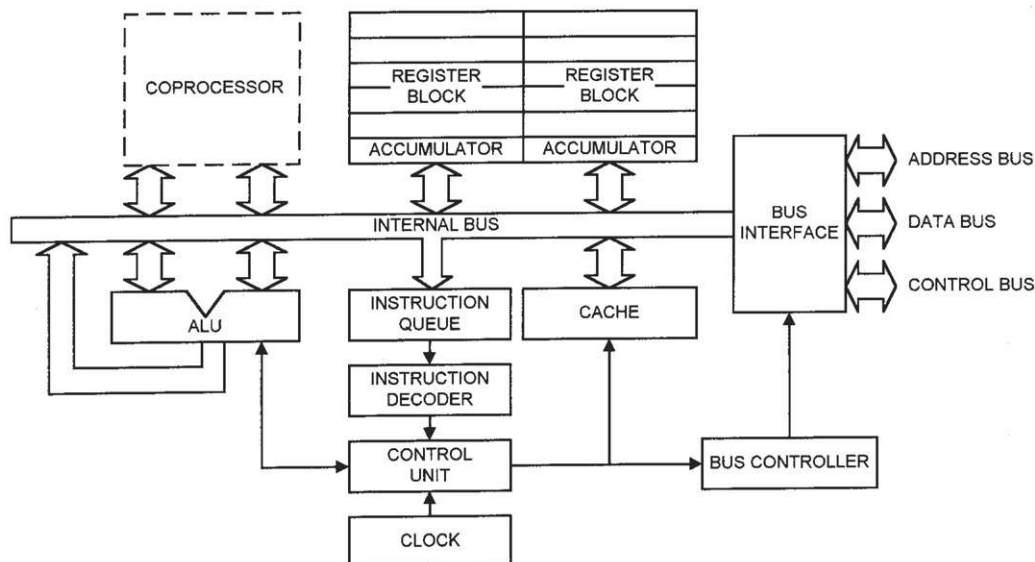


Figure 2.3.2

data, address, and control lines between the system bus and the processor. The *coprocessor* is activated when complex mathematics is processed, for example, when using spreadsheet or CAD software packages. The internal *cache* improves the data throughput of the microprocessor by allowing frequently used data and instructions to be stored internally within the chip, thus reducing the need for access to the external bus.

### Visual display unit

The VDU screen is used to display information and, where only alphanumeric data is displayed, low-resolution monitors are adequate. However, when graphics are used high-resolution screens are usually employed, as they give a clearer picture. The high-resolution monitors are normally larger, permitting a greater area to be viewed without loss of definition.

When the design engineer wishes to display graphics there are a variety of graphics monitors, which may be chosen. Considerations such as resolution, computer power requirements, speed of operation, and animation requirements dictate which type of VDU to use. The three most common types of monitor, that all use the cathode ray tube (CRT) as the display device are: the direct view storage tube (DVST), the vector refresh system, and the raster refresh system. *Refresh* is a system whereby the image has to be constantly regenerated in order to avoid it visibly flickering on screen.

The CRT is a vacuum tube in which an image is made visible on a phosphorescent viewing screen that glows when a narrow beam of electrons strikes the phosphor. There are three basic elements to a CRT these are:

- the electron gun which produces and focuses the beam of electrons;
- some means of deflecting the beam according to the signals present;
- the screen which converts the screen into visible light.

### The DVST system

In this system, information is sent to the screen once only, by a flood of electrons which converge onto the main electron beam and light the

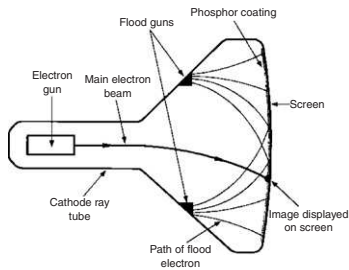


Figure 2.3.3

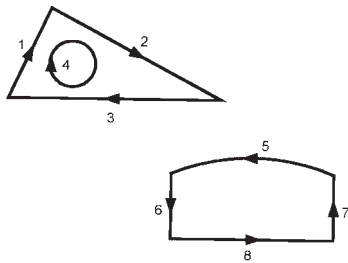


Figure 2.3.4

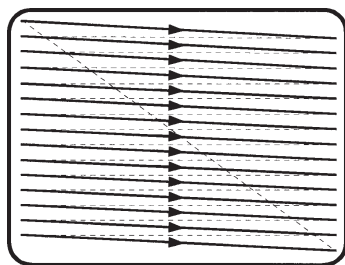


Figure 2.3.5

required area even after the beam has moved onto another position (Figure 2.3.3). This system produces excellent line quality and is able to retain the projected image for a period of time, but this system has several disadvantages. These disadvantages include: the requirement to re-draw the entire screen if just a single line is edited, thus slow in operation; animation is difficult to achieve; colour images are not usually available; the system is unable to create filled-in areas on the drawing; and the system needs to be used in dimly lit areas since flood lighting is insufficiently bright under normal lighting conditions.

### Vector refresh

This system was developed to overcome many of the disadvantages of the DVST system. The screen is regularly updated within a fraction of a second, under CPU control. There are no flood guns as with the DVST system. Editing of single elements is easily achieved. This system has excellent line quality, high-drawing speed, and can accommodate 2D and 3D animation.

Disadvantages include high costs, although with improving technology, prices are now almost comparable with other graphic display systems. Screen flicker can present problems with complex drawings, if the refresh rate slows sufficiently to become less than the flicker threshold of the human eye. Difficulties may also be experienced when trying to in-fill areas on the screen.

The generation of images using this system relies on vector drawing in sequential steps, the refresh system also uses a vectored line approach (see Figure 2.3.4).

### Raster refresh system

This system is similar in operation to that of a normal television, where thousands of picture element dots (pixels) may be illuminated and the brightness of the pixels controlled by the intensity and direction of a narrow electron beam. The refresh mechanism zig-zags from the top right to the bottom left of the screen approximately 50 times a second (Figure 2.3.5). The picture resolution may be varied according to how many pixels are available, these typically vary from  $320 \times 240$  low resolution to  $1024 \times 1024$  high resolution. Advantages include bright and clear monochrome and colour picture quality, no flickering problems, image in-fill easily achieved, and low costs making this a widely used system. Disadvantages include 'staircasing' effect on lines and curves, the fact that the pixel display must be calculated for each refresh which makes heavy demands on computer memory and animation is only available when extensive computing facilities exist.

### Output devices

The output devices most relevant to a CAD workstation are pen plotters, electrostatic plotters, and printers. There are essentially two types of pen plotter: flat bed and drum (see Figure 2.3.1(d)). Both types of plotter use an ink pen which produces lines on the paper, as the pen and paper move relative to one another, when signalled. Multi-colour production is achieved by using a numbers of pens. Drum plotters are generally less expensive than their flat bed counterpart. Drawings of virtually any length can be produced on the drum plotter, although the width is limited by the length of the drum. Drum and flat bed plotters have relatively slow plotting

speeds, although the drum plotter is normally faster than the flat bed, however, both types of plotter are highly accurate.

Electrostatic plotters have fast plotting speeds but low accuracy, they are ideal for producing draft copies quickly. These plotters can only process data that is presented in the raster format, unless some form of conversion equipment is added. They are, like the drum plotter, capable of producing drawings of any length, the width being limited to around 1.5 m.

Many CAD systems can now output to printers or onto photographic film. There are three main types of printer that are used commercially, the dot matrix, laser, and ink jet. Laser printers have high-quality reproduction and can now print in colour. Ink jet printer have slightly poorer definition, when compared with the laser, but are cheaper. Dot matrix printers, have until recently, given relatively poor reproduction but with the advent of 24 pin systems, the quality of reproduction has improved. All of these printing devices are still limited to relatively small paper sizes, when compared to conventional plotters.

### ***Storage devices***

Mention has already been made of magnetic tape and floppy disk storage devices. These storage devices and others are required for the files generated by the CAD system. For personal computers (PCs) and workstations, there is typically storage available near to the processor, or a common storage device provided through a network. Local storage can be achieved by saving to the hard disk or a floppy disk, whereas storage for a network might be saved to a common magnetic tape system, with back-up on floppy disks. On central systems, the storage devices are normally attached to the central processor.

The use of networked or central storage provides greater flexibility and the rapid transfer of information between users. When information is available to all users, the question of accessibility and confidentiality becomes a problem. Systems need to be introduced which safeguard the integrity of stored information, this is most easily achieved through the use of networked or centralised storage systems, where individual control is maintained through the use of access codes.

## ***Computer-aided drafting and modelling***

### ***Computer graphics***

At the heart of any CAD system is the software which enables the artefact to be graphically represented and described. This software enables the computer to understand geometric shapes, when the operator inputs commands. A CAD and drafting package needs to be able to perform the following operations:

- Generate graphic elements, such as points, lines, arcs, circles, and tangents.
- Allow objects to be transformed by scaling, moving, mirroring, or rotating them.
- Display components by windowing, providing alternative views, layering, etc.
- Allow common shapes to be drawn and stored for future use in a component library.
- Ability to program the computer to generate a component drawing or model automatically by specifying as a set of dimensions (user defined parametric macros).

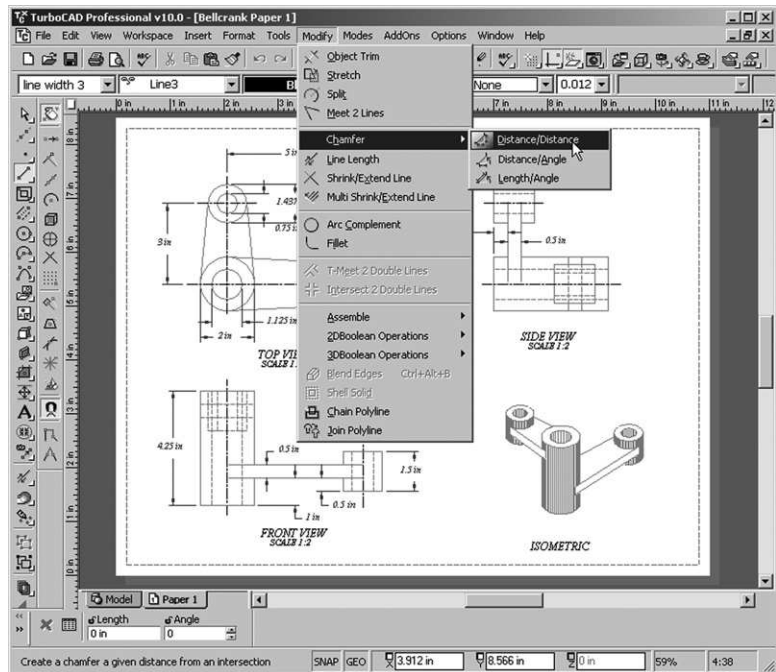


Figure 2.3.6



Figure 2.3.7



Figure 2.3.8

- Provide standard macros.
- Allow non-graphical information to be stored with the drawing or model (attributes).
- Ability to program package to suit individual needs.

In Figure 2.3.6 is an illustration of a TurboCAD computer drafting package window showing a pull-down menu which is accessed from the main menu, using the cursor or keyboard. The tools menu, for example, is used for pinpointing the ends of lines, centres of circles, etc. Note also the icon menus, typical examples of screen icons used with a graphics package are shown in Figure 2.3.7. Macros, you may remember, allow a sequence of operations to be executed in one instruction. Two macro programs are given below, one is a 'script file' used with TurboCAD, whilst the other is an AutoSketch 'macro'. Both result in the production of the nine-sided polygon shown in Figure 2.3.8!

*TurboCAD script file:*

```
DrawLinePolygon(9) : Polygon will have nine sides
ClickAt(5,5)       : Centre it on 5,5
ClickAt(5,7)       : Starting point for drawing
SetTextJustify(5)  : Justify label text
SetTextCSize(0.5)  : Set up text size and
DrawText(POLYGON) : place the text in
ClickAt(5,5)       : the centre of the polygon
EditMode
```

*Autosketch macro:*

```
DRAWPOLYGON
SETPOLYGON
SET POLYGONSIDES 9
```

```

DialogBoxReturn 1
DRAWPOLYGON
POINT 5,5
POINT 5,7
DRAWQUICKTEXT
SETTEXT
SET TEXTHEIGHT 0.500000
SET TEXTJUSTLEFT 0
SET TEXTJUSTCENTER 1
DialogBoxReturn 1
DRAWQUICKTEXT
POINT 4.808734,4.896243
STRING POLYGON\013

```

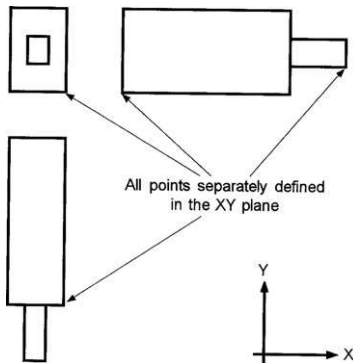


Figure 2.3.9

### Computer modelling

There are three basic ways in which we may graphically model an object. *2D* representation, which most closely resembles the method of working on the drawing board. This representation has limitations, only the plane of the screen or paper is used (the  $xy$  plane), so the object has to be drawn using different elevations (see Figure 2.3.9). The points that go to make up a view are separately produced and so are not interrelated, thus changing features in one view will not automatically produces changes in the other views.

$2\frac{1}{2}/D$  representation, portrays a 3D object, but with the third dimension being constant in section or having a constant section which is allowed to revolve around its axis, for example as in turning operations (Lock, 1992).

*3D* representation which can completely define the shape of the component (Figure 2.3.10). 3D systems are subdivided into: *wire-frame*, *surface*, and *solid modelling* which grow, respectively, in sophistication and so require a corresponding increase in available computer power. 3D software packages store the finished component in the computer's memory as a real shape, even though the CAD system is only able to display it on a 2D screen.

### Three-dimensional wire-frame modelling

This is considered to be the lowest level of 3D modelling, where the component is described in terms of points and lines only (see Figure 2.3.11). This system was originally developed to model simple machining operations. It has a number of limitations which include: no detail or shading of faces; the inability to distinguish between the interior and exterior of a solid object, thus causing problems in visualising different orientations; inability to detect interference between component parts; and unreliable calculations with respect to physical properties, such as mass, surface area, centre of mass, etc.

### Surface modelling

In this system the points, lines, and faces are defined, this system is more powerful than wire-frame modelling and as a consequence requires more computer power, such as that available from a 32-bit minicomputer. This system is able to recognise and display complex curves and profiles (see Figure 2.3.12). Since faces are recognised, shading is possible to help enhance the computer image. Holes are also easily identified.

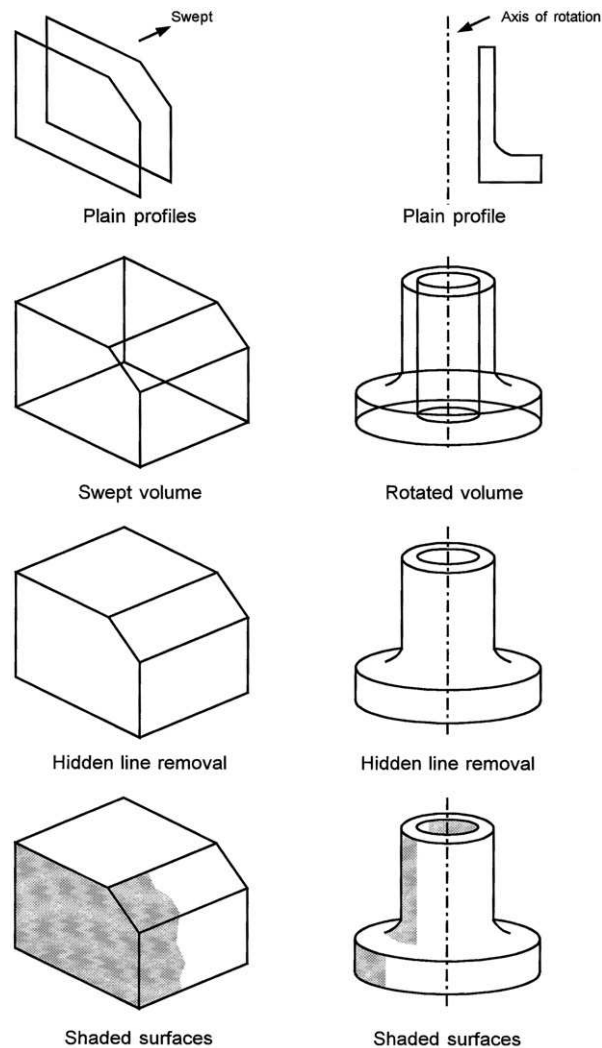


Figure 2.3.10

Applications of the system include complex tool path simulation and robot simulation. Complex curved surfaces, such as automobile body panels, plant ducting, cowlings, and fairings, may be designed using this system.

Limitations of the system include: the inability to comprehend solid volumes and so subsequent volume data may be unreliable; hidden lines cannot be easily removed; and internal sections are difficult to display.

### **Solid modelling**

This is the most sophisticated of all computer-aided modelling systems requiring a large amount of computing power, from at least a 32-bit computer. Solid modelling enables the user to describe fully and unambiguously 3D shapes (Figure 2.3.13).

The advantages of this system are numerous. All 3D features are easily defined, including internal and external detail. There is a facility for automatic line removal. Shading and variable colour graphics are available. Finite element analysis (FEA) is easily carried out (see later). Component sectioning is easily and effectively achieved. Physical parameters may be accurately calculated. The system provides excellent simulation

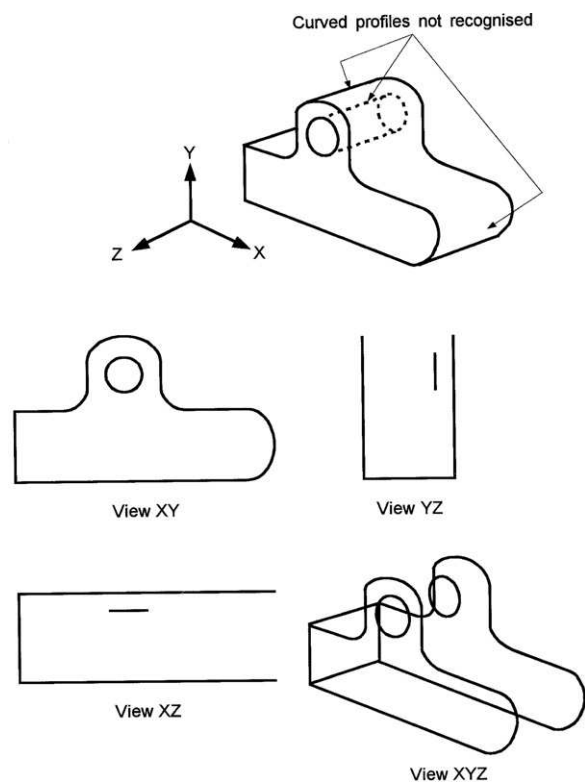


Figure 2.3.11

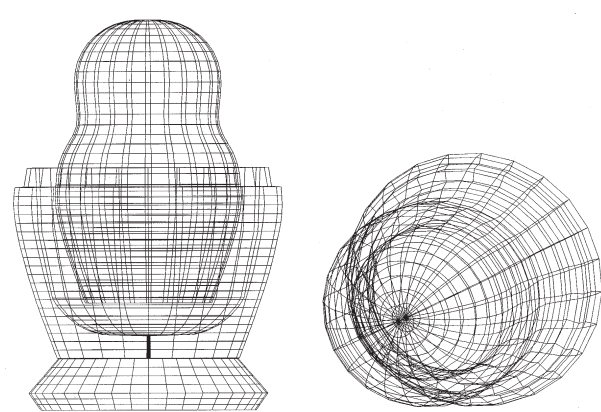


Figure 2.3.12



Figure 2.3.13

facilities for motion parameters, such as mechanism dynamics and robot configurations.

The obvious disadvantage has already been mentioned, a large amount of computer power is required so the system is very expensive. However, as technology continues to advance, cheaper hardware and even more sophisticated software will become available.

### Computer-aided analysis

We have discussed a variety of tasks that may be undertaken by typical design analysis packages, to illustrate this area of CAD we will discuss a method known as *FEA*. This widely used technique predicts the mechanical characteristics of a design under load. These loads may result from imposed stresses, temperature and pressure changes, or other forces acting on the system.

Before a FEA can be carried out, the finite element model (FEM) of the design must be determined. This is then used as an input to the FEA program. The purpose of an FEM is to represent the proposed design in a testable form.

FEM is the process of simplifying the design, for analysis by subdividing the entire object into a large number of *finite elements* usually rectangular or triangular for 2D work, and cubic or tetrahedral for 3D work. These elements are discrete (individually recognisable) and as such form an interconnecting network of concentrated *nodes*, as shown in Figure 2.3.14.

FEA consists of software programs that use an FEM to perform a variety of tests when given a set of parameters. FEA software programs are separate from CAD programs but use CAD generated models to perform the specific tests. In addition to FEM these programs require the input of other information variables, such as the material specification, the loads to be imposed on the structure, dimensions, and other geometric factors, heat flow characteristics, and so on. FEA is thus able to analyse the entire component for stress, strain, heat flow, etc. by calculating the interrelating behaviour of each node in the system.

Today, modern advanced CAD systems have the capability to automatically model the network (meshes and nodes), without the design engineer having to specify the FEM. This saves valuable time and ensures the relevance of the model selected. The output of the FEA provides an indication as to whether or not the design conditions have been met. The results also include a graphical output which is displayed on the workstation monitor. Such outputs may be in the form of contour plots for stress, strain, or heat flow, among other factors (see Figure 2.3.15).

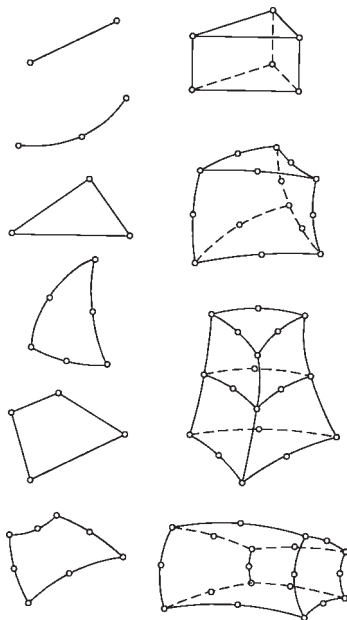


Figure 2.3.14

### CAD and manufacture

Little has been said about the relationship between CAD and computer-aided manufacture (CAM), where in a CAD/CAM system data is exchanged between each. CAM systems cover activities which convert the design of the product into instructions that define how the component will be produced. CAM is thus any automated manufacturing process which is controlled by computer. These will include CNC lathes, millers and punches, plasma flame cutters, process systems, and robotic cells, to name but a few. Other more sophisticated computer controlled manufacturing techniques, such as flexible manufacturing systems (FMS) and computer-integrated manufacturing systems (CIMS), all come under the umbrella of CAM.



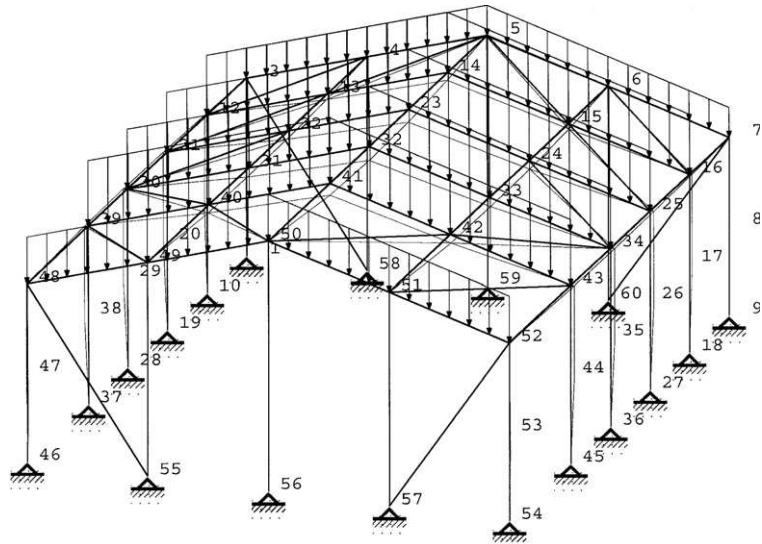


Figure 2.3.15

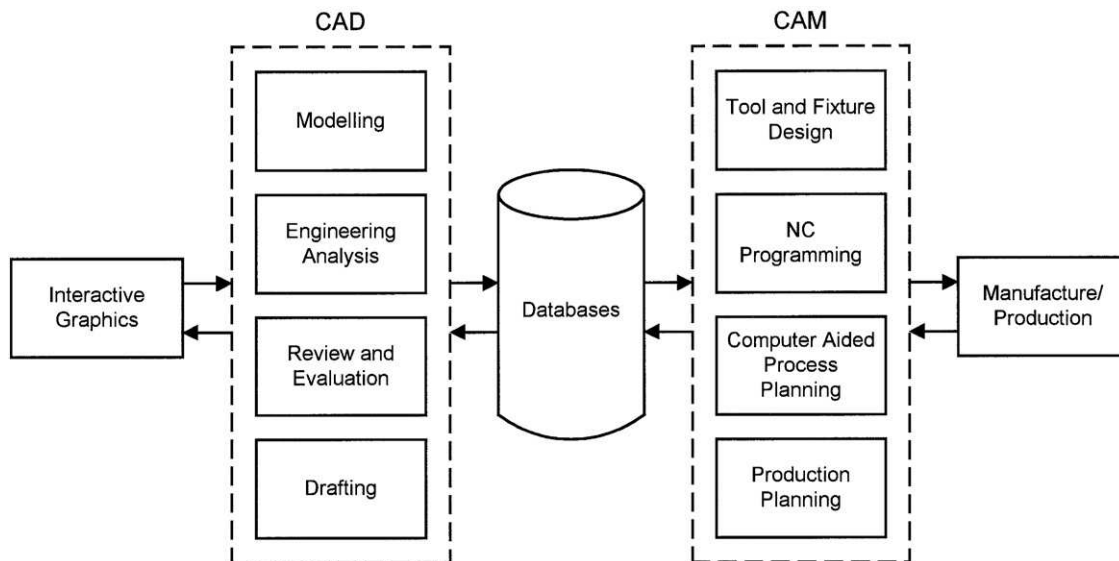


Figure 2.3.16

CAM systems are also used for generating machine tool programs, which vary in sophistication from simple 2D applications, such as turning or drilling, to more complex software programs for multi-axis operations.

A CAD/CAM system would include the following features: an interactive graphics system and the associated software for design; software packages for manufacture; and a common CAD/CAM database organised to serve both design and manufacture. The CAD interactive graphics system is likely to have facilities for engineering analysis, automatic drafting, design review, and evaluation and modelling. The CAM software might include packages for tool and fixture design, NC part programming, automated process planning, and production planning (Groover, 1984). Figure 2.3.16 shows the desired features of a typical CAD/CAM system.

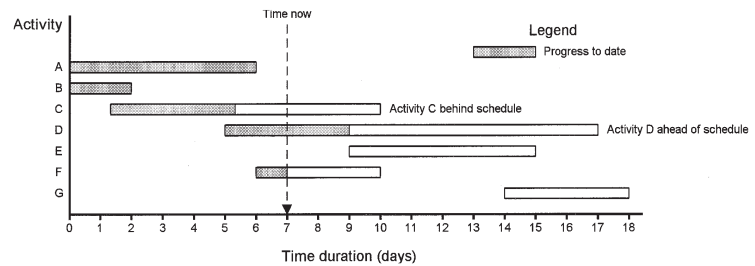


Figure 2.3.17

### Computer-aided planning and resource scheduling

It is not the intention here to discuss in any detail the management techniques available for planning and scheduling, these are discussed fully in Chapter 1, Business Management Techniques. However, in order to see how the computer may help us to plan and schedule a design project there is set out below, a brief reminder of one or two important *time planning* methods which you need to be acquainted with in order to *plan the design process*.

As you may be aware, the simplest planning aid is the Gantt chart, which is a form of bar chart where each bar represents an individual activity necessary to complete a project. The bar shows the scheduled start and finish time for each activity and the vertical arrows indicated their current status (Figure 2.3.17).

To produce a bar chart for a particular design project you would need to:

- analyse the project and break it down in terms of activities;
- estimate the time to perform each activity;
- place the activities onto the chart in strict time sequence, having determined which activities must be performed sequentially and which may be performed sequentially;
- ensure that the total time for the activities is planned to meet the specified completion date, if not adjust accordingly.

The primary advantage of the bar chart is that the plan, schedule, and progress of the project are all represented together on the one diagram. In spite of this important advantage, bar charts have not been too successful with complex engineering projects. The simplicity of the bar chart precludes sufficient detail to determine early slippage of scheduled activities. Also the bar chart is essentially a manual-graphical procedure, which is difficult to modify quickly, as project changes occur.

As a result of the disadvantages associated with bar charts, a *network-based methodology* had to be developed. The network model, like the bar chart, is a graphical representation of the planning process necessary for the project. It is based on the logic of *precedence planning*, which simply means that all activities which precede a given activity need to be completed, before the given activity may commence.

Figure 2.3.18 shows a typical *project network*. The *bracketed numbers* give an estimate of time for each activity (the units of time must be stated, hours, days, weeks, etc.). The *solid lines* denote activities which usually require resources to complete – time, manpower, equipment, the *dashed line* indicates a *dummy activity*, that is one which shows precedence only. Order of precedence is also shown for example, activities A<sub>3</sub> and B<sub>2</sub> must be completed before activity E is started. Whereas activities A<sub>1</sub>, B<sub>1</sub>, and D<sub>1</sub> may be commenced simultaneously. Each activity starts and finishes in a unique pair of *nodes* called *events*.

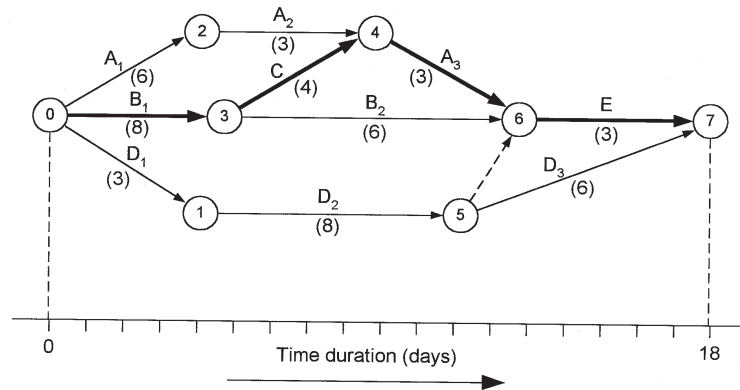


**Figure 2.3.18** Note: (i) Critical path ( $B \rightarrow C \rightarrow A_3 \rightarrow E$ ) total 18 days; (ii) dummy activity ( $5 \rightarrow 6$ ) precedence only

The events denote a point in time, their occurrence signifies the completion of all activities terminating with the event in question. For example, the occurrence of event 3 signals the completion of activities  $1 \rightarrow 3$  and  $2 \rightarrow 3$ . Time flows from the tail to the head of the arrow.

The essential ingredient in all networks is the *critical path* which by definition is a sequence of activities which determines the duration of the whole project. If we again consider Figure 2.3.18, the activities  $B_1$ ,  $C$ ,  $A_3$ , and  $E$  total 18 days, which is the same as the time duration for the whole project (events 0–6), so this pathway has no *float*, for this reason it is referred to as the *critical path*. You should refer to Chapter 1 for a full explanation of this process (see also Lockyer, 1991).

### Summary of planning and scheduling process

Many computer programs exist for analysing critical path networks, and some of these can schedule resources. These programs are particularly useful for assisting with the planning of the design process. Below is summarised the planning and scheduling process for a design project, which encompasses the use of computers:

- (1) Start by planning the work sequence *manually*, and producing a project network. Your sequence will include all the major steps given in Figure 2.2.1. Each of these major steps will be broken down into sub-activities. So for example, when considering *concept design*, this may include: information searching, lateral thinking exercises, group activities, the development and presentation of alternative design solutions, the evaluations of design solutions and selection of the optimum solution, and so on.
- (2) Estimate the time duration and resource requirements for each activity. Remember that some of the activities will be able to run concurrently, while others will need to be completed prior to the start of subsequent activities. Resource requirements should include: costs, materials, tooling, jigs and fixtures, machinery, production requirements, and human resources. Time scales will depend on the complexity and number of networks required for a particular design.
- (3) Load the network into a suitable computer system, with appropriate software installed. Provide the system with supplementary data, such as time units, required start and finish dates, and the milestones for the design project, if applicable.
- (4) Load the computer with information on total resource capacities, cost rates, calendar information, and all other relevant data. Safety margins with respect to time, physical, and human resources should be

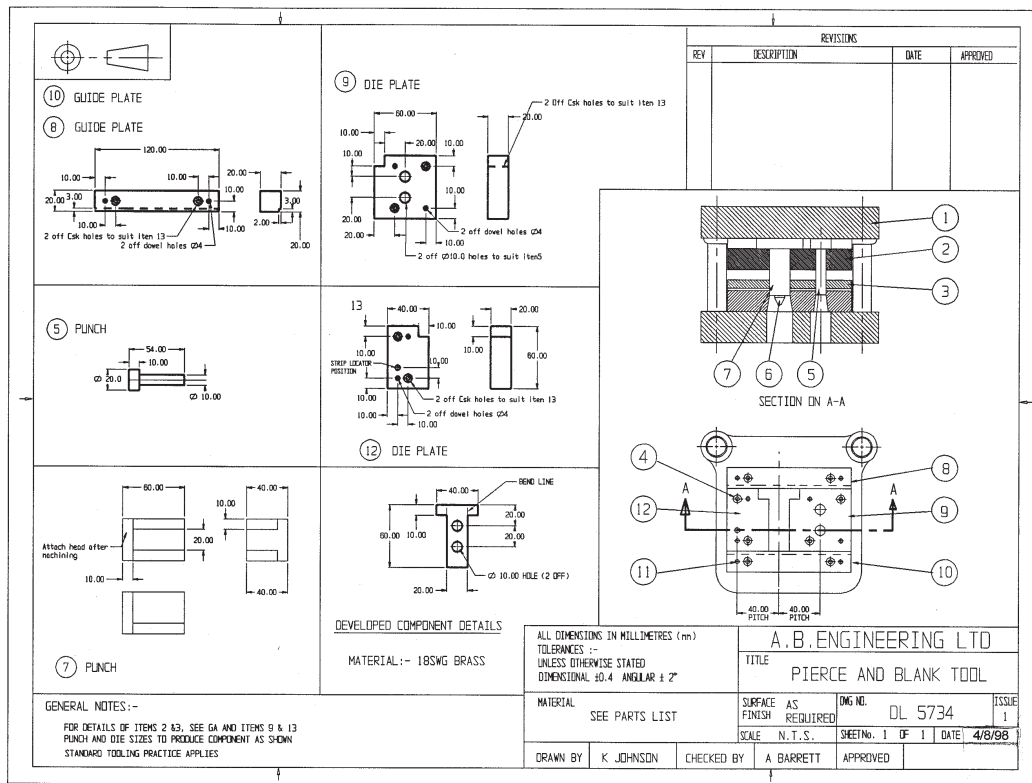


Figure 2.3.19

considered. So for example, you may not wish to declare all human resources that are available, but hold some back, so that in the event of the project running behind time you have additional human resources in reserve, that you can use to bring the project back on schedule.

- (5) Start the time-analysis run and rectify input errors, if they occur. Make sure that the network can be time-run, without error.
- (6) Carry out the resource scheduling run and produce the network print-outs, work schedules, and resource lists for monitoring and review by, marketing, administration, accounting, design, production and management staff, as required.
- (7) Continually update the computer, with information parameters, as they occur, repeating, as necessary, steps 1–6.

This concludes this short section on the use of computers in engineering design. Many aspects of the use of computers have been omitted, in the interests of space. No mention has been made of computer programming itself, software design QC, or dynamic system modelling such as vibration analysis or control system design. Spreadsheets and databases have not been mentioned in any detail, these packages are particularly useful in fulfilling the numerous administrative and management accounting functions that surround all engineering design projects. Nevertheless it is hoped that some indication of the importance of the computer to the design engineer, and the design process, has been emphasised.

We leave this section with a few examples of the variety and detail of engineering drawings and diagrams (Figures 2.3.19–2.3.21), which may be achieved using commercially available drafting packages. Figures 2.3.22, 2.3.24, and 2.3.25 illustrate the use of specialist software for the

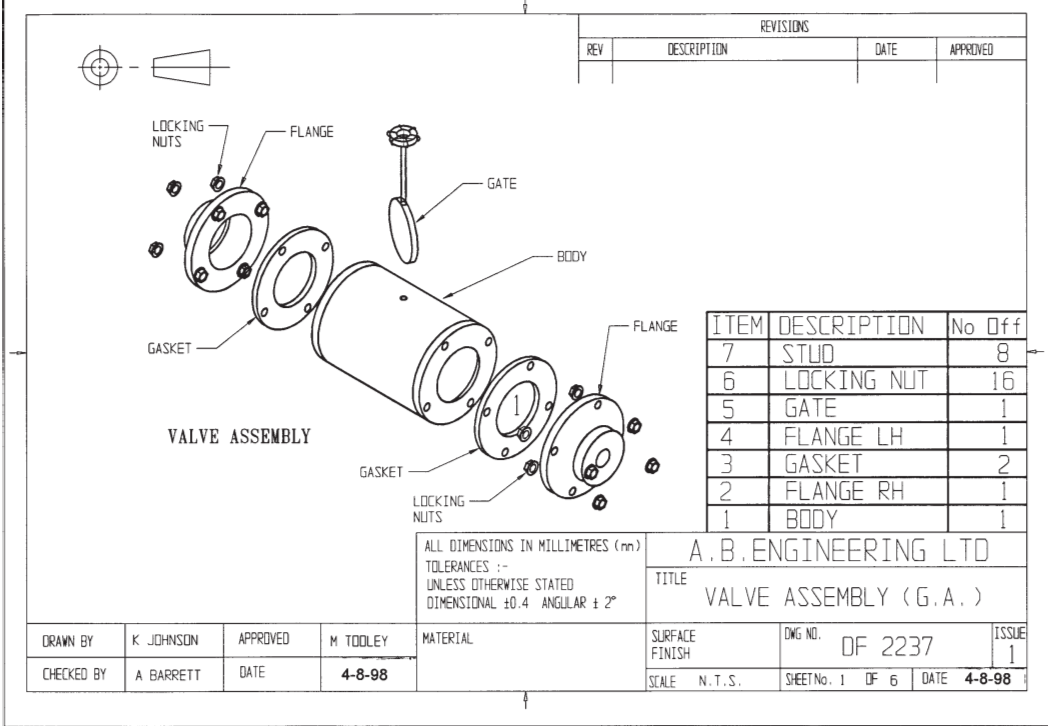


Figure 2.3.20

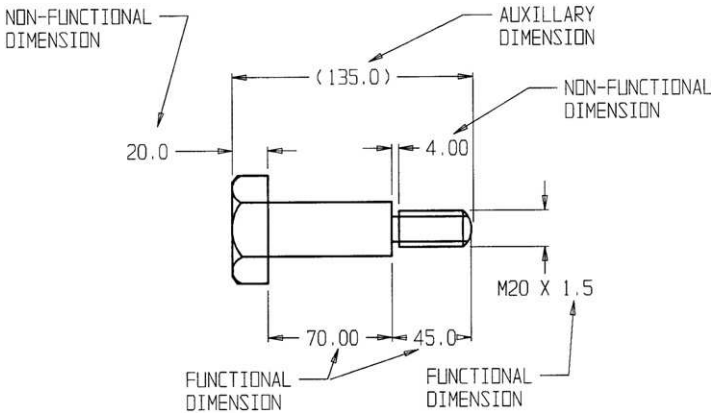


Figure 2.3.21

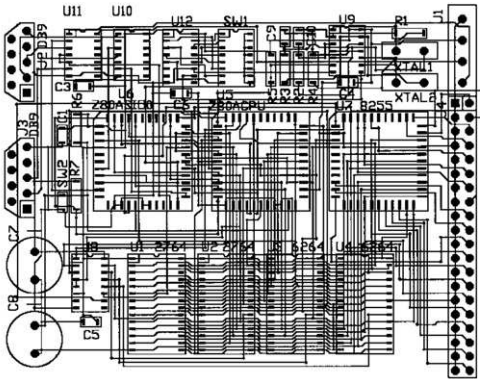


Figure 2.3.22

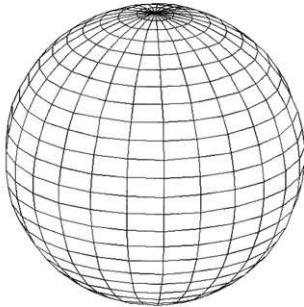
Create a surface plot of a unit sphere:

$N = 25$

$i := 0..N \quad \phi_i := i \cdot \frac{\pi}{N} \quad j := 0..N \quad \theta_j := j \cdot 2 \cdot \frac{\pi}{N}$

$X_{i,j} := \sin(\phi_i) \cdot \cos(\theta_j) \quad Y_{i,j} := \sin(\phi_i) \cdot \sin(\theta_j)$

$Z_{i,j} := \cos(\phi_i)$



X, Y, Z

Figure 2.3.23

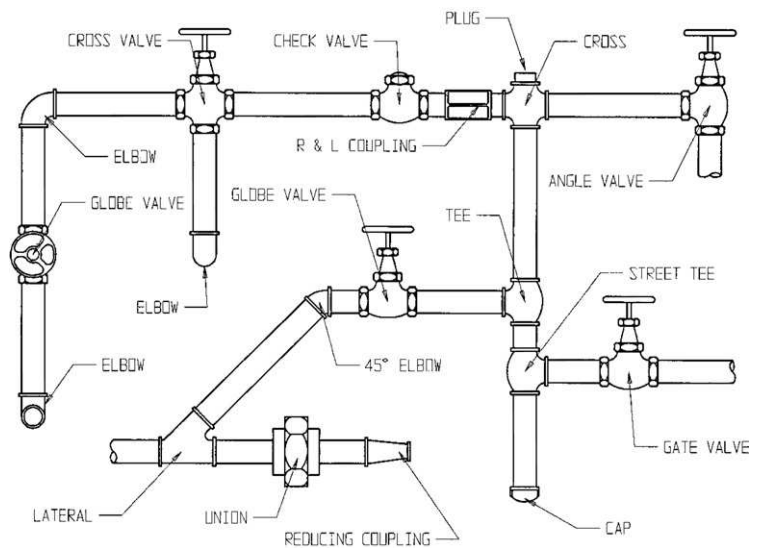


Figure 2.3.24

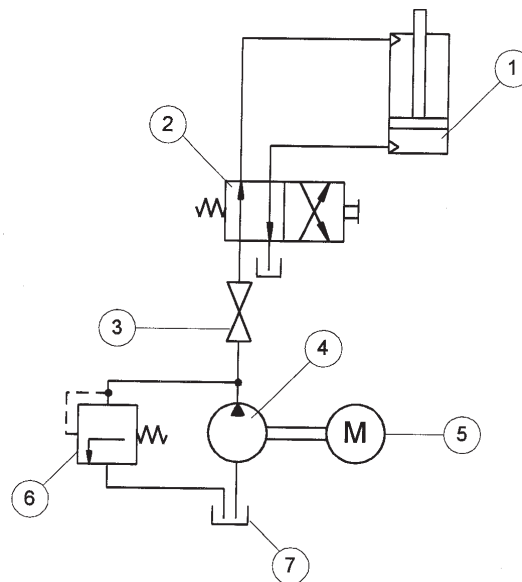


**Figure 2.3.25** KEY: (1) double-acting actuator; (2) 4/2-way directional control valve; (3) shut-off valve; (4) hydraulic pump; (5) motor; (6) pressure relief valve; and (7) tank (reservoir)

production of engineering system diagrams. The graphical output from a standard mathematical modelling package is shown in Figure 2.3.23.

### Questions 2.3.1

- (1) State the advantages of 3D drafting packages over their 2D counterpart.
- (2) Explain the operation of a CRT and explain the differences in operating principle between the DVST; vector refresh; and raster refresh systems.
- (3) Describe the likely computer storage systems required with:
  - (i) a stand-alone PC workstation;
  - (ii) a networked CAD system, with local workstations.

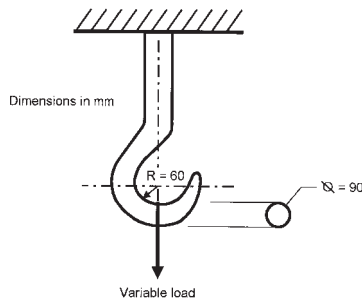


Figure 2.3.26

- (4) You have been tasked with the re-design of the front wing fairings of a new model of motor car, which encompasses new front light clusters. State, giving explicit reasons, which type of computer modelling package you would choose to assist you.
- (5) Investigate an FEA package that will assist in the design and development of the load-carrying gantry hook, shown in Figure 2.3.26. Consider the analysis in terms of the ability of the hook to withstand the loading conditions, from the point of view of strength and rigidity.
- (6) Using the design specification you produced for the 'hydraulic hose connector' in Question 2.1.2.
  - (i) *Draw-up a network* for the design process from final specification to the production of the agreed final solution.
  - (ii) With help from your tutor, load the network into a computer equipped with suitable software, carry out steps 1–5, of the summary of the planning and scheduling process given above.

## Summary of the design process

We conclude this short chapter with a summary which highlights the philosophy and thinking necessary during the design process in order to produce an engineering product or system, *on-time, at the right price and of the right quality*.

The first step in this process was to *analyse the design problem*, this included considerations such as the likely market, need, and economic viability of the design. The production of a suitable *design specification*, to suit the needs of the customer, whether internal or external, was emphasised as being of paramount importance.

In producing the design specification all the design parameters would need to be considered. These include: performance, ergonomics, manufacture and materials, maintenance, safety, installing and commissioning, as well as organisational factors such as legal implications, costs, transportation, quality, and so on.

In order to produce a number of possible *design alternatives*, the use of lateral thinking techniques was discussed, these include methods such as brainstorming, systematic searching, literature searching, and the gallery method to name but a few. Next, evaluation techniques might need to be used to select possible design alternatives. Finally, matrix methods, with appropriate ranking methods could be used to assist in selecting the *optimum design solution*. It is at this stage that the *design report* is written, and design ideas communicated to those in a position to make management decisions.

Use of the above techniques could be repeated for all phases of the design, assisting with the problems that might be encountered during the *layout, detail*, after sales and service phases of the project.

During the layout and detail phases, design engineers will be required to draw upon their knowledge of engineering principles and applications, materials, mathematics, machine component operation and selection, drafting techniques, and the nature of numerous technological advances,



## Another view

Mention has already been made of the complex thought processes that may be required to produce a successful design. The exploration and development of ideas is often more successful when a team approach is adopted. There will, of course, always be occasions when individuals are able to cry 'Eureka!' however, in the vast majority of cases a pre-planned and highly systematic approach to the whole process of design will be required to ensure success.

in order to produce the final design solution. Layout and detail design must also take into consideration the issues surrounding packaging and transportation, such as the need for modular construction, if appropriate.

Throughout the *whole design process* a constant check must be maintained to ensure that the design solution is produced within budget, and progresses on schedule. This may be achieved through the use of appropriate *computer software packages*, as discussed in Section 2.3.

Legal and safety aspects must be continually monitored to ensure that the design complies with all relevant design standards and, where appropriate, meets the requirements of European and other International legislation.

Finally, an efficient *after-sales service* must be set up, and appropriate help given with installing and commissioning the product, plant, or system.

In order to draw together some of the ideas expressed in this chapter we finish with one or two design questions, which will help you put into practice some of the concepts needed to ensure a systematic approach to design. To assist you with these exercises you should refer to the references listed at the end of this book.

## Test your knowledge

- (1) Your company intends to produce a steam iron, which is able to use ordinary tap water to fill the reservoir, without incurring fouling, scaling, or discolouration. The power rating for the iron should not exceed 1.8 kW and, the iron should be light and, capable of easy handling by both right- and left-handed users.  
Using the information provided in this chapter and any other information from the reference sources:
  - (i) Produce a design specification which includes at least 15 requirements and constraints.
  - (ii) Produce at least four preliminary design solutions, in the form of a concept sketch with accompanying explanation, for each possible solution.
- (2) A device to test the mechanical properties of toothpaste tubes is required by a manufacturer:
  - (i) Identify at least four possible sources of energy that may be used for the testing device.
  - (ii) Produce as many possible design alternatives as you can using different power sources, giving details of the design principles, and proposed operation for each design solution.
  - (iii) Using a suitable method, rank your possible design solutions.
  - (iv) Select your optimum solution and produce a general arrangement drawing (layout drawing), using a computer-aided drafting package.
- (3) There are currently a large number of microcomputers on the market, which utilise the *Pentium* processor. All these computers aim to fulfil the same function, but they differ in price and detail design features. After studying consumer magazines and/or other appropriate literature:
  - (i) Prepare an evaluation matrix for six different models with no more than 10 selection criteria.
  - (ii) Select the best product and explain in detail why, in your opinion, it is superior to the others.



# 3

# Engineering Science

## Summary

This unit aims to provide you with an understanding of the scientific principles that underpin the design and operation of modern engineering systems. Unlike other core units, this unit covers both mechanical and electrical principles. It thus provides a valuable introduction to engineering science for anyone who has not studied engineering before.

It is important to realise that this unit contains the basic underpinning knowledge required for further study of the two principles core units: Electrical and Electronic Principles and Mechanical Principles. This unit also provides a grounding for several of the specialist Options Units. It is divided into four sections, corresponding to the unit Outcomes, these deal with: static engineering systems, dynamic engineering systems, DC and single phase AC theory, and information and energy control systems.

## 3.1 STATIC ENGINEERING SYSTEMS

The study of static structures such as bridges, buildings, and frameworks, together with dynamic systems that include engines, power plant, electrical machines and air, sea, and rail transport, is of the utmost importance to engineers. The subject matter of static and dynamic engineering systems makes their study essential as a foundation for all engineers irrespective of their specialisation.

This section introduces the reader to some fundamental static's, that includes a study of, *simply supported beams and columns* (or struts). These structural members are commonly used in the production of bridges, buildings, and pylons, they are primarily designed to take bending and compressive loads. The selection of beams and columns, for a particular engineering purpose is also looked at. This section ends with a study of torsion, looking particularly at the shear stresses imposed on shafts and the effect that altering shaft geometry, has on the distribution of such stresses.

### Simply supported beams

#### ***Revision of fundamentals***

Before we introduce the idea of the simply supported beam and the methods we adopt to determine the shear forces and bending moments in such beams. We will first, review and revise the fundamental concepts of,

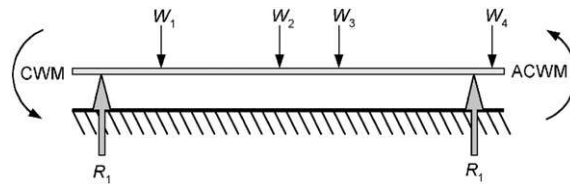


**Figure 3.1.1** Conditions for equilibrium

static equilibrium and the principle of moments, which are needed for any study of beams.

For a beam to remain in static equilibrium, two conditions must apply:

1. The upward forces must equal the downward forces
2. The sum of the clockwise moments (CWM) must equal the sum of the anticlockwise moments (ACWM).

You will know from Newton's second law that for every action there is an equal and opposite reaction. Thus for the beam shown in Figure 3.1.1, the actions are the points of application of the downward loads  $W_1$  to  $W_4$ . While the reactions, that balance these loads are shown acting vertically upwards through the supports ( $R_1$  and  $R_2$ ).

Also shown, is the other condition for equilibrium, where the algebraic sum of the moments created by the forces acting along the beam, must be in balance, that is must be equal. In other words, the principle of moments states that: *the sum of the CWM = the sum of the ACWM*.

You should also remember, that a *uniform beam* is such, that it has an equal cross-section along its whole length and is manufactured from a *consistent* material, that is one with an even density throughout. Under these circumstances, the *center of gravity* of the beam acts vertically down, through its geometric center.

When solving shear force and bending moment problems on beams, the first step is often to determine the reactions at the supports. If the beam is in static equilibrium (which it will be during our study of static's), then the method we use to find these reactions, is based on the application of the principle of moments. An example will demonstrate the technique.

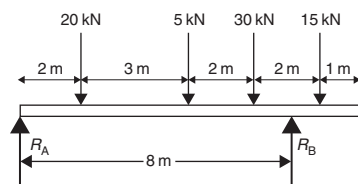


**Figure 3.1.2** Uniform beam for Example 3.1.1

### Example 3.1.1

A uniform horizontal beam is supported as shown in Figure 3.1.2. Determine the reactions at the supports  $R_A$  and  $R_B$ .

Now in order to produce one equation to solve one unknown, it is necessary to take moments about one of the reactions, in order to eliminate that reaction from the calculation. So taking moments about  $R_A$ , we get:

$$(2 \times 20) + (5 \times 5) + (7 \times 30) + (8 \times 15) = 8 \times R_B.$$

The above equation corresponds to the sum of the CWM = the sum of the ACWM,

$$\text{and } 40 + 25 + 210 + 135 = 8R_B$$

$$\text{then } R_B = 51.25 \text{ kN}$$

and using the first condition for equilibrium, that is, upward forces = downward forces, we get:

$$R_A + 51.25 \text{ kN} = 70 \text{ kN} \quad \text{so that } R_A = 18.75 \text{ kN.}$$

Instead of using the first condition we can take moments about  $R_B$ , to find the reaction at  $R_A$ . Then:

$$(1 \times 15) + (8R_A) = (6 \times 20) + (3 \times 5) + (1 \times 30)$$

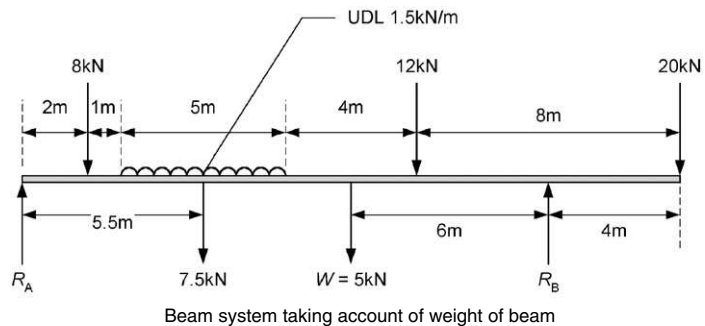
$$\text{and } 8R_A = 120 + 15 + 30 - 15$$

$$\text{then } R_A = 18.75 \text{ kN as before.}$$

Now it is not always the case that the loads on a beam are *point loads*. Beams can also be subjected to loads that are distributed for all, or part, of their length. These loads are known as, *uniformly distributed loads* (UDLs). For UDLs the whole mass of the load is assumed to act at a point through the center of the distribution.

### Example 3.1.2

For the beam system shown below, determine the reactions at the supports  $R_A$  and  $R_B$ , taking into consideration the weight of the beam.



So from what has been said, the UDL acts as a point load of magnitude  $(1.5 \text{ kN} \times 5 = 7.5 \text{ kN})$  at the center of the distribution, which is 5.5 m from  $R_A$ .

In problems involved with reaction it is essential to eliminate one reaction from the calculations because only one equation is formed and only one unknown can be solved at any one time. This is achieved by taking moments about one of the reactions and then, since the distance from that reaction is zero, its moment is zero and it is eliminated from the calculations.

So taking moments about A (thus eliminating A from the calculations), we get:

$$(2 \times 8) + (5.5 \times 7.5) + (10 \times 5) + (12 \times 12) + (20 \times 20) = 16R_B$$

$$\text{or } 651.25 = 16R_B$$

$$\text{so the reaction at B} = 40.7 \text{ kN.}$$

We could now take moments about B in order to find the reaction at A. However, at this stage, it is easier to use the fact that for static equilibrium:

upward forces = downward forces

$$\text{so } R_A + R_B = 8 + 7.5 + 5 + 12 + 20$$

$$R_A + 40.7 = 52.5$$

$$\text{and so the reaction at A} = 11.8 \text{ kN.}$$

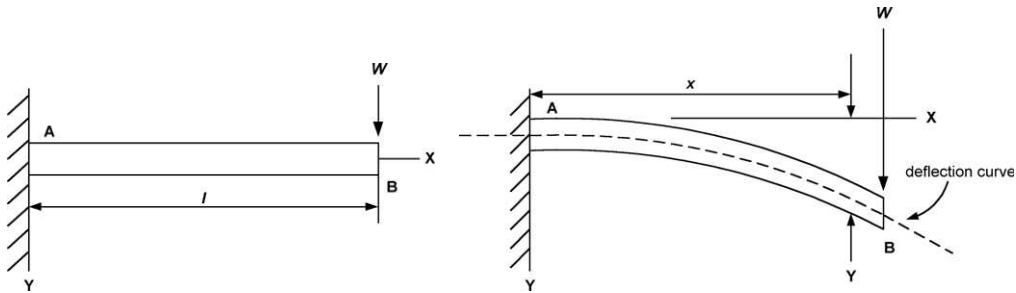


**Figure 3.1.3** Cantilever subject to symmetric bending

### Fundamental terminology

Before we look at shear force and bending in beams, we need to be familiar with one or two important terms and the methods we use to classify the type of beam we are considering.

The loads that act on a beam create *internal* actions, in the form of *shear stresses* and *bending moments*. The lateral loads that act on a beam cause it to bend, or flex, thereby deforming the axis of the beam into a curve (Figure 3.1.3) called the *deflection curve* of the beam.

We will only consider beams in this section that are symmetric about the  $xy$  plane, which means that the  $y$ -axis as well as the  $x$ -axis is an *axis of symmetry* of the cross-sections. In addition, all loads are assumed to act in the  $xy$  plane, that is the plane of the paper. As a consequence bending deflections occur in this same plane, which is known as the *plane of bending*. Thus the deflection curve AB of the beam in Figure 3.1.3, is a plane curve that lies within the plane of bending.

Beams are normally classified by the way in which they are supported. The *cantilever* (Figure 3.1.4(a)), is a beam that is rigidly supported at one end. The *simply supported beam* is either supported at its ends on rollers or smooth surfaces (Figure 3.1.4(b)) or by one of the ends being pin-jointed and the other resting on a roller or smooth surface (Figure 3.1.4(c)). It is the *simply supported beam* that we concentrate on in this section, as the heading suggests!

Figure 3.1.4(d), shows a built-in or *encastre beam*, where both of its ends are rigidly fixed. Finally, Figure 3.1.4(e), illustrates a simply supported beam with overhang, where the supports are positioned some distance from the ends.

Now, as already mentioned, when a beam is loaded reactions and resisting moments occur at the supports. While, shear forces and bending moments of varying magnitude and direction occur along the length of the beam. We now consider these shear forces and bending moments.

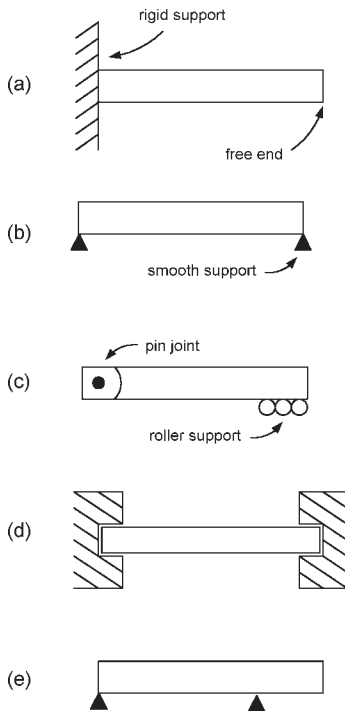


**Figure 3.1.4** Types of beam support

### Shear force and bending moment

We have mentioned shear force and bending moment but what are they and exactly what causes them? Consider the simply supported uniform beam shown in Figure 3.1.5.

If we cut the beam at X–Y and consider the section to the left of the cut (Figure 3.1.5(b)), then the effect of the upward force from the support  $R_A$ , is to try to bend the beam about the neutral axis AB, moving it upwards. However the portion of the beam to the right of the section X–Y prevents this movement by applying an equal and opposite force downwards (Figure 3.1.5(c)). The horizontal forces result from the beam being compressed above the neutral axis AB and stretched below the AB axis. It is these components of force that resist the bending and set up the bending moment within the beam material (Figure 3.1.5(c)).

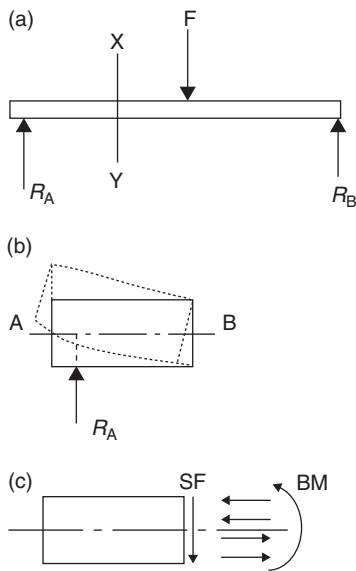


**Figure 3.1.5** Simply supported uniform beam

Remember that the shear force and resulting bending moment we are considering, is that which results from the flexing or bending of the beam created by the load.

### **Determination of shear force and bending moment**

We can use the ideas presented above to determine the bending moment and shear force set up in a beam, at any particular cross-section. These cross-sections being taken wherever there exists a support, point load or distributed load.

The *bending moment* at any section of a beam, just uses the principle of moments for its solution about a chosen cross-section. That is: *the algebraic sum of the moments of all external loads acting on the beam to one side only of the section.*

The shear force that acts at a section, only changes along the beam where a load normal to the beam, acts. Also, *the shear force is considered positive, when acting upwards and negative when acting downwards.*

#### **Example 3.1.3**

Ignoring the weight of the beam. Find the shear force and bending moment at AB and at CD, for the simply supported uniform beam shown below (Figure 3.1.6).

We first find the reactions at the supports. Then taking moments about  $R_A$ , gives:

$$(3 \times 10) + (8 \times 8) = 10R_B$$

$$\text{or } 10R_B = 94 \text{ and } R_B = 9.4 \text{ kN.}$$

Then using first condition for equilibrium,  $R_A + 9.4 = 18 \text{ kN}$  or  $R_A = 8.6 \text{ kN}$ .

Now the shear force acting at the section AB is simply that due to the reaction  $R_A$ , that is **SF = +8.6 kN**. The bending moment *to the left* of section AB is simply the moment due to the reaction that is., **BM =  $(2 \times 8.6) = 17.2 \text{ kN m}$** .

The SF to the left of section CD, is that due to the reaction which is +8.6 kN opposed by that due to the 10 kN load, which is acting down, then:

$$\text{SF} = 8.6 - 10 \text{ kN} = -1.4 \text{ kN.}$$

The BM to the left of the section CD is the algebraic sum of the moments, to the left of the section line i.e.,  $(6 \times 8.6) - (3 \times 10) \text{ kN m}$  or **BM = 21.6 kN m**.

*Note that it does not matter, which side of the section line we take moments, left or right, the bending moment will be the same!*

So for example, the BM to the right of CD is  $(-2 \times 8) + (4 \times 9.4) = 21.6 \text{ kN m}$ , as before.

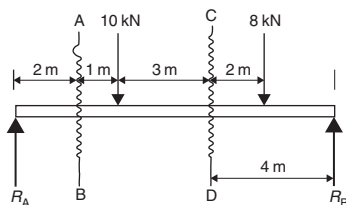


**Figure 3.1.6** Beam for Example 3.1.3

If instead of point load or loads, we have a UDL or combination of point loads and UDLs, we treat the UDL as a point load with its center at the mid-position of the UDL and then find shear force and bending moments, in the same manner as that given above.

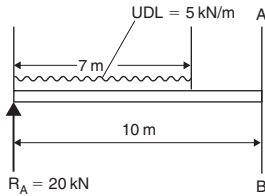


**Figure 3.1.7** Figure for Example 3.1.4

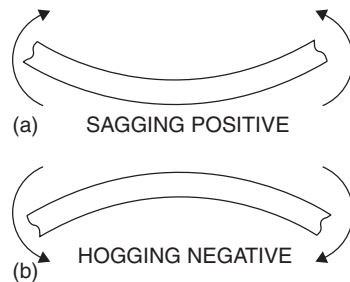


**Figure 3.1.8** Sign convention for bending moment

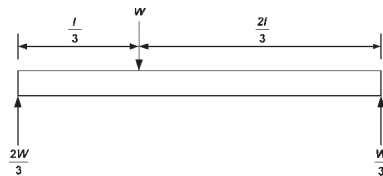


**Figure 3.1.9** Figure for Example 3.1.5 – simply supported beam with concentrated load

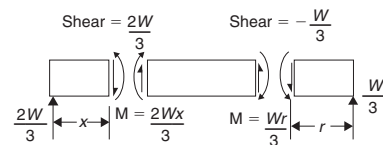


**Figure 3.1.10** Figure for Example 3.1.5 – free body diagram for cuts either side of the load

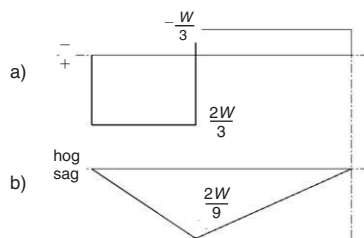


**Figure 3.1.11** Figure for Example 3.1.5 – shear force and bending moment diagram

### Example 3.1.4

Find the BM for the section AB, shown in Figure 3.1.7.

Then as before to find the BM to the left of the cut AB, we consider the algebraic sum of the moments to the left.

The UDL = 5 kN/m acting over 7 m, so the total load is  $(5 \times 7) = 35$  kN, and can be represented by a point load acting 3.5 m from the end of the beam or 6.5 m from AB.

Then taking moments at AB, we get:  $(10 \times 20) - (6.5 \times 35)$  kN m  
or **BM = -27.5 kN m.**

When determining shear force and bending moments it is convenient to use a sign convention. We have already used such a convention in the above example. *Bending moments* will be considered *positive*, if they compress the top section of the beam and stretch the lower section that is, the beam sags (Figure 3.1.8(a)) and *negative* when the loads cause the beam to stretch on its top surface or hog (Figure 3.1.8(b)).

The corresponding *shear forces* may be thought of as positive when the loads on the section of interest causes the beam to sag and negative when the loads cause the beam to hog. This convention is illustrated in Example 3.1.6.

The next example generalises the method for finding shear force and bending moment.

### Example 3.1.5

Draw the shear force and bending moment diagrams for the beam shown in Figure 3.1.9.

We have already been given the reactions at the left-hand and right-hand supports, they are  $2W/3$  and  $W/3$ , respectively. So we may now determine the shear force and bending moment values by producing our free body diagram, making cuts either side of the load at the arbitrary points  $x$  and  $r$  (Figure 3.1.10).

For vertical equilibrium, using our sign convention, then at left-hand cut the shear force  $F = 2W/3$  and at right-hand cut  $F = -W/3$ , note also the direction of the shear force pairs at the cuts, either side of the transverse load.

At the left-hand end of beam the reaction is  $2W/3$  (positive), at the left hand cut the shear force is again  $2W/3$ , but this changes on the right-hand side of the load to  $-W/3$ . So at the load the shear force on the beam is the sum of these two shear forces, as shown in Figure 3.1.11(a).

Note that the shear force only changes where a transverse load acts. Figure 3.1.11(b) shows the resulting bending moment diagram, where the bending moment under the load is obtained by putting  $x = \frac{l}{3}$  or  $r = \frac{2l}{3}$  into the equation  $M = \frac{2Wx}{3}$  or

$$M = \frac{Wr}{3}, \text{ respectively.}$$

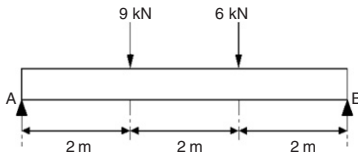


**Figure 3.1.12** Figure for Example 3.1.6 – simply supported beam with two concentrated loads

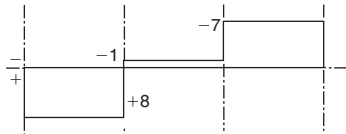


**Figure 3.1.13** Figure for Example 3.1.6 – shear force distribution

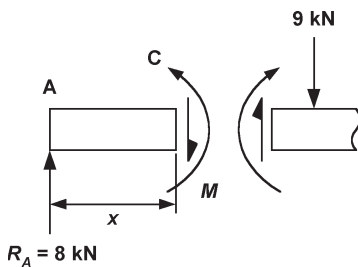


**Figure 3.1.14** Figure for Example 3.1.6 – free body diagram

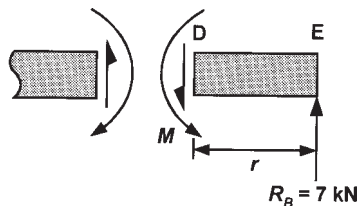


**Figure 3.1.15** Figure for Example 3.1.6 – bending moment diagram

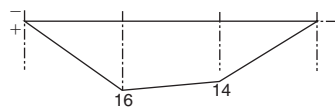


**Figure 3.1.16** Figure for Example 3.1.6 – bending moment diagram

## Shear force and bending moment diagrams

So far we have concentrated on finding shear force and bending moments at a single cut point, for a section. It is beneficial for engineers to know how the shear forces and bending moments are distributed across the whole length of the beam. To do this we construct shear force and bending moment diagrams, that give the whole picture. Computer software may also be used to establish these forces and moments, for more complex loading cases.

### Shear force and bending moment diagrams for concentrated loads

When constructing these diagrams all we need do is extend the techniques we have already mastered for finding the SF and BM at a section cut. Remembering that the shear force will only vary, where there is a transverse (normal) point load acting on the beam. The technique for beams subject to point loads is illustrated in the next example.

#### Example 3.1.6

Draw the shear force and bending moment diagram for the beam shown in Figure 3.1.12.

The first step is to calculate the reactions  $R_A$  and  $R_B$ .

The reactions may be found in the normal manner by considering rotational and vertical equilibrium. Then taking moments about  $R_A$  gives:

$$(9 \times 2) + (6 \times 4) = 6R_B \quad \text{and} \quad R_B = 7 \text{ kN.}$$

Also for vertical equilibrium  $R_A + 7 = 15$  and so  $R_A = 8 \text{ kN}$ .

Now considering the shear forces at the 9 kN load, and remembering that the shear force only changes where a transverse loads acts. Then on the left-hand side of the 9 kN load we have a positive shear force  $R_A = +8 \text{ kN}$ . This is opposed at the load by the 9 kN force acting in a negative direction, so the shear force to the right of this load is  $= -1 \text{ kN}$ . This shear force continues until the next transverse load is met, in this case the 6 kN load, again acting in a negative direction, producing a net load of  $-7 \text{ kN}$  which continues to  $R_B$ . This shear force distribution is clearly illustrated in Figure 3.1.13.

The bending moment diagram can be easily determined from the shear force diagram by following the argument given in Example 3.1.5. If we consider the free body diagram for AC (Figure 3.1.14), then taking moments about C gives:

$$R_A x - M = 0 \quad \text{and at the 9 kN load where } x = 2 \text{ m we have } (8 \times 2) = M = 16 \text{ Nm (sagging).}$$

Similarly, considering the free body diagram for DE (Figure 3.1.15) and taking moments about D, when  $r = 2$ , then from  $-R_B r + M = 0$ , we have  $-(7 \times 2) = -M = 14 \text{ Nm}$ .

These results can now be displayed on the bending moment diagram (Figure 3.1.16).

Note that the loads produce a sagging moment downwards, which intuitively you would expect.

## Mathematics in action

Take a small element, of length  $dx$  from a beam carrying a load  $w$  (downwards) per unit length (Figure 3.1.21).

Then we can see that if  $F$  is the shear force due to load at the point  $x$ , then at the point  $x + \delta x$  the shear is  $F + \frac{dF}{dx} \times \delta x$

and similarly if  $M$  is the bending moment at  $x$ ,  $M + \frac{dM}{dx} \times \delta x$  is the bending moment at the point  $x + dx$ .

Now for vertical equilibrium we have (adding downward forces)

$$-F + \omega \delta x + \left( \frac{F + \frac{dF}{dx} \delta x}{dx} \right) \delta x = 0.$$

(where  $\omega$  is the distributed load) giving

$$\frac{dF}{dx} = -\omega. \quad (3.1.1)$$

This indicates that a downward load  $\omega$ /unit length gives a negative rate of shear.

Next we consider the tendency of the element to rotate, that is equilibrium of moments (taken clockwise about the point  $x$ ).

Then:

$$M + \omega \delta x \frac{\delta x}{2} + \left( \frac{F + \frac{dF}{dx} \delta x}{dx} \right) \delta x - \left( \frac{M + \frac{dM}{dx} \delta x}{dx} \right) \delta x = 0$$

giving

$$M + \omega \frac{(\delta x)^2}{2} + F \delta x + \frac{dF}{dx} (\delta x)^2 - M - \frac{dM}{dx} \delta x = 0$$

neglecting second order terms in  $\delta x$  this gives:

$$F \delta x - \frac{dM}{dx} \delta x = 0 \quad \text{or} \quad \frac{dM}{dx} = F. \quad (3.1.2)$$

Check that you follow the above argument! This mathematical argument leads to two important results.

- (i) For a maximum value of  $M$ ,  $dM/dx = 0$ , thus  $F = 0$  (from Equation 3.1.1), so if we have already drawn an SF diagram we know that the position of the maximum BM is at the point where the SF is zero.
- (ii) If we integrate Equation 3.1.2 we get  $M = \int F dx$ , so we can find the values for the BM diagram by adding numerically the areas of the SF diagram (i.e. integrating) and this may well be quicker than plotting numerical values found from the algebraic expression for the BM. Also, from Equation 3.1.1,  $S = -\int \omega dx$  and so substituting this into Equation 3.1.2 we get:

$$\int \omega dx = -\frac{dM}{dx} \quad \text{or} \quad \omega = \frac{d^2 M}{dx^2}. \quad (3.1.3)$$



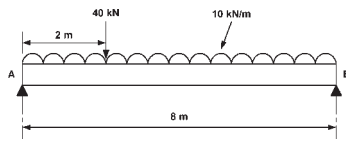


**Figure 3.1.17** Beam subject to concentrated load and UDL

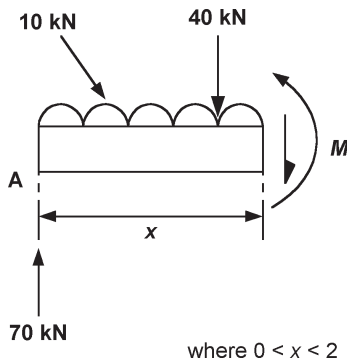


**Figure 3.1.18** Free body diagram of beam to left of cut

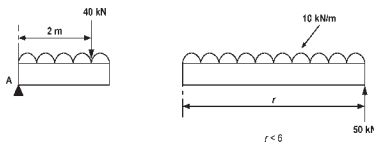


**Figure 3.1.19** Free body diagram of beam to right of cut

### Shear force and bending moment diagrams for UDL

We will now consider drawing shear force (SF) and bending moment (BM) diagrams for a beam subject to a UDL. These loads are normally expressed as force per unit length.

The beam shown in Figure 3.1.17 is subject to a UDL of 10 kN/m in addition to having a point load of 40 kN acting 2 m from A.

As usual we first find the reaction at A and B. To do this you will remember from your previous work, to treat the UDL as a point load acting at the center of the uniform beam. This gives a load of 80 kN acting 4 m from A. Then the reactions may be calculated as  $R_A = 70$  kN and  $R_B = 50$  kN. We now cut the beam and, consider the free body on the left of the cut (Figure 3.1.18).

Our first cut is made at a distance  $x$  from the left-hand support and to the left of the point load. We consider the left-hand part because it is the simpler of the two (you would of course get the same values for SF and BM if you considered the right-hand part).

Then, from the convention used in Figure 3.1.18 we have:

- (i)  $SF = (70 - 10x)$  kN;
- (ii) and taking moments about the cut gives a  $BM = \left( 70x - \frac{10x^2}{2} \right)$  kN m sagging.

Next consider a cut to the right of the point load (Figure 3.1.19).

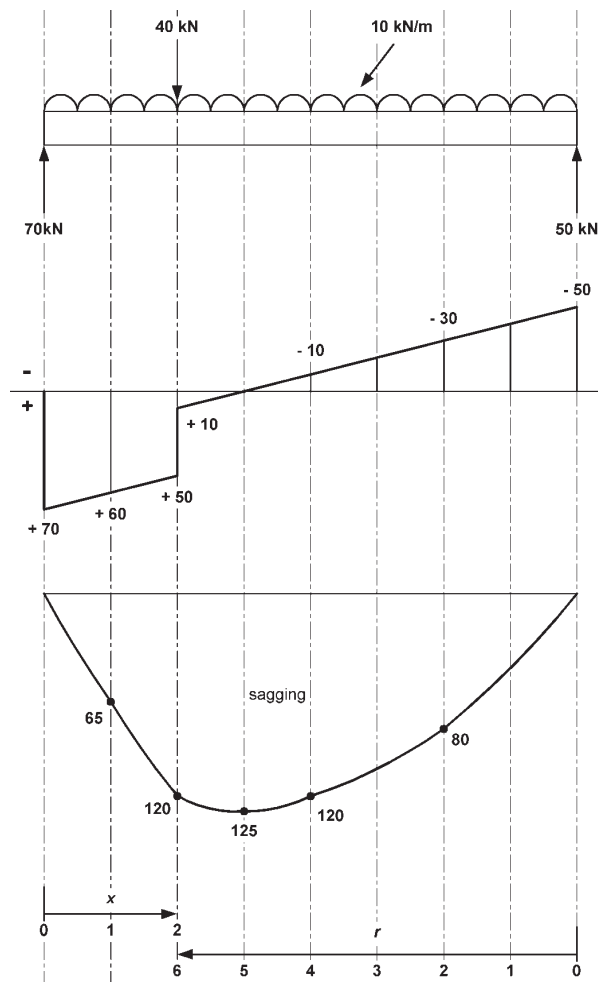
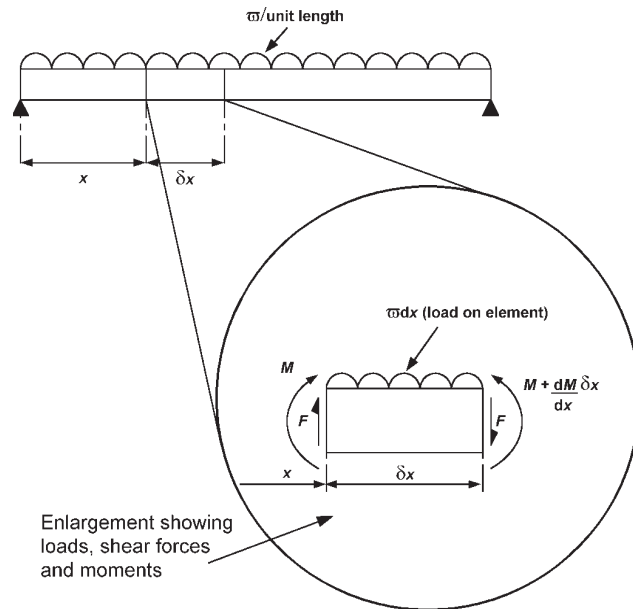


**Figure 3.1.20** Shear force and bending moment diagram for UDL

**Figure 3.1.21** Small element from beam carrying a UDL



Again using the convention shown, we have:

- (i)  $SF = (10x - 50) \text{ kN}$ ;
- (ii)  $\left( 50x - \frac{10x^2}{2} \right) \text{ kNm sagging}$ .

You will note that the curve given by the expression for BM is a parabola. This is true for any section of a beam carrying only an UDL.

If we plot the SF at A and then subsequently, at 1 m intervals for the whole length of the beam we will obtain the SF and BM diagrams illustrated in Figure 3.1.20.

You should ensure that you follow the above argument by verifying one or two of these SF and BM values for yourself.

Plotting diagrams point by point like this is a tedious business! Fortunately there are ways around it, based on relations between load, shear force, and bending moment that we have not yet considered (Figure 3.1.21).

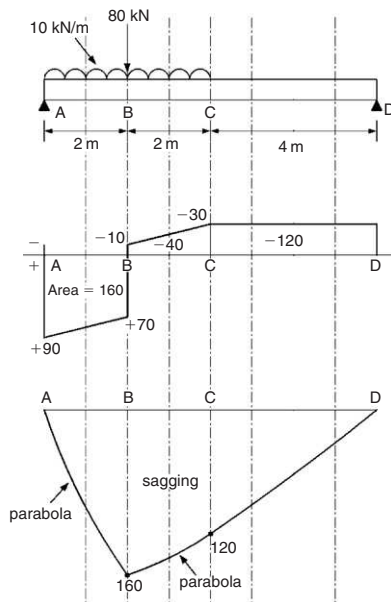


**Figure 3.1.22** Figure for Example 3.1.7

### Example 3.1.7

Consider a beam with loads as shown in Figure 3.1.22 and draw the SF and BM diagrams for the beam.

Again the reactions may be found in the usual way, where  $R_A = 90 \text{ kN}$  and  $R_B = 30 \text{ kN}$  both acting upwards.

To obtain the SF diagram we start at the left-hand end of beam where  $R_A = +90 \text{ kN}$ . Moving along to point B and adding up the loads we have,  $+90 - \int w dx$  which becomes  $90 - 20 = +70$ . At point B a downward load of 80 acts, giving  $+90 - 20 - 80 = -10$ . At D we have an SF of  $-30$ , resulting from the reaction  $R_B$ , this is constant until it meets the

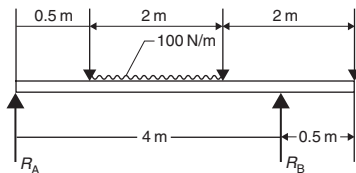


Figure 3.1.23 Beam system

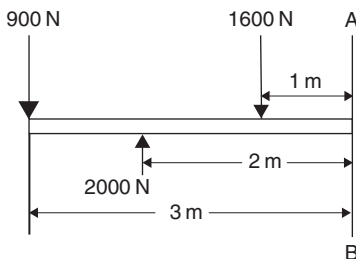


Figure 3.1.24 Section AB

Figure 3.1.25 Simply supported beam

distributed load at C. Now moving from C to B the SF changes constantly from  $-30$  to  $-10$ , as indicated on the diagram. You will note that on the SF diagram distributed load gives a straight sloping line between A and B, and between B and C.

To obtain the BM diagram we again start at A, where the BM is zero. *Adding up the area of the shear force diagram* between A and B gives a BM at B of  $+160$ . Continuing we may say that the sum of the areas between A and C is given by  $\int_A^C F dx = 160 - 40 = 120$ . At D it can easily be seen that the BM  $= 160 - 160 = 0$ , by again summing the areas on the SF diagram between A and D.

Note that the SF is zero at point B, so the maximum BM occurs at B. The slope of the BM diagram is everywhere given by the ordinate of the SF diagram and because the SF varies linearly, the BM diagram is parabolic (between A and B, B and C), while between C and D where the SF is constant the BM is a straight line, remembering your integration!

### Problems 3.1.1

1. Find the reactions at the supports for the beam system shown (Figure 3.1.23).
2. Find the value of the BM for the section AB, shown (Figure 3.1.24).
3. Find the reactions at the supports and the bending moment at the section AB, for the simply supported beam shown (Figure 3.1.25)
4. Draw the SF and BM diagrams for the simply supported beam shown (Figure 3.1.26)
5. A beam is simply supported over a span of 8 m and overhangs the right-hand support by 2 m. The whole beam carries a UDL of  $10 \text{ kN/m}$ , together with a point load of  $120 \text{ kN}$ ,  $2.5 \text{ m}$  from the left-hand support. Draw the SF and BM diagrams showing appropriate numerical values.

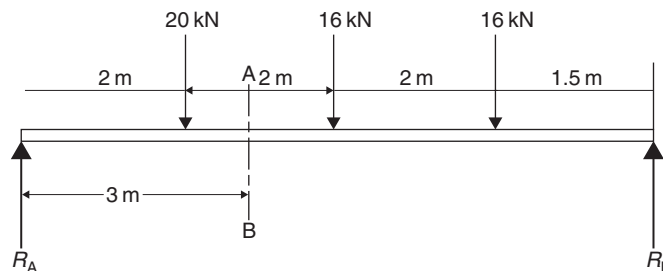
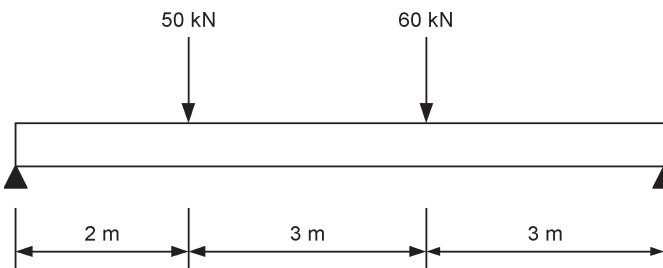


Figure 3.1.26 Figure for Problem 3.1.1(4)



## Properties and selection of beams and columns

In this section we look first at the theory of bending related to beams,

where the important relationship  $\frac{M}{I} = \frac{\sigma}{y} = \frac{E}{R}$  is developed. We then

consider the important properties of beams and the use of these properties. We then look in particular at the Second Moment of Area, for a number of beam cross-sections, before looking at the properties needed to select a beam for a particular function.

In the last section we move on to **Columns** or Struts, and consider the concentric (axial) loading of columns and the properties needed for their selection.

### Engineers' theory of bending

When a transverse load is applied to a beam, we have seen that bending moments are set up which are resisted internally by the beam material. If the bending moment resulting from the load is sufficient, the beam will deform, creating tensile and compressive stresses to be set up within the beam (Figure 3.1.27). Between the areas of tension and compression there is a layer within the beam, which is unstressed termed the *neutral layer* and its intersection with the cross-section is termed the *neutral axis*.

Engineers' theory of bending is concerned with the relationships between the stresses, the beam geometry and the curvature of the applied bending moment. In order to formulate such relationships we make several assumptions, which simplifies the mathematical modelling. These assumptions include:

- (i) the beam section is symmetrical across the plane of bending and remains within the plane after bending (refer back to the introduction);
- (ii) the beam is straight prior to bending and the radius of bend is large compared with the beam cross-section;
- (iii) the beam material is uniform (homogenous) and has the same elastic modulus in tension and compression (this is not the case, for example, with ceramic materials).

### Deformation in pure bending

Consider the beam subject to pure bending (Figure 3.1.28), which shows an exploded view of the beam element where the neutral axis, discussed above, is clearly seen. Figure 3.1.28a shows the general arrangement of the beam before and after bending, where we consider the deformations

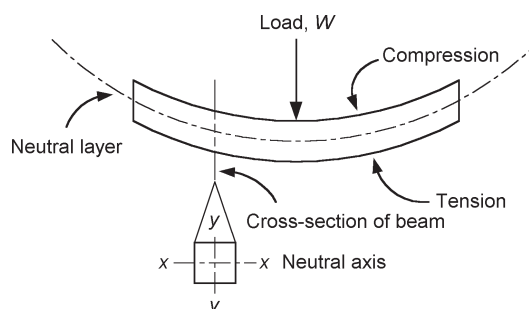


**Figure 3.1.27** Deformation of beam, resulting in tensile and compressive stresses

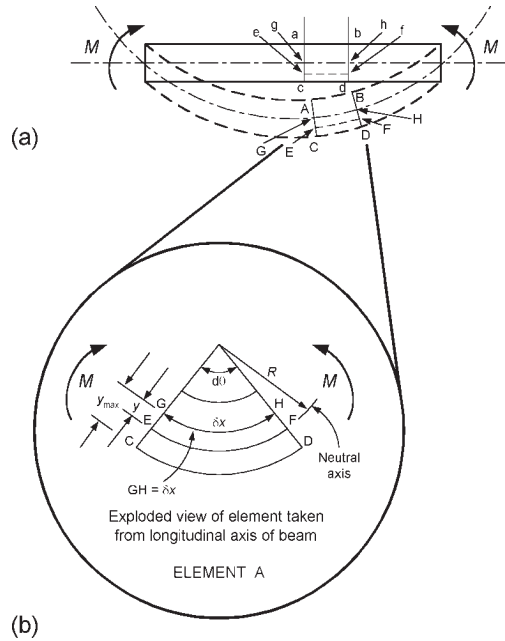


**Figure 3.1.28** Deformation in pure bending

that occur between the sections  $ac$  and  $bd$ , which are  $\delta x$  apart, when the beam is straight. The longitudinal fibre of the beam material  $ef$  which is at a distance  $y$  from the neutral axis has the same initial length as the fibre  $gh$  at the neutral axis, prior to bending.

Now during bending we can see from Figure 3.1.28 that  $ef$  stretches to become  $EF$ . However,  $gh$ , which is at the neutral axis is not strained (since stress is deemed to be zero) and so it has the same dimensions when it becomes  $GH$ . So, by simple trigonometry, when  $R$  equals the radius of curvature of  $GH$  we may write:

$$GH = gh = \delta x \text{ where } \delta x = R d\theta \text{ (}\theta \text{ in radians) and } EF = (R + y)d\theta.$$

We also see from the figure that the longitudinal strain of  $EF$  is given by:

$$\epsilon_x = \frac{EF - ef}{ef}.$$

Now since we know that prior to bending  $ef = gh$  and that after bending because  $gh$  is on the neutral axis  $gh = GH$ , then  $GH$  must equal  $ef = R d\theta$  (from above). Also, since  $EF = (R + y)d\theta$  we have:

$$\epsilon_x = \frac{(R + y)d\theta - R d\theta}{R d\theta}$$

$$\text{or } \epsilon_x = \frac{R d\theta + y d\theta - R d\theta}{R d\theta}$$

and on division by  $d\theta$

$$\epsilon_x = \frac{y}{R}. \quad (3.1.4)$$

We also showed that  $dx = R d\theta$  or  $R = dx/d\theta$ , and on substitution into Equation 3.1.4 we get:

$$\epsilon_x = \frac{d\theta}{dx} y. \quad (3.1.5)$$

From our earlier work in this chapter on Poisson's ratio we established relationships for three-dimensional strains. So if we consider the strains

$\varepsilon_y$  and  $\varepsilon_z$  which are transverse to the longitudinal strain  $\varepsilon_x$  we can show that they may be determined from the relationship:

$$\varepsilon_y = \varepsilon_z = -\frac{\nu\sigma}{E}x$$

where  $\nu$ , you will remember, is Poisson's ratio.

*Note:* If you find the verification of these formula difficult do not worry, just concentrate on the results.

### **Stresses due to bending**

We have already shown (Equation 3.1.4) that  $\varepsilon_x = \frac{y}{R}$  and knowing that

$\varepsilon = \frac{\sigma}{E}$  then substituting into the above equation yields the relationship:

$$\frac{\sigma}{E} = \frac{y}{R}. \quad (3.1.6)$$

We can also show, by considering the cross-section of element A in Figure 3.1.28 that the bending moment  $M$  imposed on the beam at the element is related to the stress on the element, the distance  $y$  from the neutral axis and the second moment of area  $I$  by:

$$\frac{M}{I} = \frac{\sigma}{y}. \quad (3.1.7)$$

The second moment of area  $I$  is a measure of the bending efficiency of the structure, in that it measures the resistance to bending loads. The greater the distance and amount of mass of the structure away from the bending axis, the greater the resistance to bending. The derivation of  $I$  will be found in Chapter 6 under the applications of the calculus, which you should consult if you have difficulties.

If we combine Equations 3.1.6 and 3.1.7 we can write:

$$\frac{M}{I} = \frac{\sigma}{Y} = \frac{E}{R} \quad (3.1.8)$$

The above relationship is known as the general bending formula or *engineers' theory of bending*.

For a beam the maximum bending stress ( $\sigma_{\max}$ ) will occur where the distance from the neutral axis is a maximum ( $y_{\max}$ ) and from Equation 3.1.8 we may write:

$$M = \frac{I\sigma_{\max}}{y_{\max}}$$

the quantity  $I/y_{\max}$  depends only on the cross-sectional area of the beam under consideration, and is known as the *section modulus*  $Z$ , therefore:

$$M = Z\sigma_{\max}.$$

*Note:* Tables of section module  $Z$ , together with other important properties of beams (discussed next), may be found by consulting past and present British Standards (BS5950-1: 2000 and BS 4-1 1993).

**Example 3.1.8**

Figure 3.1.29 shows a simplified diagram of a cantilever lifting gantry, 1.8 m long, designed to raise loads of up to 3 kN. If the maximum design stress of the cantilever beam material is 250 MPa. Determine a suitable diameter for the bar.

Now the design load is given as 3 kN, however, in design work we should always consider a factor of safety, in view of the fact that the suspended loads are involved we must ensure the complete safety of the operators so we will choose a factor of safety of 2. Thus the maximum design load = 6 kN.

The maximum bending moment for the cantilever beam is furthest from the support, in this case the maximum BM =  $6 \times 1.8 = 10.8 \text{ kN m}$ .

Now the diagram shows a circular cross-section beam. We need to calculate the second moment of area  $I$  for this beam. This is given by the formula:

$$I = \frac{\pi d^4}{64} \text{ (refer to Chapter 6).}$$

The maximum allowable design stress ( $\sigma_{\max} = 250 \text{ MN/m}^2$ ) on the beam occurs at the maximum distance from the neutral axis ( $y_{\max}$ ), in this case at distance  $d/2$ .

So using engineers' theory of bending

$$\text{where } \frac{M}{I} = \frac{\sigma}{y} \text{ then:}$$

$$\frac{10\,800}{\pi d^4/64} = \frac{250 \times 10^6}{d/2}$$

from which we get  $d = 76 \text{ mm}$ .

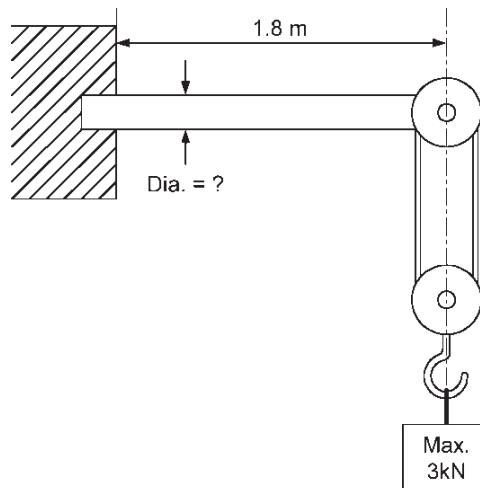


**Figure 3.1.29** Cantilever lifting gantry

**Test your knowledge**

- (1) A simply supported beam (Figure 3.1.30) carries the UDL as shown. The maximum allowable bending stress for the beam is 120 MPa.
  - (i) Construct the shear force diagram and so determine the maximum bending moment.
  - (ii) Calculate the required section modulus  $Z$ . Ignore the weight of the beam.
- (2) A simply supported beam (Figure 3.1.31) carries the UDL as shown:
  - (i) Construct the shear force diagram.
  - (ii) Sketch the bending moment diagram.
  - (iii) Determine the maximum bending moment.
  - (iv) Given that the section modulus for the beam is  $3500 \text{ cm}^3$ , determine the allowable bending stress. Ignore the weight of the beam.

**Centroid and second moment, of area**

Beams have to resist the loads which act upon them, causing high stresses due to these loads and resulting bending moments. The resistance of a beam to bending moments and shear forces depends on:

- the inherent properties of the beam material, such as; homogeneity, ultimate tensile strength and elastic modulus, and
- the shape and size of the beam section.

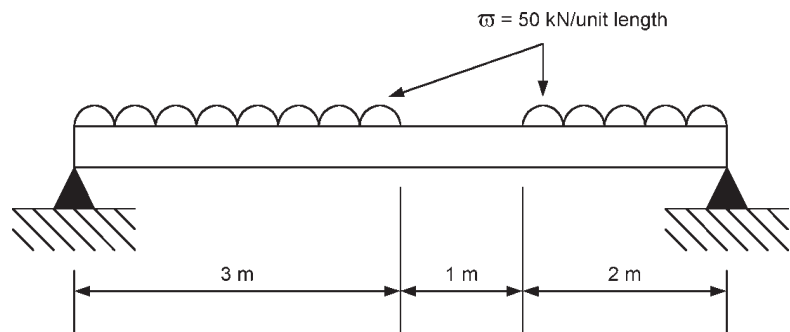


**Figure 3.1.30** Figure for Test your knowledge (1)

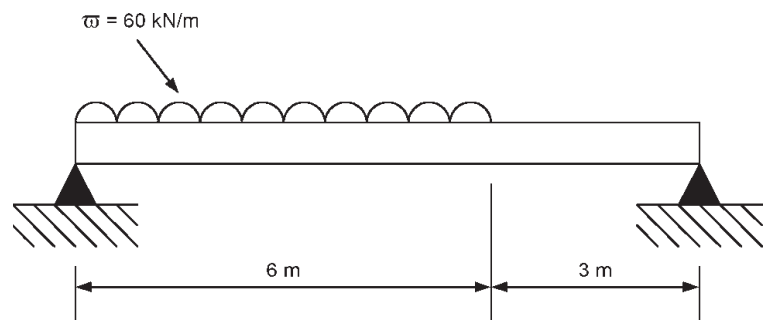


**Figure 3.1.31** Figure for Test your knowledge (2)



Properties of a beam which can be attributed to *their shape* include: cross-sectional area, *centroid of area* (their center of gravity), *second moment of area* or moment of inertia and their *section modulus*  $Z$ , which was mentioned earlier.

The *centroid of area* is, for calculation purposes, the point at which the total area is considered to act. While the *second moment of area* of a beam, or for that matter any shape, is a measure of the resistance to bending of that shape.

The mathematical derivation of one or two centroids of area and second moments of area are given in the engineering applications of the calculus section, starting on page 503, with centroids. Here, we concentrate on the practical techniques necessary to determine these properties.

### **Calculating centroids of area**

To determine the center of gravity of a particular shaped cross-section, we adopt the following procedure.

1. Divide the shape into its component parts
2. Determine the area of each part
3. Assume the area of each part to act at its center of gravity and
4. Take moments about a convenient point or axis, to determine the center of gravity (centroid) of the whole area.

*Note:* If we were finding the *moment of inertia of a body*, or the *second moment of mass of a body*, then the center of gravity of the weight of the whole body would be determined. The moment of inertia or second moment of mass is analogous with the second moment of area for beams. Where, with beams, we are interested in cross-sectional areas to determine their resistance to bending.

The above procedure is best illustrated by example.

#### **Example 3.1.9**

Determine the centroid of area of the beam section shown in Figure 3.1.32.

Now the figure is symmetrical about the center line X–X, therefore the centroid of area ( $\bar{x}$ ) must lie on this line, some distance  $x$ , from the datum face A.

Following the procedure, we first find the individual areas of the beam cross-sections 1, 2, and 3. Noting that, the centroids for these individual areas occur at their geometric centers. This will be the case for homogenous materials, such as steels.

$$\begin{aligned}\text{Then: Area 1} &= (20 \times 100) = 2000 \text{ mm}^2; \\ \text{Area 2} &= (30 \times 250) = 7500 \text{ mm}^2; \\ \text{Area 3} &= (30 \times 150) = 4500 \text{ mm}^2; \\ \text{Total Area} &= 14\,000 \text{ mm}^2.\end{aligned}$$

Again following the procedure, we take moments about the datum face A equating the moment for total area, to the moments for the individual area.

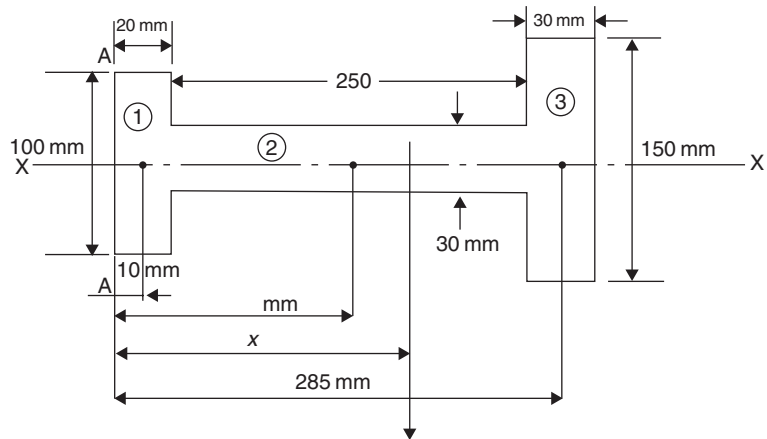


**Figure 3.1.32** Beam section for Example 3.1.9

$$\text{That is: } 14\,000x = (2000 \times 10) + (7500 \times 145) + (4500 \times 285)$$

$$\text{Or } 14\,000x = (20\,000) + (1\,087\,500) + (1\,282\,500)$$

and

$$x = \frac{2\,390\,000}{14\,000} = 170.7 \text{ mm}$$

so the centroid of area,  $\bar{x}$  is on the center line, at a distance  $x = 170.7 \text{ mm}$  from A.

Note that this procedure is similar to that which we used for finding the reaction at the supports of a uniform beam, where areas have been substituted in place of loads. We may choose any convenient point or face of the cross-section, to act as a datum.

Here is one more example, make sure you can follow the argument.

### Example 3.1.10

Find the centroid of area for the cross-section of the uniform beam shown (Figure 3.1.33).

$$\text{Then: Area 1} = (50 \times 74) = 3700 \text{ mm}^2;$$

$$\text{Area 2} = (100 \times 24) = 2400 \text{ mm}^2;$$

$$\text{Area 3} = 0.5(24 \times 24) = 288 \text{ mm}^2;$$

$$\text{Total Area} = 6388 \text{ mm}^2.$$

Taking moments about B–B to find position of centroid  $\bar{x}$ , we get:

$$6388x = (3700 \times 25) + (2400 \times 75) + (288 \times 158), \text{ which gives:}$$

$$\bar{x} = \frac{319\,004}{6388} = 49.94 \text{ mm}.$$

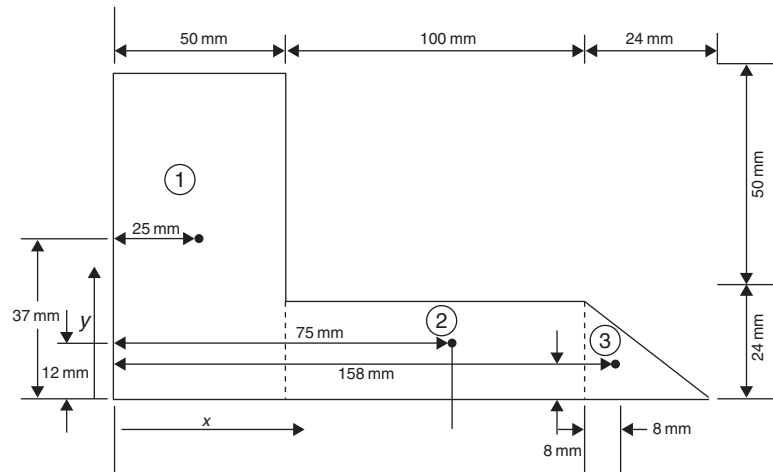


**Figure 3.1.33** Cross-section of beam for Example 3.1.10

Similarly, taking moments about A–A, we get:

$6388y = (3700 \times 37) + (2400 \times 12) + (288 \times 8)$  which gives:

$$\bar{y} = \frac{168004}{6388} = 26.3 \text{ mm.}$$

Then centroid is **49.94 mm** to the right of the B–B datum and **26.3 mm** above the A–A datum.

### Second moment of area (I) and its determination

We have already said that the *second moment of area* or *mass moment of inertia*, is a measure of the efficiency of the beam to resist bending loads. Consider the Universal ‘I’-section Beam shown in Figure 3.1.34. The beam is shown with its principal axes X–X and Y–Y. These axes intersect at the centroid.

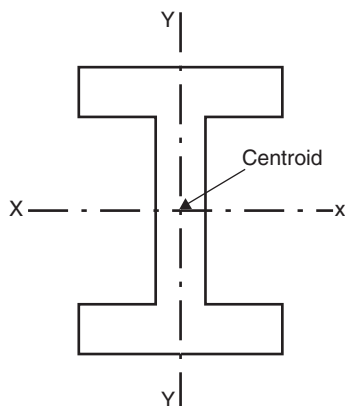


**Figure 3.1.34** Universal beam showing principal axes

Normally when considering the second moment of area, it is necessary to find this, about the X–X and Y–Y axes. This is shown in a later example, where on substitution of values into the formulae for the second moment of area, it is found that when the second moment of area is taken about the X–X axis, with the greater amount of material further away from the reference axis. *Then the higher will be the value of the second moment of area and the higher will be the resistance to bending of the beam.*

In general then, the resistance to bending of a beam, will be greater, as its area and depth increase.

This is easily demonstrated by considering a steel rule (Figure 3.1.35), where it is easy bent around cross-section (A–A) but it is difficult to bend around cross-section (B–B). The second moment of area about the principal axes for a rectangular cross-section beam is shown next, in Example 3.1.11.

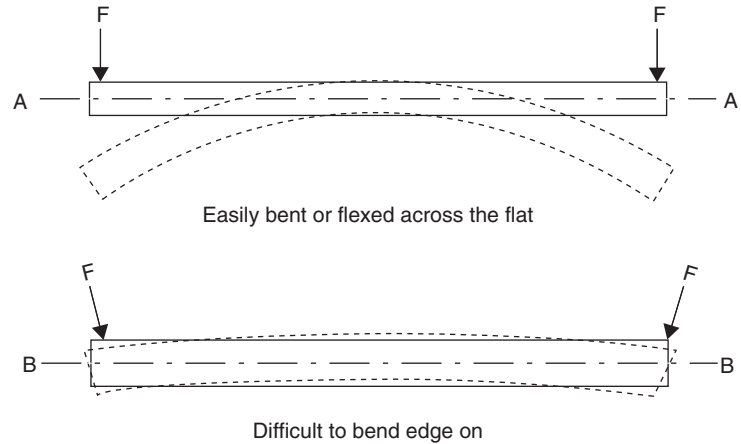


**Figure 3.1.35** Resistance to bending of a steel rule

### Example 3.1.11

Derive the formulae for the second moment of area  $I_{xx}$  about the X–X axis for the rectangular cross-section, shown (Figure 3.1.36a).

We are required to find the second moment of area about the X–X principal axis. Figure 3.1.36b shows the same rectangular section, with an elemental strip and its breadth and depth, marked on.

The first stage in determining the second moment of area of a beam cross-section is to find its centroid. This is easily found for a rectangular cross-section and is shown in Figure 3.1.36(a), as being at the intersection of the principal axis.

Then considering this elemental strip having breadth  $b$  and depth  $dy$ , at a distance  $y$  from the X–X axis.

Then the second moment of area of this strip is *its area multiplied by distance*  $y = (b)(dy)(y^2)$  or  $by^2 dy$ . So the second moment of half of the rectangular area, is the sum of all these quantities.

Then, using the integral calculus to sum these quantities. We have:

$$I_{xx} \text{ for half the rectangular area } \int_0^{d/2} by^2 dy = \left[ \frac{by^3}{3} \right]_0^{d/2} = \frac{bd^3}{24}.$$

$$\text{Then for complete rectangular area } I_{xx} = \frac{bd^3}{12}.$$

By symmetry the second moment of area about the Y–Y principal axis is given as:

$$I_{yy} = \frac{db^3}{12}$$

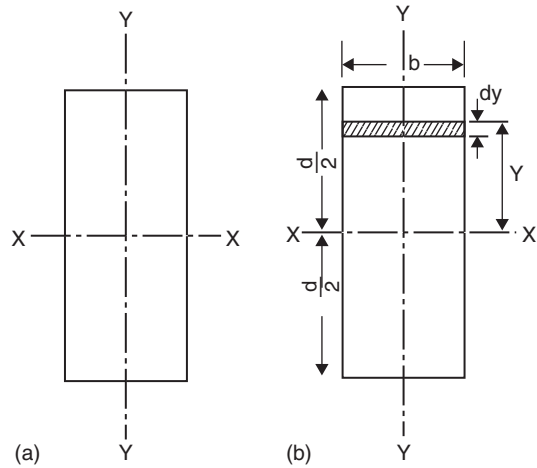


**Figure 3.1.36** (a) Rectangular section. (b) Rectangular section showing elemental strip, breadth and depth relationships

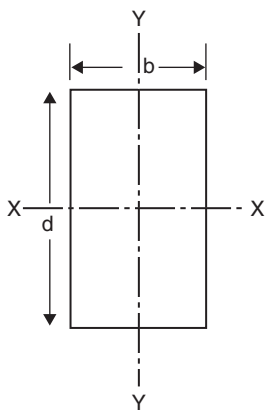


**Figure 3.1.37** Rectangular section about principal axes

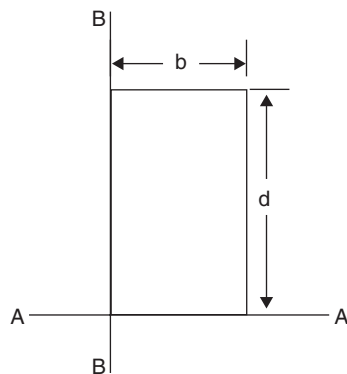


**Figure 3.1.38** Rectangular section about one edge

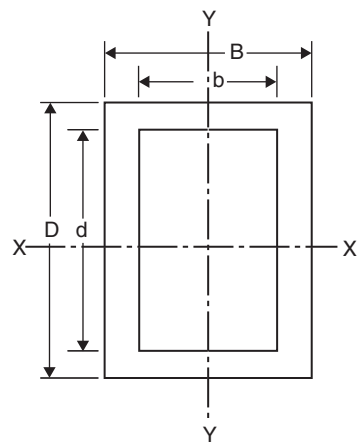


**Figure 3.1.39** Hollow rectangular section about principal axes

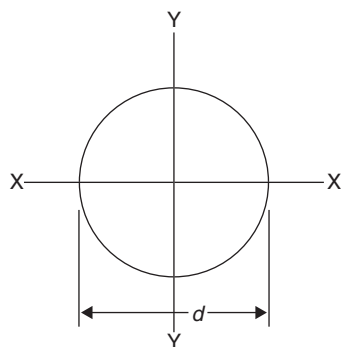


**Figure 3.1.40** Circular section about principal axes

An example for determining the second moment of area of a rectangular cross-section, when moments are taken about its base will be found in the calculus section of Chapter 6 on analytical methods, Example 6.3.26, as mentioned before.

The table given below summarises the second moment of area for rectangular and circular shapes about their various axes.

| Section detail                                                     | Second moment of area                                              |
|--------------------------------------------------------------------|--------------------------------------------------------------------|
| Rectangular section about principal axes (Figure 3.1.37)           | $I_{XX} = \frac{bd^3}{12}, I_{YY} = \frac{db^3}{12}$               |
| Rectangular section about one edge (Figure 3.1.38)                 | $I_{AA} = \frac{bd^3}{3}, I_{BB} = \frac{db^3}{3}$                 |
| Hollow rectangular section about principle axes (Figure 3.1.39)    | $I_{XX} = \frac{BD^3 - bd^3}{12}, I_{YY} = \frac{DB^3 - db^3}{12}$ |
| Circular section about principal axes (polar axes) (Figure 3.1.40) | $I_{XX} = I_{YY} = \frac{\pi d^4}{64}$                             |

**Parallel axes theorem**

To determine the second moment of area of any shape *about an axis B–B*, parallel to another axis A–A, at a perpendicular distance  $S$ , then, the area of the shape multiplied by the distance squared ( $AS^2$ ) must be added to the second moment of area about A–A, that is:

$$I_{BB} = I_{AA} + AS^2.$$

**Example 3.1.12**

Consider the I section universal beam shown in Example 3.1.9. Determine the second moment of area about the principal axes  $I_{XX}$  and  $I_{YY}$ , for this beam.

The beam is reproduced here in Figure 3.1.41, showing the I-section in its normal vertical position, with the principal axes marked on in their correct positions about the centroid of area.

We know from Example 3.1.9, that the centroid of area for the whole beam acts through the center of the web, at 170.7 mm from the A–A axis.

We can verify the centroid position by taking moments about the base B–B. So that:

$$14\,000x = (4500 \times 15) + (7500 \times 145) + (2000 \times 290) \text{ and}$$

$$14\,000x = 1\,810\,000 \text{ or,}$$

$$x = \frac{1\,810\,000}{14\,000} = 129.3 \text{ mm as expected!}$$

To find  $I_{XX}$ , we first divide the I-beam into three rectangular sections, as we did, when finding its centroid. Then using the *parallel axis theorem*, that is  $I_{XX} = I_{NN} + A_N(S_N)^2$ . We find the individual values of  $I_{XX}$  for each of the rectangular areas, then sum them to find the total value of  $I_{XX}$  for the whole beam section.

So, using the parallel axes theorem and taking moments about X–X. We get for the top flange (Area 1):

$$I_{XX_1} = I_{11} + A_1(S_1)^2 = \frac{(100)(20)^3}{12} + (2000)(160.7)^2 \text{ or}$$

$$I_{XX_1} = 51.716 \times 10^6 \text{ mm}^4.$$

Similarly, for the web, we get:

$$I_{XX_2} = I_{22} + A_2(S_2)^2 = \frac{(30)(250)^3}{12} + (7500)(25.7)^2 \text{ or}$$

$$I_{XX_2} = 44.016 \times 10^6 \text{ mm}^4.$$

And for the bottom flange, we get:

$$I_{XX_3} = I_{33} + A_3(S_3)^2 = \frac{150(30)^3}{12} + (4500)(114.3)^2 \text{ or}$$

$$I_{XX_3} = 59.128 \times 10^6 \text{ mm}^4.$$

Therefore total second moment of area about X–X for the beam section is,

$$I_{XX} = 154.86 \times 10^6 \text{ mm}^4.$$

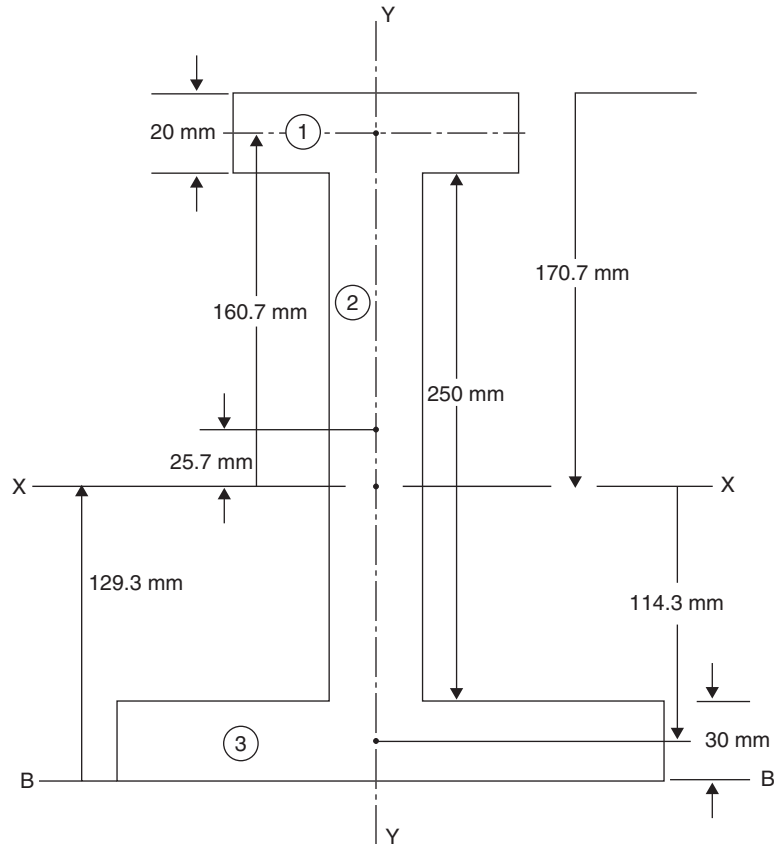


**Figure 3.1.41** I – Section universal beam for example 3.1.12

Now to find  $I_{YY}$ , we note that all the individual rectangular second moments of area, lie on the Y–Y axis, therefore,  $S = 0$  and in all cases, the total second moment of area about Y–Y reduces to the sum of the individual rectangular second moments, that is:

$$I_{YY} = I_{YY_1} + I_{YY_2} + I_{YY_3} \text{ or,}$$

$$I_{YY} = \frac{(20)(100)^3}{12} + \frac{(250)(30)^3}{12} + \frac{(30)(150)^3}{12} \text{ therefore}$$

$$I_{YY} = 10.67 \times 10^6 \text{ mm}^4.$$

Make sure that you are able to follow the whole argument and obtain, the same numerical values.

### Beam selection

Now, you will have noted that it is quite tedious to calculate values of  $I$  immediately the beam is anything other than rectangular, and yet, this is only *one* of the criteria necessary for **beam selection**.

As mentioned earlier tables of beams exist, which may be used to select a beam, for a specified purpose, to sustain a particular loading regime. These tables quote key dimensions and inherent properties such as the elastic and plastic modulus, together with the second moment of area (mass moment of inertia) and radius of gyration (which you will meet a little later), of particular beam sections.





**Example 3.1.13**

A rolled steel universal I-section beam with a serial size of  $406 \times 178$ , has a mass of 60 kg/m. What is the maximum safe allowable bending moment this beam can sustain, given that the maximum allowable bending stress in tension and compression must not exceed  $165 \text{ N/mm}^2$ .

This example requires us to determine the appropriate section modulus from tables and then use engineers theory of bending to determine  $M_{\max}$ .

Then from above extract, for this beam section,

$$Z_{X-X} = 1059 \text{ cm}^3 \quad \text{or} \quad 1059 \times 10^3 \text{ mm}^3.$$

$$\text{Now: } \frac{M}{I} = \frac{\sigma}{y} = \frac{E}{R}, \text{ or, ignoring the last term and remembering that } Z = \frac{I}{y}, \text{ we get, } \frac{M}{\sigma} = \frac{I}{y} = Z \text{ and}$$

$$\text{bering that } Z = \frac{I}{y}, \text{ we get, } \frac{M}{\sigma} = \frac{I}{y} = Z \text{ and}$$

$M = \sigma Z$ , this will be a maximum, when  $Z$  is a maximum, that is when we use the beam correctly about the X-X axis.

$$\text{So, } M_{\max} = (165 \text{ N/mm}^2)(1.059 \times 10^6 \text{ mm}^3), \text{ or}$$

$$\mathbf{M_{\max} = 174.7 \times 10^6 \text{ N mm.}}$$

**Example 3.1.14**

A roof support, is made from a wooden rectangular section beam 100 mm wide and 200 mm deep. If the maximum allowable bending moment, anywhere along the beam is  $14 \times 10^6 \text{ N mm}$ . Determine the maximum allowable bending stress.

In this case we have no tabulated data for such a beam, the properties of wooden beams, are dependent not only on beam geometry but also on the particular species of wood being used. So we need first, to *calculate*  $Z$ .

Figure 3.1.43 shows the beam with distance 'y' marked on. This is the distance to the edge of the beam material, from its neutral axis and because the beam is symmetrical about the X-X axis, the distance  $y$  is equal in either direction, from this axis.

Therefore knowing that  $Z = \frac{I_{XX}}{y}$ ,  $y = \frac{d}{2}$  (that is half the depth) and

$$\text{that, } I_{XX} = \frac{bd^3}{12}, \text{ then } Z = \frac{I_{XX}}{y} = \frac{\frac{bd^3}{12}}{\frac{d}{2}} = \left( \frac{bd^3}{12} \right) \left( \frac{2}{d} \right) = \frac{bd^2}{6}.$$

$$\text{Thus } Z = \frac{(100)(200)^2}{6} = 0.667 \times 10^6 \text{ mm}^3 \text{ and know}$$

$$M_{\max} = Z\sigma_{\max}, \text{ then: } \sigma_{\max} = \frac{M_{\max}}{Z} = \frac{14 \times 10^6 \text{ N mm}}{0.667 \times 10^6 \text{ N mm}^3} = \mathbf{21 \text{ N/mm}^2}.$$

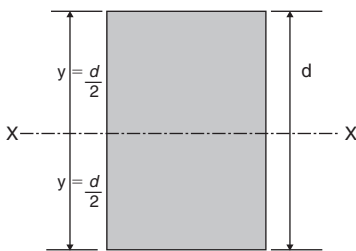


**Figure 3.1.43** Rectangular wooden beam section

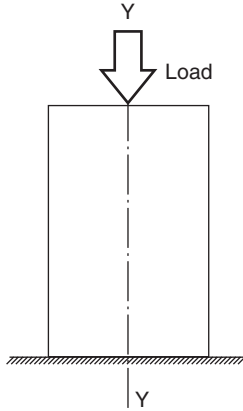


**Figure 3.1.44** Column subject to concentric compressive load

### Concentrically loaded columns

A beam turned on end and loaded directly through its center line, may be considered to be a *column*. Under these circumstances the column is subject to *concentric* or *axial compressive loads* (Figure 3.1.44).

Structural members that are subject to concentric compressive loads, are not only known as columns. Concrete *pillars*, support axial compressive loads, using the fact that concrete is around 15 times stronger in compression, than it is in tension. In steel girder construction, *struts* are used as the main compressive support members. In wooden structures, the axially loaded supports are often referred to as *posts*. We will continue to use the word column, for all structural members that are loaded axially and subject to compressive stresses.

In a similar way to beams, the maximum concentric loads that a column can support depends on the *material* from which the column is manufactured and the *slenderness* of the column. The slenderness of a column is determined by considering column height, size, and shape of cross-section. As you will see later, a measure of the shape of cross-section is given by the *radius of gyration*.

The slenderness of a column, has a direct effect on the way in which the column may fail under compressive loads. Thus if the column is very long and thin, it is likely to fail due to buckling, well before it may have failed due to compressive stresses. How do we define the slenderness of a column? Logic would suggest that we can use a ratio, the *slenderness ratio*. Where, for a column the slenderness ratio may be defined as:

$$\text{Slenderness ratio} = \frac{\text{Column length or height}}{\text{Minimum column width}}$$

Now this is fine for solid rectangular sections, such as posts used for wooden structures. However, we said earlier that the axial loads that can be sustained by columns, are dependent on the shape of their cross-section. For rectangular columns, the above ratio, provides an effective measure of whether or not a column will fail due to buckling. For other column cross-sections, such as I-sections, circular-sections, and L-sections, commonly found in steel columns, the above simple ratio has to be modified. This is where the radius of gyration, is used. Then,

$$\text{Slenderness ratio} = \frac{\text{Effective length}}{\text{Least radius of gyration}} = \frac{L_e}{r}$$

Now in order for us to be able to understand the use of this ratio, we need first to define, the least radius of gyration and then, effective length.

### Radius of gyration

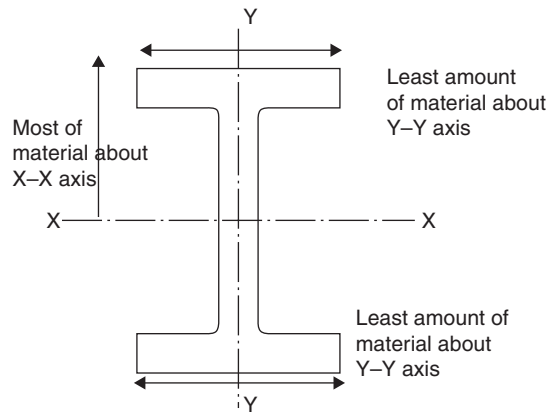
The least radius of gyration takes into account, both the size and shape of the column cross-section. The size being measured by the column cross-sectional area  $A$  and the shape effect, being determined by considering the second moment of area  $I$  (*moment of inertia*) of the column. Thus:

$$\text{Radius of gyration } r = \sqrt{\frac{I}{A}}$$

However the slenderness ratio is concerned with the *least* radius of gyration, which will be least when the moment of inertia is least. This

**Table 3.1.2**

| <i>Steel column section</i>     | <i>Least value of radius of gyration (r)</i> |
|---------------------------------|----------------------------------------------|
| Solid rectangular cross-section | $r = 0.289b$                                 |
| Hollow square section           | $r = \frac{1}{2} \sqrt{\frac{B^2 - b^2}{3}}$ |
| Annulus                         | $\frac{\sqrt{D^2 + d^2}}{4}$                 |

**Figure 3.1.45**

occurs when  $I$  is taken about the axis where the least amount of material protrudes (Table 3.1.2).

So, for example, if a column is to have an  $I$  cross-section (Figure 3.1.45), then the least moment of inertia will be about the Y–Y axis. So for the I-column section show:

**Least radius of gyration**  $r = \sqrt{\frac{\text{Least moment of inertia}}{\text{Area of cross-section}}}$ , that is

$$r = \sqrt{\frac{I_{YY}}{A}}.$$

### Example 3.1.15

A concrete column has a rectangular cross-section, with a breadth of 300 mm and a depth of 600 mm and an effective length of 3.4 m. For this column, calculate:

- the least radius of gyration;
- the slenderness ratio.

To find the least radius of gyration we must first find the least second moment of area (moment of inertia). This occurs about the Y–Y principal axis, that is the longitudinal axis, where:

$$I_{YY} = \frac{db^3}{12} = \frac{600(300)^3}{12} = 1350 \times 10^6 \text{ mm}^3. \text{ Then,}$$

a) the least radius of gyration:

$$r = \sqrt{\frac{I_{YY}}{A}} = \sqrt{\frac{1350 \times 10^6 \text{ mm}^3}{(300)(600) \text{ mm}^2}} = 86.6 \text{ mm}.$$

b) the slenderness ratio:

$$\frac{L_e}{r} = \frac{3.4 \times 10^3}{86.6} = 39.26.$$

So we can now find the slenderness ratio, for a variety of column cross-sections, but what do we do with it?

Well, there is a direct comparison between the slenderness ratio of a column and its permissive compressive strength ( $\sigma_c$ ). Thus if we know the slenderness ratio, then we can determine the compressive buckling strength of the column. These comparisons for *steel section* columns are tabulated, for example, in BS 5950.

By consulting these or similar tables, we can determine the compressive strength for a particular slenderness ratio. For example if the maximum compressive strength of a column about its Y–Y axis is 275 N/mm<sup>2</sup>. Then from tables in *BS 5950 Part 1*, when the slenderness ratio is 80, the compressive strength is reduced to 181 N/mm<sup>2</sup>. Codes of practice limit the slenderness ratio, for all practical purposes to a maximum of 180. When the slenderness ratio is 180, then for the above example the compressive strength reduces to 54 N/mm<sup>2</sup>, that is only one-fifth of the original compressive strength!

At slenderness ratios of 15 and below, the column will tend to fail at critical loads, that are the same as those required for the column to fail in direct compression. Thus as a ‘rule of thumb’ restricting the slenderness ratio to values of 15 or less, means that failure in buckling can generally be ignored.

In order to find the slenderness ratio, you will have noted from Example 3.1.15, that it was necessary to first find the least radius of gyration. The formulae for finding  $r$ , for some common steel cross-sections is tabulated in (Table 3.1.2), for your convenience.

### Effective length

So far we have not defined the *effective length*. This is directly related to the way in which the column is fixed at its ends.

Figure 3.1.46 shows the four basic methods of attachment for columns. In Figure 3.1.46(a), both the ends are pinned, this is considered to be the basic case, where for the purposes of analysis the actual length is equal to the effective length. Figure 3.1.46(b) shows both ends are rigidly fixed, under these conditions, the middle half of the column behaves in the same manner as in the basic case. For this reason the effective length is the actual length divided by two. Figure 3.1.46(c) shows the case where one end is fixed and the other end is free, under these circumstances, the column behaves as though this length is only half of the effective length, so that the effective length becomes twice the actual length. Finally Figure 3.1.46(d) shows the case where one end is fixed and the other is pin jointed. Analysis shows that under these circumstances the column has an effective length of approximately 0.7 times the actual length. These findings are tabulated in Table 3.1.3.

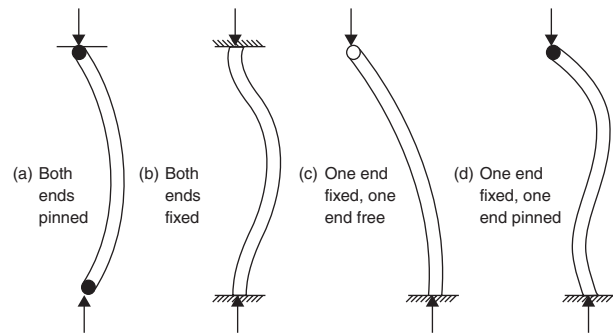


**Figure 3.1.46** End fixtures for columns under compressive load

**Table 3.1.3** Relationship between actual length  $L$  and effective length  $L_e$ , for various methods of attachment

| Method of fixing           | Both ends pinned | Both ends rigidly fixed | One end fixed, one end free | One end fixed, one end pinned |
|----------------------------|------------------|-------------------------|-----------------------------|-------------------------------|
| Effective length ( $L_e$ ) | $L$              | $\frac{L}{2}$           | $2L$                        | $0.7L$                        |

Note: BS 5950-1: 2000 provides a slight variant for conditions between total restraint and partial restraint. The latest and most appropriate British Standards or ISO standards must be consulted, if column design information is required.

**Example 3.1.16**

A hollow section steel column is shown below (Figure 3.1.47), with the dimensions indicated.

If the column is rigidly fixed at both of its ends and is subject to an axial load, through its centroid. Then, given that the compressive resistance of a column is the product of its compressive strength =  $200 \text{ N/mm}^2$  and section area, Determine for the column:

- a) the least radius of gyration;
- b) the slenderness ratio;
- c) the compressive resistance.

a) From Table 3.1.2, then least radius of gyration

$$r = \frac{\sqrt{D^2 + d^2}}{4} \text{ and } r = \frac{\sqrt{200^2 + 150^2}}{4} = \mathbf{62.5mm.}$$

If tables are *not* available then  $r$  may be found from  $I =$

$$I_{XX} = I_{YY} \text{ for an annulus, that is } I_{\text{least}} = I = \frac{\pi(D^4 - d^4)}{64}$$

and then using the relationship  $r = \sqrt{\frac{I_{\text{least}}}{A}}$ , where the

$$\text{section area of an annulus is given by, } A = \frac{\pi(D^2 - d^2)}{4},$$

you will find this yields the same result!

- b) To find the slenderness ratio, we just need to find the effective length, which is  $L/2$ , (from Table 3.1.3), that is

$$L_e = \frac{2600 \text{ mm}}{2} = 1300 \text{ mm};$$

So the **slenderness ratio**  $= \frac{L_e}{r} = \frac{1300 \text{ mm}}{62.5 \text{ mm}} = \mathbf{20.8}$ .

- c) Finally, the compressive resistance is equal to the compressive strength of the column ( $\sigma_c$ ), multiplied by the section area ( $A$ ). The section area for an annulus

$$A = \frac{\pi(D^2 - d^2)}{4} \text{ or } A = \frac{\pi(200^2 - 150^2)}{4} = 4375\pi.$$

And so, column compressive resistance ( $F_c$ ) is:

$$\mathbf{F_c = (200 \text{ N/mm}^2)(4375\pi \text{ mm}^2) = 2.7489 \text{ MN.}}$$

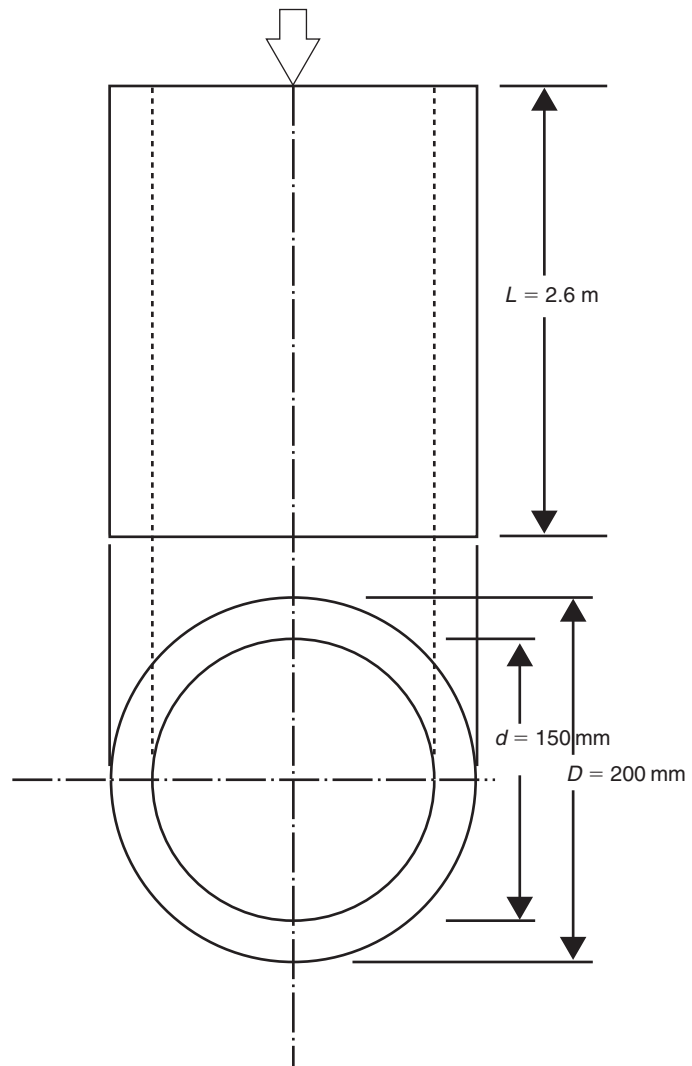


**Figure 3.1.47** *Hollow section steel column*

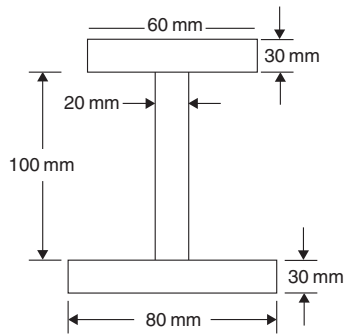


Figure 3.1.48 I-section beam

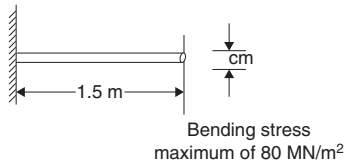


Figure 3.1.49 Cantilever beam

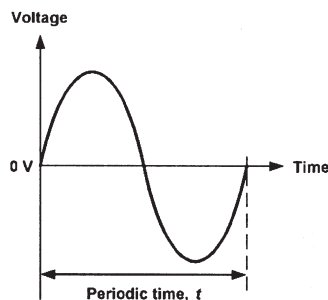


Figure 3.1.50 Axially loaded column

## Problems 3.1.2

- For the beam section shown in Figure 3.1.48. Calculate:
  - the centroid of area
  - $I_{XX}$ , the second moment of area (mass moment of inertia) about the X-X principal axis.
- A uniform beam is simply supported at its ends and is 6 m long. A concentrated load of 6 kN, acts at the centre of the beam. Given that the beam is of rectangular cross-section with a depth of 12 cm and a breadth of 8 cm and has a mass of 2 kN. Determine, using bending theory, the:
  - reactions at the supports
  - maximum second moment of area  $I_{XX}$
  - maximum bending moment
  - value and nature of the maximum stresses in the beam. compressive at top of beam tensile at bottom of beam.
- Figure 3.1.49 shows a uniform circular section cantilever, of length 1.5 m and diameter 7.0 cm. If the maximum longitudinal stress due to bending must not exceed  $80 \text{ MN/m}^2$ . Then, neglecting the mass of the bar, determine:
  - The second moment of area
  - The maximum bending stress
  - The maximum load that the cantilever can support at its free end.
- A steel cylindrical hollow tube, with outside diameter of 10 cm and 15 mm thick is used as a beam over a span of 3 m, being simply supported at its ends. Neglecting the mass of the tube determine the greatest point load that can act at the centre of the span, given that the maximum bending stress is limited to a maximum of 125 MPa.
- A uniform wooden rectangular section beam is 120 mm wide and 240 mm deep. If the maximum allowable bending moment, anywhere along the beam is  $15 \times 10^6 \text{ Nmm}$ . Determine the maximum allowable bending stress.
- A rolled steel universal I-section beam with a serial size of  $254 \times 102$ , has a mass of 25 kg/m. Using Table 3.1.1, find the maximum safe allowable bending moment this beam can sustain. Given that the maximum allowable bending stress in tension and compression must not exceed  $160 \text{ N/mm}^2$ .
- A hollow concrete column has a rectangular cross-section, with an external breadth of 300 mm and external depth of 600 mm. If the internal breadth is 240 mm and the internal depth is 520 mm. Then for a column with an effective length of 2.8 m, calculate:
  - the least radius of gyration
  - the slenderness ratio.
- The rolled steel universal I-section column shown in Figure 3.1.50 below, is subject to an axial load through its centroid. If the beam is 3 m long and is to be rigidly fixed at one end and pinned at the other end. Then, given that the compressive strength of the

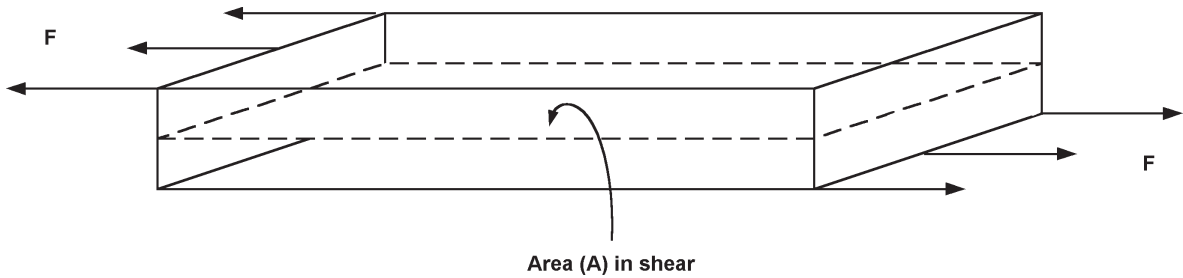


**Figure 3.1.51** Solid subject to shear stress

column is  $210 \text{ N/mm}^2$ . Determine:

- the least radius of gyration
- the slenderness ratio
- the compressive resistance.

## Torsion

Drive shafts for pumps and motors, propeller shafts for motor vehicles and aircraft as well as pulley assemblies and drive couplings for machinery, are all subject to *torsional* or *twisting* loads. At the same time shear stresses are set up within these shafts resulting from these torsional loads. Engineers need to be aware of the nature and magnitude of these torsional loads and the subsequent shear stresses in order to design against premature failure and to ensure safe and reliable operation during service.

### Review of shear stress and strain

In order to study the theory of torsion it is necessary to review the fundamental properties of shear stress and strain. For some of you this may be a new topic, for others the coverage will act as revision however, no matter what your background, it is a convenient place to start.

If two equal and opposite parallel forces, not acting in the same straight line, act on a body, then that body is said to be loaded in *shear*, Figure 3.1.52. Shear stress is defined in a similar way to tensile stress, except that the area in shear acts parallel to the load, so:

$$\text{Shear stress } (\tau) = \frac{\text{Load causing shear } (F)}{\text{Area in shear } (A)}$$

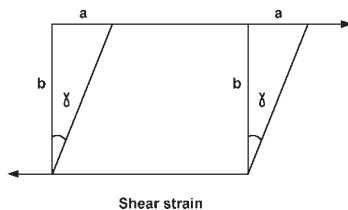


**Figure 3.1.52** Shear strain and the angle of deformation

Figure 3.1.52 illustrates the phenomenon known as shear strain. With reference to Figure 3.1.52, *shear strain* may be defined in two ways as:

- the relative distance between the two surfaces on which the shear load acts divided by the distance between the two surfaces; or
- the angle of deformation  $\theta$  (in radians).

The relationship between shear stress and shear strain is again similar to that of tensile stress and strain. In the case of shear the modulus is known as the *modulus of rigidity* or *shear modulus* ( $G$ ), and is defined as:

$$G = \frac{\text{Shear stress } (\tau)}{\text{Shear strain } (\gamma)}$$



The modulus of rigidity ( $G$ ), or shear modulus, measures the stiffness of the body, in a similar way to the elastic modulus, the units are the same,  $G$  being measured in  $\text{N/m}^2$ .

Riveted joints are often used where high shear loads are encountered, bolted joints being used where the primarily loads are tensile.

### Example 3.1.17

Figure 3.1.53 shows a bolted coupling in which the two 12 mm securing bolts act at  $45^\circ$  to the axis of the load. If the pull on the coupling is 80 kN, calculate the direct and shear stresses in each bolt.

We know that the axis of the bolts is at  $45^\circ$  to the line of action of the load ( $P$ ) = 80 kN. We will assume that the bolts are ductile and so the load is shared equally between the bolts.

Shear area of each bolt is  $= \frac{\pi(12)^2}{4} = 113.1 \text{ mm}^2$ .

Shear force ( $F_s$ ) from diagram  $= P \sin 45 = 80 \times 10^3 \times 0.7071 = 56\,568.5 \text{ N}$ . This shear force is shared between the two bolts.

Thus shear force on each bolt  $= \frac{56\,568.5}{2 \times 113.1} = 250.08 \text{ N/mm}^2$ .

Similarly:

the direct tensile force  $= P \cos 45 = 80 \times 10^3 \times 0.7071 = 56\,568.5 \text{ N}$  and the tensile force on each bolt  $= \frac{56\,568.5}{2 \times 113.1} = 250.08 \text{ N/mm}^2$ .

*So direct and shear stresses are, respectively,  $250 \text{ MN/m}^2$  and  $250 \text{ MN/m}^2$ .*

### Engineers' theory of torsion

Shafts are the engineering components which are used to transmit torsional loads and twisting moments or *torque*. They may be of any cross-section but are often circular, since this cross-section is particularly suited to transmitting torque from pumps, motors and other power supplies used in engineering systems.

We have been reviewing shear stress because if a uniform circular shaft is subject to a torque (twisting moment) then it can be shown that

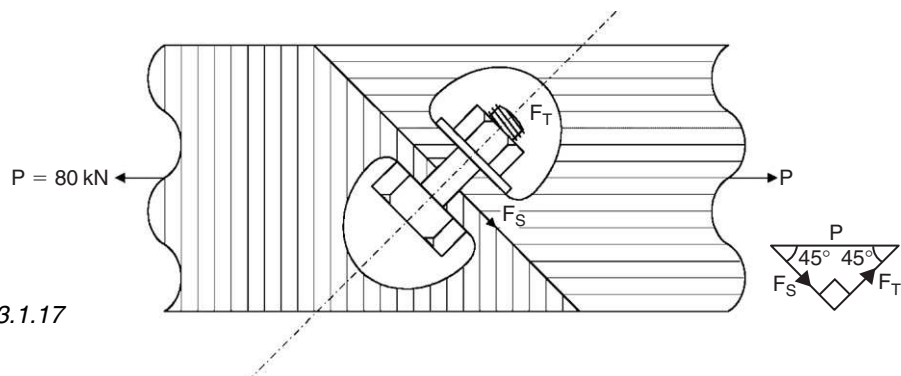
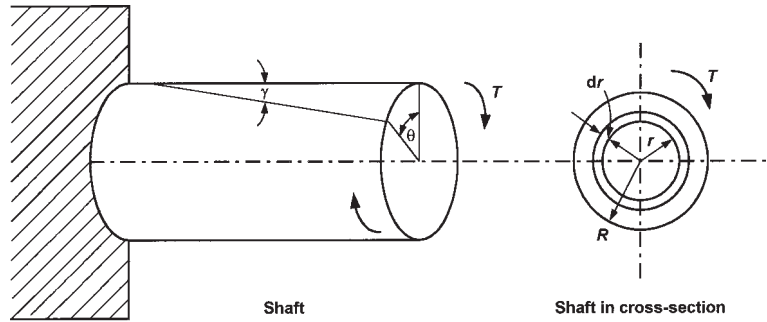


**Figure 3.1.53** Example 3.1.17  
bolted coupling subject to  
tensile load

**Figure 3.1.54** Circular shaft subject to torque



every section of the shaft is subject to a state of pure shear. In order to help us derive the engineers' theory of torsion which relates torque, the angle of twist and shear stress, we must first make the following fundamental assumptions.

- (i) The shaft material has uniform properties throughout.
- (ii) The shaft is not stressed beyond the elastic limit, in other words it obeys Hooke's law.
- (iii) Each diameter of the shaft carries shear forces which are independent of and, do not interfere with their neighbours.
- (iv) Every cross-sectional diameter rotates through the same angle.
- (v) Circular sections which are radial before twisting are assumed to remain radial after twisting.

Let us first consider torsion from the point of view of the *angle of twist*.

Figure 3.1.54 shows a circular shaft of radius  $R$  which is firmly fixed at one end and at the other is subject to a torque (twisting moment)  $T$ . Imagine that on our shaft we have marked a radial line of length  $L$ , which when subjected to torque  $T$ , twists from position  $p$  to position  $q$ . Then the angle  $\theta$  is the same angle of deformation we identified in Figure 3.1.52 and is therefore the *shear strain* of our shaft.

You will remember that the angle of distortion is measured in radians. When considering torsion, shear strain or the angle of twist is measured in radians per unit length, that is in *rad/m*.

Thus from basic trigonometry:

arc length  $pq$  = the radius  $R$  multiplied by  $\theta$  the angle of distortion seen in cross-section. It also follows that the arc length  $pq$  = the length of the radial line  $L$  multiplied by the angle of distortion  $\gamma$  so:

arc length  $pq = R\theta = L\gamma$  from which

$$\gamma = \frac{R\theta}{L}. \quad (3.1.9)$$

Now from our previous study of the modulus of rigidity we know that:

$$G = \frac{\tau}{\gamma} \quad \text{or} \quad \gamma = \frac{\tau}{G}. \quad (3.1.10)$$

Then combining Equations 3.1.9 and 3.1.10 and rearranging gives the relationship:

$$\frac{\tau}{R} = \frac{G\theta}{L}. \quad (3.1.11)$$

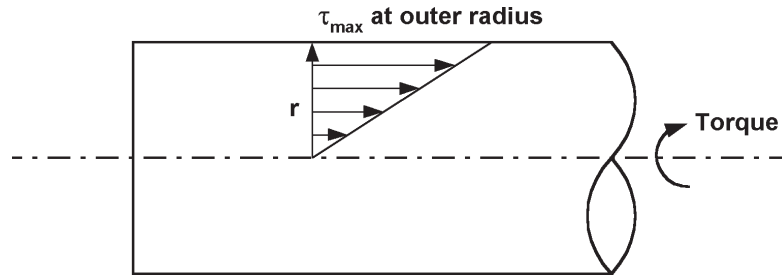


**Figure 3.1.55** Shear stress distribution for shaft subject to a torque

The above relationship is independent of the value of the radius,  $R$ , so that any intermediate radius, emanating from the centre of the cross-section of the shaft can be considered. The related shear strain can then be determined for that radius.

With a little more algebraic manipulation we can also find expressions which relate the *shear stresses*, developed in a shaft subject to pure torsion. Their values are given by Equations (3.1.10) and (3.1.11), if these two equations are combined and rearranged we have:

$$\tau = \gamma G = \frac{G\theta R}{L}. \quad (3.1.12)$$

Equation (3.1.12) is useful because it relates the shear stress and shear strain with the angle of twist per unit length. It can also be seen from this relationship that the shear stress is directly proportional to the radius, with a maximum value of shear stress occurring at the outside of the shaft at radius  $R$ . Obviously the shear stress at all other values of the radius will be less, since  $R$  is less. Other values of the radius apart from the maximum are conventionally represented by lower case  $r$ , see Figure 3.1.54.

Figure 3.1.55 shows the cross-section of a shaft subject to torsional loads, with the corresponding shear stress distribution, which increases as the radius increases.

Although not part of our study at this time, it should be noted that the shear stresses shown in Figure 3.1.55 have complimentary shears of equal value running normal to them, i.e. parallel with the longitudinal axis of the shaft.

We will now derive the relationship between the *stresses* and the *torque* that is imposed on a circular shaft, this will involve the use of the integral calculus. Figure 3.1.54 shows the cross-section of our shaft, which may be considered as being divided into minute parts or *elements*. These elements having radius  $r$  and thickness  $dr$ . From elementary theory we know that:

Force set up on each element = stress  $\times$  area.

The area of each of these cross-sectional elements of thickness  $dr$  at radius (approximately) equal to  $r$  is:

$$2\pi r dr$$

now if  $\tau$  is the shear stress at radius  $r$  then the shear force on the element is:

$$2\pi r dr \tau$$

so the torque carried by the element is:

$$2\pi r^2 \tau dr \quad (3.1.13)$$

also from Equation (3.1.12)  $\tau = \frac{G\theta r}{L}$  and so substituting Equation (3.1.12) into Equation (3.1.13) for the shear stress  $t$  then:

$$\text{Torque carried by element} = 2\pi r^2 \times \frac{G\theta}{L} \times dr = 2\pi \frac{G\theta}{L} r^3 dr.$$

Now the total torque carried by the solid shaft is the sum of all the element torques, from the centre of the cross-section to the extremities of the shaft where  $r = R$ . The integral calculus may now be used to sum all the elements.

$$\text{Then total torque } T = 2\pi \frac{G\theta}{L} r^3 dr.$$

Since all circular sections (elements) remain radial before and after torque is applied then  $\theta$  remains the same for all the elements making up the solid shaft. Also because  $G$  and  $L$  are constants then:

$$T = \frac{G\theta}{L} 2\pi r^3 dr \quad \text{where } J = 2\pi r^3 dr = \text{polar second moment of area}$$

$$T = \frac{G\theta}{L} J$$

$$\text{or } \frac{T}{J} = \frac{G\theta}{L}. \quad (3.1.14)$$

Now combining Equations (3.1.11) and (3.1.14) gives the relationship known as the *engineers' theory of torsion*.

$$\frac{T}{J} = \frac{\tau}{R} = \frac{G\theta}{L} \quad (3.1.15)$$

In practice, it has been shown that the engineers' theory of torsion, based on the assumptions given earlier, shows excellent correlation with experimental results. So although the derivation of Equation 3.1.15 has been rather arduous, you will find it very useful when dealing with problems related to torsion!

We will now consider in a little more detail the polar second moment of area  $J$ , identified above, that is

$$J = 2\pi r^3 dr.$$

One solution of the integral between the limits the polar second moment of area for a solid shaft is:

$$J = 2\pi[r^4/4] \quad \text{and so} \quad J = \frac{2\pi R^4}{4} \quad \text{or} \quad J = \frac{\pi D^4}{32}. \quad (3.1.16)$$

The polar second moment of area defined above is a measure of the resistance to bending of a shaft.

The polar second moment for hollow shafts is analogous to that for solid shafts, except the area is treated as an annulus.

Therefore the polar second moment of area for a *hollow shaft* is:

$$J = \frac{\pi}{2}(R^4 - r^4) \quad \text{or} \quad J = \frac{\pi}{32}(D^4 - d^4) \quad (3.1.17)$$

When the difference between the diameters is very small as in a very thin walled hollow shaft, the errors encountered on subtraction of two very large numbers close together, prohibits the use of Equation (3.1.16). We then have to use an alternative relationship, which measures the polar second moment of area for an individual element. For very thin walled hollow shafts this is a much better approximation to the real case.

That is, the polar second moment of area for a thin walled hollow cylinder is:

$$J = 2\pi r^3 t \text{ (approximately, where } t = dr). \quad (3.1.18)$$

### **Power transmitted by shafts**

The power transmitted by shafts is the final topic we need to consider, in order to be equipped to put theory into practice. The most useful definition of power for a rotating shaft carrying a torque, relates this torque with the angular velocity of the shaft.

Then the power transmitted by a shaft in watts is given by:

$$\text{Power (watts)} = \text{torque} \times \text{angular velocity}$$

where the torque is measured in newton metres (Nm) and the angular velocity is measured in radians/second (rad/s). Thus:

$$\text{Power} = T\omega \text{ watts} \quad (3.1.19)$$

We are now ready to look at one or two applications of our theory.

#### **Example 3.1.18**

A solid circular shaft 40 mm in diameter is subjected to a torque of 800 Nm.

- (i) Find the maximum stress due to torsion.
- (ii) Find the angle of twist over a 2 m length of shaft given that the modulus of rigidity of the shaft is 60 GN/m<sup>2</sup>.
- (i) The maximum stress due to torsion occurs when  $r = R$ , that is at the outside radius of the shaft. So in this case  $R = 20$  mm. Using the standard relationship:

$$\frac{T}{J} = \frac{\tau}{R}$$

We have the values of  $R$  and  $T$ , so we only need to find the value of  $J$  for our solid shaft and then we will be able to find the maximum value of the shear stress  $\tau_{\max}$ .

$$\text{Then for a solid circular shaft } J = \frac{\pi D^4}{32}$$

$$\text{so } J = \frac{\pi(40)^4}{32} = 0.251 \times 10^6 \text{ mm}^4$$

and on substitution into the standard relationship given above we have:

$$\tau = \frac{(25)(800 \times 10^3)}{0.251 \times 10^6} \frac{(\text{mm})(\text{Nmm})}{\text{mm}^4}$$

giving  $\tau_{\max} = 79.7 \text{ N/mm}^2$ .

This value is the maximum value of the shear stress, which occurs at the outer surface of the shaft.

Notice the manipulation of the units, care must always be taken to ensure consistency of units, especially where powers are concerned!

(ii) To find  $\theta$  we again use the engineers' theory of torsion, which after rearrangement gives:

$$\theta = \frac{LT}{GJ}$$

and substituting our known values for  $L$ ,  $T$ ,  $J$ , and  $G$  we have:

$$\begin{aligned}\theta &= \frac{(2000)(800 \times 10^3)}{(60 \times 10^3)(0.251 \times 10^6)} \frac{(\text{mm})(\text{N mm})}{(\text{N mm}^{-2})(\text{mm}^4)} \\ &= 0.106 \text{ radians}\end{aligned}$$

So angle of twist = 6.07 degrees.

Note once again the careful manipulation of units!

### Example 3.1.19

Calculate the power which can be transmitted by a hollow circular propshaft, if the maximum permissible shear stress is  $60 \text{ MN/m}^2$  and it is rotating at 100 rpm. The propshaft has an external diameter of 120 mm and internal diameter of 60 mm.

Again we use the engineers' theory of torsion and note that the maximum shear stress ( $60 \text{ MN/m}^2$ ) will be experienced on the outside surface of the propshaft, where  $R = 60 \text{ mm}$ .

Then using:  $T = \frac{\tau J}{R}$  where  $J = \frac{\pi}{32}(D^4 - d^4)$ .

Then  $J = 19.09 \times 10^6 \text{ mm}^4$  (you should check this result)

So torque

$$\begin{aligned}T &= \frac{(60)(19.09 \times 10^6)}{60} = 19.09 \times 10^6 \text{ N mm} \\ &= 19.09 \times 10^3 \text{ Nm}.\end{aligned}$$

Now the angular velocity in rad/s is  $= \frac{2\pi 100}{60} = 10.47 \text{ rad/s}$

and we know that power is  $= T\omega$ .

So maximum power transmitted by propshaft:

$$= (19.06 \times 10^3)(10.47) = 199.6 \text{ kW}.$$

In this example not all working has been shown, you are strongly advised to check all results and ensure again that the units correspond.

### Problems 3.1.3

---

1. Find the maximum torque that can be transmitted by a solid drive shaft 75 mm in diameter, if the maximum allowable shear stress is  $80 \text{ MN/m}^2$ .
2. Calculate the maximum power that can be transmitted by a solid shaft 150 mm in diameter, when rotating at 360 rpm. If the maximum allowable shear stress is  $90 \text{ MN/m}^2$ .
3. A solid shaft rotating at 140 rpm has a diameter of 80 mm and transmits a torque of 5 kNm. If  $G = 80 \text{ GN/m}^2$ . Determine the value of the maximum shear stress and the angle of twist per metre length of the shaft.
4. A circular hollow shaft has an external diameter of 100 mm and internal diameter of 80 mm. It transmits 750 kW of power at 1200 rpm. Find the maximum and minimum shear stress acting on the shaft and determine the angle of twist over a 2 m length. Take  $G = 75 \text{ GN/m}^2$ .

## 3.2 DYNAMIC ENGINEERING SYSTEMS

This section is concerned with the application of force and energy transfer to dynamic systems and the linear, angular and oscillatory motions, that result. We start our study of dynamics with the revision of uniformly accelerated linear motion and the application of Newton's laws, to such motion. We then consider angular motion, momentum and inertia, together with the nature and effects of friction, that acts on linear and angular motion machines and systems.

Mechanical energy transfer is then considered and its relationship to linear and angular motion is explored. Finally we consider oscillatory motion and the effects that external forces and system damping have on the resulting motion and outputs of oscillatory systems.

### Uniformly accelerated motion

---

We start, by considering Newton's laws of motion. In order to do this we need briefly to revise the concepts of speed, velocity, acceleration and momentum, which are fundamental to a proper understanding of these laws. You may already be familiar with these concepts, but they are presented here to assist those who, for whatever reason, have never met them before.

#### **Fundamental definitions**

##### **Speed, velocity and acceleration**

*Speed* may be defined as *distance per unit time*, or as, *rate of change of position*. Speed takes no account of direction and is therefore a *scalar* quantity. The common SI units of speed are: metres-per-second (m/s) or kilometres-per-hour (km/h).

*Velocity* is defined as *rate of change of position in a specified direction*. Therefore, velocity is a *vector quantity* and the SI units for the magnitude of velocity are the SI units for speed, that is (m/s). The direction of a velocity is not always quoted but it should be understood that the velocity is in some defined direction, even though that direction is not always stated.

*Acceleration* is defined as *rate of change of velocity*, acceleration is also a *vector quantity* and the SI unit of acceleration is m/s/s or (m/s<sup>2</sup>).

### **Equilibrium, momentum, and inertia**

You have already met the concept of equilibrium before in your study of Static's, but for completeness, its relationship to dynamics, is given here. A body is said to be in *Equilibrium* when it remains at rest or when it continues to move in a straight line with constant velocity.

*Momentum is the product of the mass of a body and its velocity (momentum =  $mv$ ).* Any change in momentum requires a change in velocity, that is an acceleration. This is why momentum is sometimes described as the quantity of motion of a body. It may be said that for a fixed quantity of matter to be in equilibrium, it must have constant momentum. The use of momentum, will be seen next when we consider Newton's laws.

All matter resists change. *The force resisting change in momentum (that is acceleration) is called inertia.* The inertia of a body depends on its mass, the greater the mass, the greater the inertia. The inertia of a body is an innate (in-built) force that only becomes effective when acceleration occurs. An applied force acts against inertia so as to accelerate (or tend to accelerate) a body.

### **Force**

Before we consider Newton's laws we will revise the concept of force. You have already meet the idea of force in your study of Static Systems, when we considered the forces acting on beams and columns. We now look at the concept of force with respect to its effects on dynamic systems. The applied force is called the *action* and the opposing force it produces is known as the *reaction*.

*Force is that which changes, or tends to change, the state of rest or uniform motion of a body.* Forces that act on a body may be external (applied from outside the body) such as weight, or internal (such as the internal resistance of a material subject to a compression).

The effects of any force depend on its three characteristics – *magnitude, direction, and point of application*, as again, you will remember from your Static's (refer back to Figure 3.1.1).

The difference between the forces tending to cause motion and those opposing motion is called the *resultant* or *out-of-balance force*. A body that has no out-of-balance external force acting on it is in equilibrium and will not accelerate. A body that has such an out-of-balance force will accelerate at a rate dependent on the mass of the body and the magnitude of the out-of-balance force. The necessary opposition that permits the existence of the out-of-balance force is provided by the force of inertia.

### **Newton's laws of motion**

Sir Isaac Newton (1642–1727) formulated, among other things three laws of motion. These deal with the acceleration produced on a body, in opposition to an external force. These long standing laws assist engineers in many motion-related design problems.

#### **Newton's first law of motion**

States that: *a body remains in a state of rest, or of uniform motion in a straight line, unless it is acted upon by some external resultant force.*

Thus, if a body is moving it requires a force to cause it to accelerate or decelerate. The reason why a body behaves in accordance with Newton's first law is because of its inertia, which causes the body to resist the change of motion.



**Newton's second law of motion**

States that: *the rate of change of momentum of a body is directly proportional to the force producing the change, and takes place in the direction in which the force acts.*

We also know, from our above definitions, that acceleration may be defined as, change in velocity per unit time or rate of change in velocity. If a force  $F$  (N) acts on a body of mass  $m$  (kg) for a time of  $t$  (s) then the velocity changes uniformly from  $u$  (m/s) to  $v$  (m/s),

Then:

$$\begin{aligned}\text{The rate of change of momentum} &= \frac{\text{change of momentum}}{\text{time taken}} \\ &= \frac{mv - mu}{t} \\ &= \frac{m(v - u)}{t}\end{aligned}$$

$$\text{but, } a = \frac{(v - u)}{t},$$

$$\therefore \text{Rate of change of momentum} = ma$$

and according to Newton's second law, force is proportional to the rate of change of momentum,

$$\text{then } F \propto ma \quad \text{or} \quad F = kma$$

where  $k$  is a constant.

The unit of force is chosen, as that force which will give a mass of 1 kg an acceleration of  $1 \text{ m/s}^2$ . Therefore, by substitution into  $F = kma$ , we get,  $1 = k \times 1 \times 1$  where  $F = \text{force (N)}$  and  $m = \text{mass (kg)}$

$$\therefore F = ma \quad (3.2.1a)$$

**Newton's third law**

States that: *to every action there is an equal and opposite reaction.*

So, for example, the compressive forces that result from the weight of a building, the *action*, are held in equilibrium by the *reaction* forces that occur inside the materials of the building's foundation. Another example is that of propulsion. An aircraft jet engine produces a stream of high velocity gases at its exhaust, the *action*, these act on the airframe of the aircraft causing a *reaction*, which enables the aircraft to accelerate and increase speed for flight.

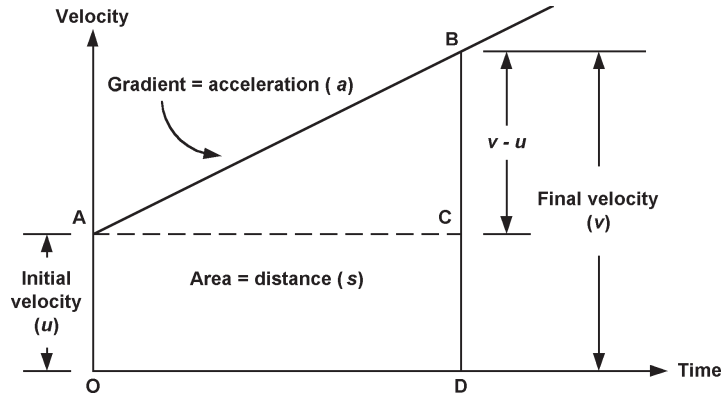
**Linear equations of motion and the velocity/time graph**

You have already been introduced to the concept of a force, velocity, acceleration and now, Newton's laws. These concepts are further exploited through the use of the *linear equations of motion*. Look back now, and remind yourself of the relationship between mass, force, acceleration and Newton's laws.

The linear equations of motion rely for their derivation on the one very important fact that the *acceleration is assumed to be constant*. We will now consider the derivation of the four standard equations of motion using a graphical method.

Even simple linear motion, motion along a straight line, can be difficult to deal with mathematically. However in the case where acceleration is constant it is possible to solve problems of motion by use of a *velocity/*

**Figure 3.2.1** Velocity/time graph



*time graph*, without recourse to the calculus. The equations of motion use standard symbols to represent the variables, these are shown below:

$s$  = distance in metres (m)

$u$  = initial velocity (m/s)

$v$  = final velocity (m/s)

$a$  = acceleration ( $\text{m/s}^2$ )

$t$  = time (s)

The *velocity* is plotted on the vertical axis and *time* on the horizontal axis. Constant velocity is represented by a *horizontal straight line* and acceleration by a *sloping straight line*. Deceleration or *retardation* is also represented by a sloping straight line but with a *negative slope*.

By considering the velocity/time graph shown in Figure 3.2.1, we can establish the equation for distance.

The distance travelled in a given time is equal to the velocity (m/s) multiplied by the time (s), this is found from the graph by the *area under the sloping line*. In Figure 3.2.1, a body is accelerating from a velocity  $u$  to a velocity  $v$  in time  $t$  seconds.

Now the distance travelled  $s = \text{area under graph}$

$$s = ut + \frac{v - u}{2} \times t$$

$$s = ut + \frac{vt}{2} - \frac{ut}{2}$$

$$s = \frac{(2u + v - u)t}{2}$$

$$\text{and so } s = \frac{(u + v)t}{2} \quad (3.2.1b)$$

In a similar manner to above, one of the velocity equations can also be obtained from the velocity/time graph. Since the acceleration is the rate of change of velocity with respect to time, the value of the acceleration will be equal to the gradient of a velocity time graph. Therefore, from Figure 3.2.1, we have:

$$\text{Gradient} = \frac{\text{acceleration}}{\text{time taken}} = \text{velocity}$$

therefore acceleration is given by:

$$a = \frac{v - u}{t} \quad (3.2.2)$$

$$\text{or } v = u + at \quad (3.2.3)$$

The remaining equations of motion may be derived from the two equations found above.

### Example 3.2.1

(a) Transpose the formula for the equation  $a = \frac{v-u}{t}$  for  $t$

(b) Derive the equation  $s = ut + \frac{1}{2}at^2$ , using the above equations.

(c) Derive the equation  $v^2 - u^2 = 2as$

(a) The transposition is exceedingly simple just multiply both sides of the equation by  $t$  and divide both sides by  $a$ , to give:  $t = \frac{v-u}{a}$ .

(b) From formula 3.2.2 we get,  $at = v - u$ , then substituting  $at$  for  $v - u$  in  $s = ut + \frac{v-u}{2} \times t$  gives:  
 $s = ut + \frac{at}{2} \times t$ , and

$$s = ut + \frac{1}{2}at^2 \quad (3.2.4)$$

as required.

(c) From formula 3.2.1 we get,  $s = \frac{1}{2}(v+u)t$  and substituting  $t = \frac{v-u}{a}$  into this equation gives:

$$s = \frac{1}{2}(v+u)\frac{(v-u)}{a} \quad \text{so } 2as = (v+u)(v-u) \text{ and on expansion of the brackets we get:}$$

$$v^2 - u^2 = 2as. \quad (3.2.5)$$

Here is an example of the use of the velocity/time graph.

### Example 3.2.2

A body starts from rest and accelerates with constant acceleration of  $2.0 \text{ m/s}^2$  up to a speed of  $9 \text{ m/s}$ . It then travels at  $9 \text{ m/s}$  for  $15 \text{ s}$  after which time it is retarded to a speed of  $1 \text{ m/s}$ . If the complete motion takes  $24.5 \text{ s}$ , find:

- the time taken to reach  $9 \text{ m/s}$ ,
- the retardation,
- the total distance travelled.

The solution is made easier if we sketch a graph of the motion, as shown in Figure 3.2.2.

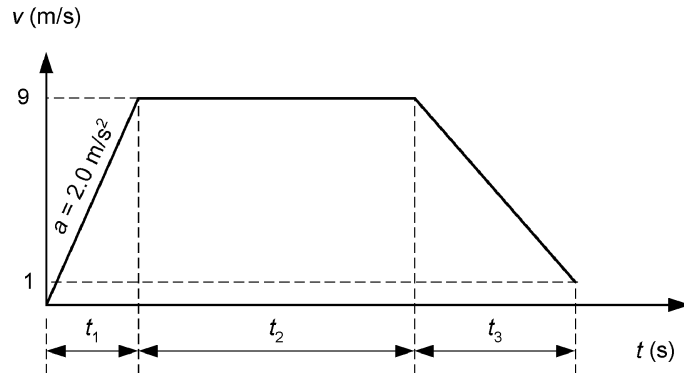


**Figure 3.2.2** Velocity/time graph of the motion

(a) We first tabulate the known values:

$$u = 0 \text{ m/s (we start from rest)}$$

$$v = 9 \text{ m/s}$$

$$a = 2.0 \text{ m/s}^2$$

$$t = ?$$

All we need to do is, select an equation which contains all the variables listed above, that is:

$v = u + at$  and on transposing for  $t$  and substituting the variables we get:

$$t = \frac{9 - 0}{2} \quad \text{so}$$

$$t = 4.5 \text{ s.}$$

(b) The retardation is found in a similar manner.

$$u = 9 \text{ m/s}$$

$$v = 1 \text{ m/s}$$

$$t = 5 \text{ s}$$

$$a = ?$$

We again select an equation which contains the variables, that is,  $v = u + at$  and on transposing for  $a$  and substituting the variables we get

$$a = \frac{1 - 9}{5} \quad \text{so} \quad a = -1.6 \text{ m/s}^2 \quad (\text{the } -\text{ve sign indicates a retardation}).$$

(c) The total distance travelled requires us to sum the component distances travelled for the times  $t_1$ ,  $t_2$ , and  $t_3$ . Again we tabulate the variable for each stage:

$$u_1 = 0 \text{ m/s} \quad u_2 = 9 \text{ m/s} \quad u_3 = 9 \text{ m/s}$$

$$v_1 = 9 \text{ m/s} \quad v_2 = 9 \text{ m/s} \quad v_3 = 1 \text{ m/s}$$

$$t_1 = 4.5 \text{ s} \quad t_2 = 15 \text{ s} \quad t_3 = 5 \text{ s}$$

$$s_1 = ? \quad s_2 = ? \quad s_3 = ?$$

The appropriate equation is  $s = \frac{(u + v)t}{2}$  and in each case we get:

$$s_1 = \frac{(0 + 9)4.5}{2} \quad s_2 = \frac{(9 + 9)15}{2} \quad s_3 = \frac{(9 + 1)5}{2}$$

$$s_1 = 20.25 \quad s_2 = 135 \quad s_3 = 25$$

$$\text{Then total distance } S_T = 20.5 + 135 + 25 = 180.25 \text{ m.}$$

We leave our short revision of linear motion, with one final example.

### Example 3.2.3

A racing car of mass 1965 kg accelerates from 160 to 240 km/h in 3.5 s. If the air resistance is 2000 N/tonne, find the:

- (a) average acceleration,
  - (b) force required to produce the acceleration,
  - (c) inertia force on the car.
- (a) We first need to convert the velocities to standard units.

$$u = 160 \text{ km/h} = \frac{160 \times 1000}{60 \times 60} = 44.4 \text{ m/s}$$

$$v = 240 \text{ km/h} = \frac{240 \times 1000}{60 \times 60} = 66.6 \text{ m/s}$$

also  $t = 3.5 \text{ s}$ , and we are required to find the acceleration  $a$ .

Then using the equation  $v = u + at$  and transposing for  $a$  we get:

$$a = \frac{v - u}{t} \text{ and substituting values } a = \frac{66.6 - 44.4}{3.5}$$

$$a = 6.34 \text{ m/s}^2.$$

- (b) The accelerating force is readily found using Newton's second law, where:

$$F = ma = 1965 \text{ kg} \times 6.34 \text{ m/s}^2 \\ = 12.46 \text{ kN}.$$

- (c) From what has already been said you will be aware that the inertia force = the accelerating force, therefore the inertia force = 12.46 kN.

## Angular motion

You previously met the equations for linear motion. A similar set of equations exists to solve engineering problems that involve angular motion as experienced for example in the rotation of a drive shaft. The linear equations of motion may be transformed to represent angular motion using a set of equations that we will refer to as the *transformation equations*. These are given below, followed by the equations of angular motion, which are compared with their linear equivalents.

### Transformation equations

$$s = r\theta \quad (3.2.6)$$

$$v = r\omega \quad (3.2.7)$$

$$a = r\alpha \quad (3.2.8)$$

where  $r$  is the radius of body from centre of rotation and  $\theta$ ,  $\omega$ , and  $\alpha$  are the angular distance, angular velocity and angular acceleration, respectively.

### Angular equations of motion

Equations (3.2.6)–(3.2.8) provide a relationship between linear and angular distance, velocity, and acceleration respectively. They enable us

to use the linear equations of motion for constant acceleration and transpose them to find their equivalents for angular motion, with constant acceleration. These *angular equations of motion* are summarised with their linear equivalents in the table given below.

| Angular equation of motion                              | Linear equation of motion   |
|---------------------------------------------------------|-----------------------------|
| $\theta = \frac{(\omega_1 + \omega_2)t}{2}$ (3.2.9)     | $s = \frac{(u + v)t}{2}$    |
| $\theta = \omega_1 t + \frac{1}{2} \alpha t^2$ (3.2.10) | $s = ut + \frac{1}{2} at^2$ |
| $\omega_2^2 = \omega_1^2 + 2\alpha\theta$ (3.2.11)      | $v^2 = u^2 + 2as$           |
| $\alpha = \frac{\omega_2 - \omega_1}{t}$ (3.2.12)       | $a = \frac{v - u}{t}$       |

where

$\theta$  = angular distance (rad);

$t$  = time (s);

$\omega_1, \omega_2$  = initial and final angular velocities (rad/s);

$\alpha$  = angular acceleration (rad/s<sup>2</sup>).

### Example 3.2.4

A 540 mm diameter wheel is rotating at  $1500/\pi$  rpm. Determine the angular velocity of wheel in rad/s and the linear velocity of a point on the rim of the wheel.

Now remembering that  $2\pi \text{ rad} = 360^\circ$ . All we need do to find the angular velocity is convert from rpm to rad/s, that is:

Angular velocity (rad/s) =  $1500/\pi \times 2\pi/60 = \mathbf{50 \text{ rad/s}}$ .

Now from the transformation equations, linear velocity,  $v = \text{angular velocity, } \omega \times \text{radius, } r = 50 \text{ rad/s} \times 0.270 \text{ m}$  so  $\mathbf{v = 13.5 \text{ m/s}}$ .

Here is an example that involves angular acceleration.

### Example 3.2.5

A pinion shown in Figure 3.2.3 is required to move with an initial angular velocity of 300 rpm and final angular velocity of 600 rpm. If the increase takes place over 15 s, determine the linear acceleration of the rack. Assume a pinion radius of 180 mm.

In order to solve this problem we first need to convert the velocities into radians per second.

$300 \text{ rpm} = 300 \times 2\pi/60 = 31.4 \text{ rad/s}$

$600 \text{ rpm} = 600 \times 2\pi/60 = 62.8 \text{ rad/s}$ .

We can use the equation  $\alpha = \frac{\omega_1 - \omega_2}{t}$  to find the angular acceleration.

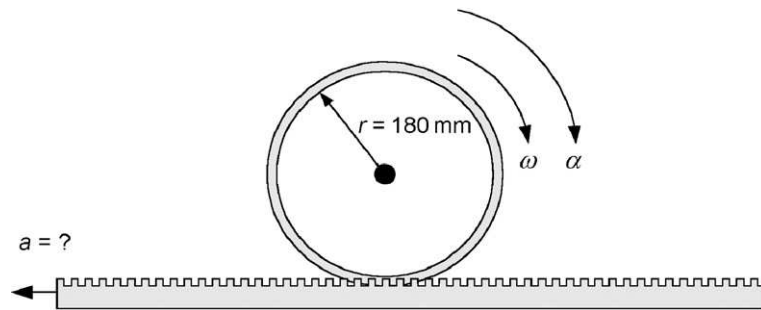


Figure 3.2.3 Rack and pinion

$$\text{So } \alpha = \frac{62.8 - 31.4}{15} = \mathbf{2.09 \text{ rad/s}^2}$$

Now we can use the transformation equation  $a = r\alpha$  to find the linear acceleration, that is:

$$a = (0.18 \text{ m})(2.09 \text{ rad/s}) = \mathbf{0.377 \text{ m/s}^2}.$$

Here are two more examples that use the angular equations of motion.

### Example 3.2.6

The armature of an electric motor rotating at 1500 rpm accelerates uniformly until it reaches a speed of 2500 rpm. During the accelerating period the armature makes 300 complete revolutions. Determine the angular acceleration and the time taken.

Solving this type of problem is best achieved by first writing down the *knowns* in the correct SI units.

So in this case we have,

initial angular velocity of armature = 1500 rpm or

$$\omega_1 = \frac{1500 \times 2\pi}{60} = 157 \text{ rad/s.}$$

Make sure you can convert rpm into rad/s and vice-versa!

$$\text{Similarly, } \omega_2 = \frac{2500 \times 2\pi}{60} = 262 \text{ rad/s}$$

$$\theta = 300 \times 2\pi = 1885 \text{ rad/s.}$$

Using our angular motion equations, we must select the equation which allows us to use the most information and enables us to find one or more of the unknowns.

Then using the equation

$$\omega_2^2 - \omega_1^2 = 2\alpha\theta \text{ to find } \alpha \text{ rad/s}^2$$

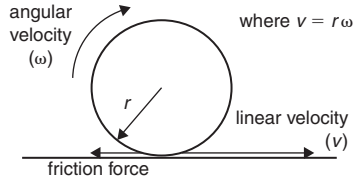
$$262^2 - 157^2 = 2\alpha(1885)$$

$$\text{from which } \alpha = 11.7 \text{ rad/s}^2$$

and using the equation  $\omega_2 = \omega_1 + \alpha t$  to find the time (s)

then,  $262 = 157 + 11.7t$  which gives

$$\mathbf{t = 9 \text{ s.}}$$



*Relationship between linear and angular velocity*

### Example 3.2.7

An aircraft sits on the runway ready for take-off. It has 1.4 m diameter wheels and accelerates uniformly from rest to 225 km/h (take-off speed) in 40 s.

Determine:

- (i) the angular acceleration of the undercarriage wheels;
- (ii) the number of revolutions made by each wheel during the take-off run.

Apart from identifying all the knowns in the correct SI units, in this example it will also be necessary to consider a combination of linear and angular motion. If we study figure (shown on left) we note that, in general, the angular rotation of the wheel causes linear motion along the ground, provided there are frictional forces sufficient to convert the rotating (torque) at the wheel into linear motion.

We have,  $v_1 = 0$  and  $v_2 = \frac{225 \times 1000}{60 \times 60} \text{ m/s} = 62.5 \text{ m/s}$ .

Then using the transformation equation for velocity we have,

$$v = r\omega \text{ and so } \omega_1 = 0 \text{ and } \omega_2 = \frac{62.5}{0.7} = 89.29 \text{ rad/s}$$

and the angular acceleration may be found using the equation:

$$\omega_2 = \omega_1 + \alpha t, \text{ that is } \frac{89.29}{40} = \alpha = 2.23 \text{ rad/s}^2$$

and the number of radians turned through by each wheel may be found using the equation:

$$\theta = \frac{1}{2}(\omega_1 + \omega_2)t \text{ then } \theta = \frac{1}{2}(0 + 89.29)40 = 1785.8 \text{ rads.}$$

So the number of revolutions turned through by each wheel

$$\text{is } \frac{1785.8}{2\pi} = 284.2 \text{ revs.}$$

## Torque, moment of inertia and angular motion

### Torque

From Newton's third law, we know that to accelerate a mass we require a force such that:

$$F = ma$$

Now in dealing with angular motion, we know that this force would be applied at a radius  $r$ , from the centre of rotation and would thus create a turning moment or more correctly a **torque**  $T$ , thus:

$$T = Fr \text{ or } T = mar$$

Since the linear acceleration,  $a = r\alpha$ , then:

$$T = m(r\alpha)r$$

or

$$T = m\alpha r^2 \quad (3.2.13)$$



**Moment of inertia and angular motion**

Now from our equation for torque the quantity  $mr^2$  has a special significance. It is known as the **moment of inertia of the body about its axis of rotation**. It is given the symbol  $I$ , thus:  $I = mr^2$  and the units of  $I$  are ( $\text{kg m}^2$ ), this is because the inertia of a body, from Newton's first law, is proportional to its mass multiplied by the distance squared. A mathematical derivation of the moment of inertia will be found in Chapter 6, Section 3. It should be remembered *that the moment of inertia of a rotating body is equivalent to the mass of a body in linear motion*.

Now from Equation (3.2.13), we may write the equation for torque as:

$$T = I\alpha \quad (3.2.14)$$

The axis of rotation if not stated, is normally obvious. For example, a flywheel or electric motor rotates about its centre, which we refer to as its *polar axis*. When giving values of  $I$  they should always be stated with respect to the reference axis.

**Example 3.2.8**

A flywheel has a moment of inertia of  $130 \text{ kg m}^2$ . Its angular velocity drops from 12 000 to 9000 rpm in 6 s, determine (a) the retardation (b) the braking torque.

$$\begin{aligned} \text{Then, } \omega_1 &= 12\,000 \times 2\pi/60 = 1256.6 \text{ rad/s} \\ \omega_2 &= 9000 \times 2\pi/60 = 942.5 \text{ rad/s} \end{aligned}$$

$$\begin{aligned} \text{and from } \alpha &= \frac{\omega_2 - \omega_1}{t} \\ &= \frac{942.5 - 1256.6}{6} \\ &= -52.35 \text{ or retardation} = 52.35 \text{ r/s}^2 \end{aligned}$$

$$\begin{aligned} \text{Now torque } T &= I\alpha \\ T &= (130)(52.35) \\ \text{so braking torque } T &= 6805.5 \text{ Nm.} \end{aligned}$$

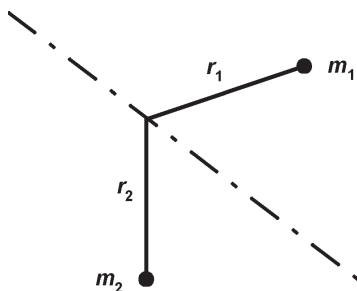


**Figure 3.2.4** Rotating masses concentrated at a point

**Radius of gyration ( $k$ )**

In order to use the above definition for the moment of inertia  $I$ , we need to be able to determine the radii at which the mass or masses are situated, from the centre of rotation of the body (see Figure 3.2.4).

For most engineering components the mass is *distributed* and *not* concentrated at any particular radius, so we need some way of finding an equivalent radius about which the whole mass of the rotating body is deemed to act.

The radius of gyration  $k$  is the radius at which a concentrated mass  $M$  (equivalent to the whole mass of the body) would have to be situated so

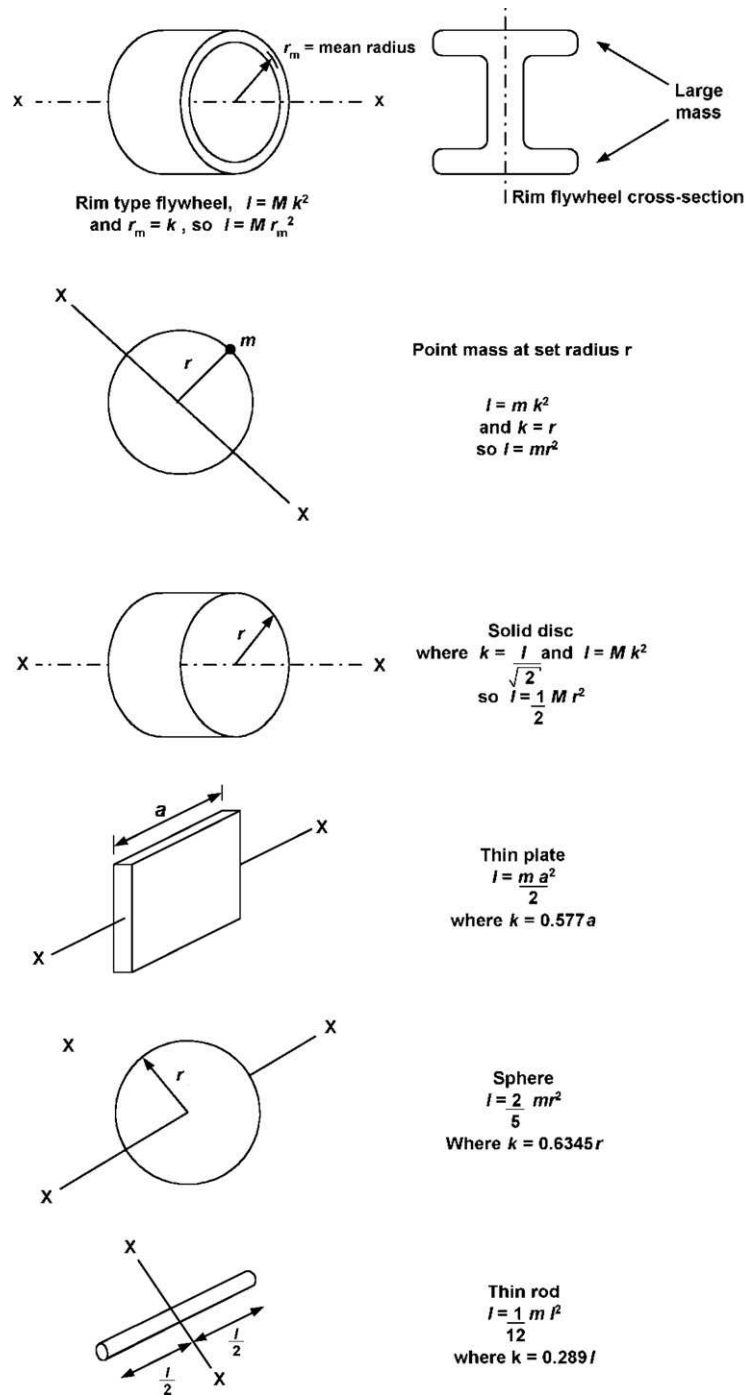


**Figure 3.2.5** Definition of the moment of inertia for some engineering components

that its moment of inertia is equal to that of the body. So the moment of the body

$$I = M k^2 \quad (3.2.15).$$

All of this might, at first, appear a little confusing! In practice tables of values of  $k$  for common engineering shapes may be used. To enable you to tackle problems involving the inertia of rotating bodies, the value of  $k$  for some commonly occurring situations are given in Figure 3.2.5.

**Example 3.2.9**

A solid cylinder having a total mass of 140 kg and a diameter of 0.4 m, is free to rotate about its polar axis. It accelerates from 750 to 1500 rpm in 15 s. There is a resistance to motion set up by a frictional torque of 1.1 Nm as a result of worn bearings. Find the torque which must be applied to the cylinder to produce the motion.

We need to find the torque  $T$  from  $T = I\alpha$

where for a solid disc  $I = \frac{1}{2} Mr^2$  (Figure 3.2.4).

The angular acceleration can be found using the Equation  $\omega_f = \omega_i + \alpha t$ .

So  $\omega_f = \omega_i + \alpha t$  where  $\omega_i = \frac{750 \times 2\pi}{60} = 78.5 \text{ rad/s}$  and

$\omega_f = \frac{1500 \times 2\pi}{60} = 157 \text{ rad/s}$  and  $t = 15 \text{ s}$ , then:

$157 - 78.5 = \alpha 15$  giving  $\alpha = 5.23 \text{ rad/s}^2$ .

Now for our solid circular disc  $I = \frac{1}{2} Mr^2$

$I = \frac{1}{2} (140)(0.2)^2 = 2.8 \text{ kg m}^2$

and from  $T = I\alpha$  we have:

net accelerating torque  $T = (2.8)(5.23) = 14.64 \text{ Nm}$

then, since net accelerating torque = applied torque – friction torque

we have  $14.64 = \text{applied torque} - 1.1$

giving, **applied torque = 15.74 Nm.**

**Problems 3.2.1**

1. A wheel has a diameter of 0.54 m and is rotating at  $\frac{1500}{\pi}$  rpm. Calculate the angular velocity of the wheel in rad/s and the linear velocity of a point on the rim of the wheel.
2. A flywheel rotating at 20 rad/s increases its speed uniformly to 40 rad/s in 1.0 min. Sketch the angular velocity/time graph and determine:
  - (a) the angular acceleration of the flywheel
  - (b) the angle turned through by the flywheel in 1.0 min and so calculate the number of revolutions made by the flywheel in this time.
3. The flywheel of a cutting machine has a moment of inertia of  $130 \text{ kg m}^2$ . Its speed drops from 120 rev/min to 90 rev/min in 2 s. Determine:
  - (a) the deceleration of the flywheel,
  - (b) the braking torque.

4. A turbine and shaft assembly has a mass moment of inertia of  $15 \text{ kg m}^2$ . The assembly is accelerated from rest by the application of a torque of  $40 \text{ Nm}$ . Determine the speed of the shaft after  $20 \text{ s}$ .
5. A rim type flywheel gains  $2.0 \text{ kJ}$  of energy when its rotational speed is raised from  $250$  to  $270 \text{ rpm}$ . Find the required inertia of the flywheel.
6. The armature of an electric motor has a mass of  $50 \text{ kg}$  and a radius of gyration of  $150 \text{ mm}$ . It is retarded uniformly by the application of a brake, from  $2000$  to  $1350 \text{ rpm}$ , during which time ( $t$ ) armature makes  $850$  complete revolutions. Find the retardation and braking force.

## Energy transfer

### Introduction

In this section we consider some forms of mechanical energy including, gravitational potential energy, strain potential energy, linear kinetic energy and angular kinetic energy. In addition we consider the relationship between work and energy transfers in linear and angular motions systems. We also consider the energy transfers that take place during impact.

Energy may exist in many different forms, for example, mechanical, electrical, nuclear, chemical, heat, light and sound.

*The principle of the conservation of energy states that: energy may neither be created nor destroyed only changed from one form to another.*

There are many engineering examples of devices that transform energy, these include the:

- loudspeaker which transforms electrical to sound energy,
- petrol engine which transforms heat to mechanical energy,
- microphone which transforms sound to electrical energy,
- dynamo transforms mechanical to electrical energy,
- battery transforms chemical to electrical energy,
- filament bulb transforms electrical to light energy.

Here, as mentioned earlier, we are only concerned with the various forms of mechanical energy and its conservation. Provided no mechanical energy is transferred to or from a body, the total amount of mechanical energy possessed by a body remains constant, unless mechanical work is done, this concept is looked at next.

### Work done

*The energy possessed by a body is its capacity to do work. Mechanical work is done when a force overcomes a resistance and it moves through a distance.*

Mechanical work may be defined as:

*Mechanical work done ( $J$ ) = force required to overcome the resistance ( $N$ )  $\times$  Distance moved against the resistance ( $m$ ), that is*

$$W = F \times d \quad (3.2.16)$$

*The SI unit of work is the Newton-metre ( $\text{Nm}$ ) or Joule where  $1 \text{ J} = 1 \text{ Nm}$ .*

*Note:*

- (a) No work is done unless there is both resistance and movement.
- (b) The resistance and the force needed to overcome it are equal.
- (c) The distance moved must be measured in exactly the opposite direction to that of the resistance being overcome.

The more common resistances to be overcome include: *friction*, *gravity* (the weight of the body itself) and *inertia* (the resistance to acceleration of the body) where:

The work done (WD) against friction = friction force  $\times$  distance moved

WD against gravity = weight  $\times$  gain in height

WD against inertia = inertia force  $\times$  distance moved

*Note:*

- (a) Inertia force is the out of balance force multiplied by the distance moved or: the *inertia force* = *mass*  $\times$  *acceleration*  $\times$  *distance moved*
- (b) WD in overcoming friction will be discussed in more detail later.

In any problem involving calculation of WD, the first task should be to identify the type of resistance to overcome. If, and only if, there is motion between surfaces in contact, is WD against friction. Similarly, only where there is a gain in height is there WD against gravity and only if a body is accelerated is WD against inertia (look back at our definition of inertia).

### Example 3.2.10

A body of mass 30 kg is raised from the ground at constant velocity through a vertical distance of 15 m. Calculate the WD.

If we ignore air resistance, then the only WD is against gravity.

WD against gravity = weight  $\times$  gain in height or

WD =  $mgh$  (and assuming  $g = 9.81 \text{ m/s}^2$ )

then WD =  $(30)(9.81)(15)$

**WD = 4414.5 J or 4.414 kJ**

Work done may be represented graphically and, for linear motion, this is shown in Figure 3.2.6(a). Where the force needed to overcome the resistance, is plotted against the distance moved. The WD is then given by the area under the graph.

Figure 3.2.6(b) shows the situation for angular motion, where a varying torque  $T$  in  $\text{Nm}$  is plotted against the angle turned through in *radian*. Again the WD is given by the area under the graph, where the units are  $\text{Nm} \times \text{radian}$ . Then noting that the radian has no dimensions, the unit for work done remains as  $\text{Nm}$  or *Joules*.

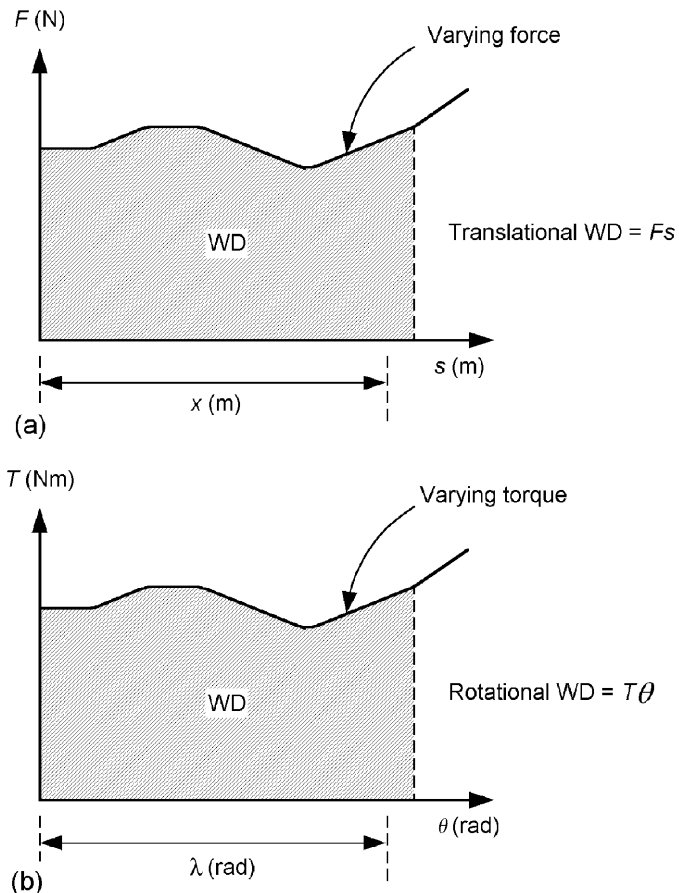


Figure 3.2.6 Work done

### Potential energy

*Potential energy (PE) is energy possessed by a body by virtue of its position, relative to some datum. The change in PE is equal to its weight multiplied by the change in height. Since the weight of a body =  $mg$ , then the change in PE may be written as:*

$$\text{Change in PE} = mg\Delta h \quad (3.2.17)$$

which of course is identical to the work done in overcoming gravity. So the work done in raising a mass to a height is equal to the PE it possesses at that height, assuming no external losses.

#### Example 3.2.11

A packing crate, weighing 150 kg, is moved 100 m by a frictionless conveyor system. If the conveyor system is inclined at  $30^\circ$  to the horizontal, what work is done?

Assuming that there are no losses due to friction, then the only work that need be done, is against gravity. Now gravitational

acceleration acts vertically down, so we need to find the *vertical height* moved by the crate.

Since the conveyor system is inclined at  $30^\circ$  to the horizontal, then the vertical component of height is given by,  $h = 100 \sin 30$  or  $\Delta h = (100)(0.5) = 50$  m. So the change in height is 50 m and the gravitational PE is:

$PE = mg\Delta h$  or  $PE = (150)(9.81)(50)$  so the **PE = 73 575 J** or 73.575 kJ.

### Strain energy

*Strain energy* is a particular form of PE possessed by an elastic body that is deformed within its elastic range, for example a stretched or compressed spring possesses strain energy. Consider the spring arrangement shown in Figure 3.2.7. The force required to compress or extend the spring is  $F = kx$ , where  $k$  is the spring constant.

Figure 3.2.7(a) shows a helical coil spring in the unstrained, compressed and extended positions. The force required to move the spring varies in direct proportion to the distance moved (Figure 3.2.7(b)). Therefore *strain energy of spring when compressed or extended* = area under graph (*force  $\times$  distance moved*) =  $\frac{1}{2}Fx$  J

and since  $F = kx$ , then substituting for  $F$  gives,

$$\text{Strain energy of spring in tension or compression} = \frac{1}{2} kx^2 \quad (3.2.18)$$

A similar argument can be given for a spring which is subject to twisting or torsion about its centre (or polar axis). It can be shown that:

$$\text{Strain energy of a spring when twisted} = \frac{1}{2} T\theta = \frac{1}{2} k_{\text{tor}}\theta^2 \quad (3.2.19)$$

where  $\theta$  = the angle of twist,  $t$  = the torque (Nm) and  $k_{\text{tor}}$  is the spring stiffness or torque rate, that is the torque required to produce unit angular displacement.

Note that the unit of strain energy, as for all other types of energy is the *Joule* (J).

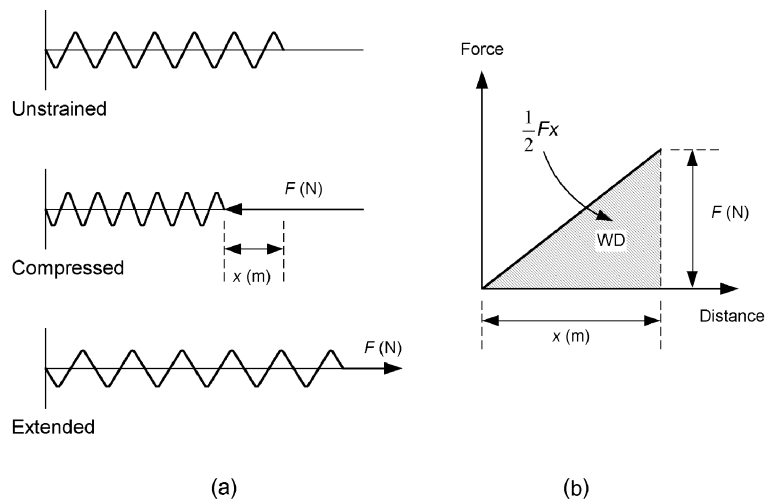


**Figure 3.2.7** Spring system demonstrating strain energy

**Example 3.2.12**

A drive shaft is subject to a torque rate of 30 MN/rad. What is the strain energy set up in the drive shaft when the angle of twist is  $5^\circ$ .

All we need do here is use the relationship: strain energy  $= \frac{1}{2} k_{\text{tor}} \theta^2$ . Remembering that the angle of twist must be in radian, then:

$$5^\circ = \frac{(5)(2\pi)}{360} = 0.087 \text{ rad, then}$$

$$\text{strain energy} = \frac{1}{2} (30 \times 10^6) (0.087)^2 = \mathbf{113.535 \text{ kJ.}}$$

**Kinetic energy**

*Kinetic energy (KE)* is energy possessed by a body by virtue of its motion. Translational KE, that is the KE of a body travelling in a linear direction (straight line), is:

$$\text{Translational KE (J)} = \frac{\text{mass (kg)} \times (\text{velocity})^2 (\text{m/s})}{2}$$

$$\text{Translation (linear) KE} = \frac{1}{2} mv^2 \quad (3.2.20)$$

Flywheels are heavy wheel shaped masses fitted to shafts in order to minimise sudden variations in the rotational speed of the shaft, due to sudden changes in load. A flywheel is therefore a store of rotational KE.

Rotational KE can be defined in a similar manner to translational (linear) KE, that is

$$\text{Rotational KE of mass} = \frac{1}{2} I\omega^2 \quad (3.2.21)$$

where  $I$  = mass moment of inertia (which you met earlier).

**Example 3.2.13**

Find the KE possessed by a railway train of mass 30 tonne, moving with a velocity of 120 km/h.

Note first, that in a lot of these examples we need to convert kph into m/s. This is easy, if you remember that there are 1000 m in a km, so that 120 km/h = (120)(1000) m/h and that there are 3600 s in an hour,

$$\therefore 120 \text{ km/h} = \frac{(120)(1000)}{3600} = \frac{120}{3.6} = 33.33 \text{ m/s}$$

Then, remembering that 1 tonne = 1000 kg, the KE of the train is given by:

$$\text{KE} = \frac{1}{2} mv^2 = (0.5)(30 \times 10^3)(33.33)^2 \quad \text{or} \quad \mathbf{KE = 16.67 \text{ MJ.}}$$



**Example 3.2.14**

Determine the total KE of a four wheel drive car which has a mass of 800 kg and is travelling at 50 kph. Each wheel of the car has a mass of 15 kg, a diameter of 0.6 m and a radius of gyration of 0.25 m.

The total KE = translational (linear) KE + angular KE

$$\text{and Linear KE} = \frac{1}{2}mv^2 \quad \text{where } v = 50 \text{ km/h} = 13.89 \text{ m/s}$$

$$= \frac{1}{2}(800)(13.89)^2 = 77.16 \text{ kJ}$$

$$\text{and angular KE} = \frac{1}{2}I\omega^2 \quad \text{where } I = Mk^2 = (15)(0.25)^2 = 0.9375 \text{ kg m}^2 \text{ (for each wheel!)} \text{ and from } v = \omega r \text{ then}$$

$$\omega = v/r = 13.89/0.3 = 46.3 \text{ rad/s}$$

$$= \frac{1}{2}(4 \times 0.9375)(46.3)^2$$

$$= 4.019 \text{ kJ}$$

Therefore total KE of the car = 77.16 + 4.019 = **81.18 kJ**.

**Conservation of mechanical energy**

From the definition of the conservation of energy we can deduce that the total amount of energy within certain defined boundaries, will remain the same. When dealing with mechanical systems, the potential energy (PE) possessed by a body is frequently converted into KE and vice versa. If we ignore air frictional losses, then we may write:

$$\text{PE} + \text{KE} = \text{a constant}$$

Thus, if a mass  $m$  falls freely from a height  $h$  above some datum, then at any height above that datum:

$$\text{Total energy} = \text{PE} + \text{KE}$$

This important relationship is illustrated in Figure 3.2.8, where at the highest level above the datum the PE is a maximum and is gradually converted into KE, as the mass falls towards the datum, immediately before impact when height  $h = 0$ , the KE is zero and the KE is equal to the initial PE.

Since the total energy is constant then:

$$mgh_1 = mgh_2 + \frac{1}{2}mv_2^2 = mgh_3 + \frac{1}{2}mv_3^2 = \frac{1}{2}mv_4^2$$

Immediately after impact with the datum surface, the mechanical KE is converted into other forms such as, heat, strain and sound energy.

If friction is present then work is done overcoming the resistance due to friction and this is dissipated as heat. Then:

$$\text{Initial energy} = \text{Final energy} + \text{Work done in overcoming frictional resistance}$$

*Note:* KE is not always conserved in collisions. Where KE is conserved in a collision we refer to the collision as *elastic*, when KE is not conserved we refer to the collision as *inelastic*.

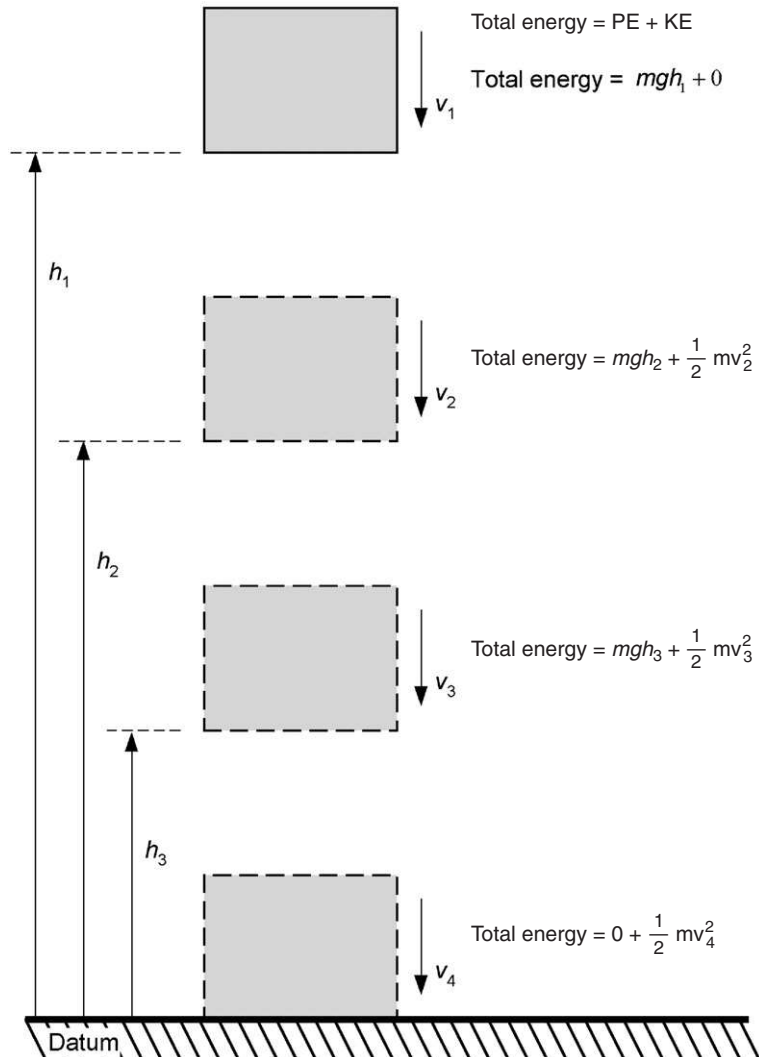


**Figure 3.2.8**  $PE + KE =$   
a constant

### Example 3.2.15

Cargo weighing 2500 kg breaks free from the top of the cargo ramp (Figure 3.2.9). Ignoring friction, determine the velocity of the cargo the instant it reaches the bottom of the ramp.

The vertical height  $h$  is found using the sine ratio, that is:

$$10 \sin 10 = h \quad \text{so} \quad h = 1.736 \text{ m}$$

$$\begin{aligned} \text{so increase in PE} &= mgh \\ &= (2500)(9.81)(1.736) \text{ J} \\ &= 42\,575.4 \text{ J} \end{aligned}$$

Now using the relationship,  $PE + KE = \text{total energy}$ . Then immediately prior to the cargo breaking away  $KE = 0$  and so  $PE = \text{total energy}$  also, immediately prior to the cargo striking the base of slope  $PE = 0$  and  $KE = \text{total energy}$  (all other energy losses being ignored).

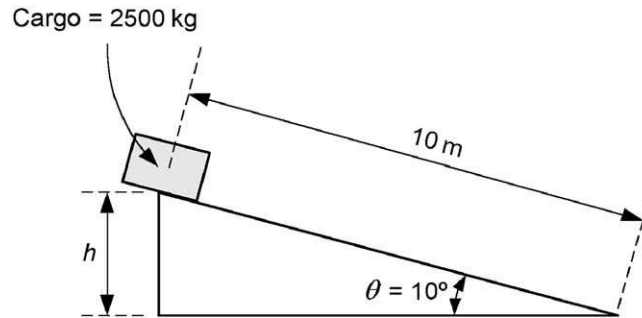


Figure 3.2.9 Cargo ramp

So at base of slope:  $42\,575.4 \text{ J} = \text{KE}$

and  $42\,575.4 = \frac{1}{2}mv^2$

so  $\frac{(2)(42\,575.4)}{2500} = v^2$

and so velocity at bottom of ramp = **5.84 m/s**.

### Impact and linear momentum

If a motor vehicle crashes into a stationary object such as a concrete post. The KE due to its motion is absorbed as work done in trying to alter the shape of post! This, in turn, produces energy in the form of heat and sound which is absorbed by the environment. Therefore, no energy has been lost due to the collision but transferred from one form into another, in accordance with the conservation of energy law. In order, as engineers to determine the loss of KE at impact, we first need to return to the concept of *momentum*, you meet in Newton's laws.

The law of the conservation of linear momentum states that *the total momentum of a system, in a particular direction, remains constant providing that the masses are unaltered and no external forces act on the system*.

The above law may be applied to collisions and stated more simply as:

*The total momentum before impact = Total momentum after impact*

Consider two bodies of different masses ( $m_1$ ,  $m_2$ ) that collide and then move off together after impact. Remembering that:

$$\text{Linear momentum} = mv \quad (3.2.22)$$

then total momentum before impact =  $m_1v_1 + m_2v_2$  and

Total momentum after impact =  $m_1v_a + m_2v_a$  (where in this case,  $v_a$  = the velocity after impact, which because they move off together will be the same for both masses.

So these quantities may be equated, to give:

$$m_1v_1 + m_2v_2 = m_1v_a + m_2v_a$$

Also from the law of the conservation of energy (assuming a perfect elastic collision), then:

$$\text{Loss of KE at impact} = \text{total KE before impact} - \text{total KE after impact}$$

**Example 3.2.16**

A curling stone of mass 20.5 kg travelling at 4 m/s collides with another curling stone of mass 20.0 kg, travelling at 3 m/s in the opposite direction. Find the loss in KE if the two stones remain together directly after impact.

Then using the momentum relationship

$$m_1 v_1 + m_2 v_2 = m_1 v_a + m_2 v_a,$$

we have:

$$(20.5)(4) + (20)(3) = 20.5 v_a + 20 v_a, \quad \text{from which we find} \\ v_a = 3.506 \text{ m/s}$$

We now have velocities before and after impact, so using:

Loss of KE at impact = Total KE before impact – Total KE after impact, we get:

Loss of KE

$$= \left[ \frac{1}{2}(20.5)(4)^2 + \frac{1}{2}(20.0)(3)^2 \right] - \frac{1}{2}(40.5)(3.506)^2$$

and on simplification,

**Loss of KE** = 164 + 90 – 249 = **5 J**. This loss may be due to energy transfer on impact such as heat and sound.

**Rotational kinetic energy and angular momentum**

From Formula 3.2.21 where, the rotational KE of mass =  $\frac{1}{2} I \omega^2$  and the knowledge that bodies lose KE when they collide. We can write a relationship for the loss of rotational KE after impact, similar to the relationship for linear KE, that is:

The loss of rotation KE at impact = total rotational KE before impact – total rotational KE after impact

and from the above formula this may be expressed as:

$$\text{The loss of rotational KE at impact} = \left[ \frac{1}{2} I_1 \omega_1^2 + \frac{1}{2} I_2 \omega_2^2 \right] \\ - \left[ \frac{1}{2} (I_1 + I_2) \omega_a^2 \right]$$

The law for the conservation of angular momentum, is given as: *The total angular momentum of a mass system, rotating in a particular direction, remains constant; providing that the moments of inertia of the rotating masses remain unaltered and that no external torques act on the system.*

So this law may be applied to collisions, where the moment of inertia of the individual masses remains constant, after collision.

Then, as with the linear case:

*Total angular momentum before impact = Total angular momentum after impact*

$$\text{or } I_1 \omega_1 + I_2 \omega_2 = (I_1 + I_2) \omega_a \quad \text{where angular momentum} = I \omega \quad (3.2.23)$$

(Compare with linear case!)

Let us look at a couple of examples that should clarify the use of the above relationships.

**Example 3.2.17**

Determine the KE energy of a flywheel having a mass of 250 kg and radius of gyration 0.3 m, when rotating at 800 rpm.

Then we may treat the flywheel as a solid disc, where from Figure 3.2.5, we find that the moment of inertia

$$I = Mk^2 \text{ or } I = (250)(0.3)^2 \text{ and } I = 22.5 \text{ kg m}^2.$$

$$\text{Also, } 800 \text{ rpm } \frac{(800)2\pi}{60} = 83.78 \text{ rad/s}$$

$$\text{Then the rotational KE} = \frac{1}{2}I\omega^2 = (0.5)(22.5)(83.78)^2 \text{ and}$$

rotational **KE = 78.96 kJ.**

**Example 3.2.18**

An electric motor drives a co-axial gearbox, through a clutch. The motor armature has a mass of 50 kg and a radius of gyration of 18 cm, while the mass of the gearbox assembly is 500 kg, with a radius of gyration of 24 cm. If the motor armature is revolving at 3000 rpm and the gearbox is rotating at 500 rpm in the *same* direction. What is the loss of KE when the clutch is engaged.

Let us first find the mass moment of inertia for the motor armature and for the gearbox. Where  $I = mk^2$ .

Then for motor armature  $I = (50)(0.18)^2 = 1.62 \text{ kg m}^2$  and for the gearbox,  $I = (500)(0.24)^2 = 28.8 \text{ kg m}^2$

Also the angular velocities are respectively:

$$\frac{(3000)2\pi}{60} = 314.16 \text{ rad/s} \text{ and } \frac{(500)2\pi}{60} = 52.34 \text{ rad/s}$$

Then to find the angular velocity after impact, we use the relationship for the conservation of angular momentum, that is:

$$\text{Total angular momentum before impact} = \text{Total angular momentum after impact}$$

where *angular momentum* =  $I\omega$  (compare with linear case)

$$\text{Then } (1.62)(314.16) + (28.8)(52.34) = (1.62 + 28.8)\omega_a$$

where we find that  $\omega_a = 66.28 \text{ rad/s}$ . We are now in a position to find the loss in KE at impact. Then from:

$$\text{The loss of rotational KE at impact} = \left[ \frac{1}{2}I\omega_1^2 + \frac{1}{2}I\omega_2^2 \right] - \left[ \frac{1}{2}(I_1 + I_2)\omega_a^2 \right].$$

We get:

$$\begin{aligned} \text{Loss of rotational KE} &= (0.5)(1.62)(314.16)^2 + (0.5)(28.8)(52.34)^2 - 0.5(30.42)(66.28)^2 \\ &= 79\,944.16 + 39\,448.45 - 66\,818.11 \\ &= 52\,574.5 \text{ J} \end{aligned}$$

Then, loss of rotational KE at impact = **52.575 kJ.**

## Power

Power is a measure of the rate at which work is done or the rate of change of energy. Power is therefore defined as: *the rate of doing work*. The SI unit of power is the watt (W) i.e.

$$\text{Power (W)} = \frac{\text{work done (J)}}{\text{time taken (s)}} = \frac{\text{energy change (J)}}{\text{time taken (s)}} \quad (3.2.24)$$

or, if the body moves with constant velocity,

$$\text{Power (W)} = \text{force used (N)} \times \text{velocity (m/s)} \quad (3.2.25)$$

*Note:* Units are Nm/s = J/s = watt (W).

### Example 3.2.19

A packing crate weighing 1000 N is loaded onto the back of a lorry by being dragged up an incline of 1 in 5 at a steady speed of 2 m/s. The frictional resistance to motion is 240 N. Calculate:

- the power needed to overcome friction,
- the power needed to overcome gravity,
- the total power needed.

$$\begin{aligned} \text{(a) Power} &= \text{friction force} \times \text{velocity along surface} \\ &= 240 \times 2 \\ &= \mathbf{480 \text{ W}} \end{aligned}$$

$$\begin{aligned} \text{(b) Power} &= \text{weight} \times \text{vertical component of velocity} \\ &= 1000 \times 2 \times 1/5 \\ &= \mathbf{400 \text{ W}} \end{aligned}$$

- Since there is no acceleration and therefore no work done against inertia,

$$\begin{aligned} \text{Total power} &= \text{power for friction} + \text{power for gravity} \\ &= 480 + 400 \\ &= \mathbf{880 \text{ W}}. \end{aligned}$$

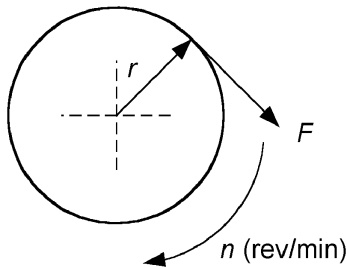


**Figure 3.2.10** Power transmitted by a torque

### Power transmitted by a torque

Let us now consider power transmitted by a torque. You have already met the concept of torque (Figure 3.2.10).

Since the WD is equal to the force multiplied by the distance, then the WD in 1 rev is given by:

$$\text{WD in 1 rev} = F \times 2\pi r \text{ J}$$

but  $Fr$  is the torque  $T$  applied to the shaft, therefore the WD in 1 rev is,

$$\text{WD in 1 rev} = 2\pi T \text{ J.}$$

In 1 min,

$$\begin{aligned} \text{WD} &= \text{WD/rev} \times \text{number of rev/min (n)} \\ &= 2\pi nT \end{aligned}$$

and WD in 1s =  $2\pi nT/60$  and since WD per second is equal to power (1J/s = 1W).

Then,

Power (W) transmitted by a torque =  $2\pi nT/60$

You have already met the Engineers Theory of Twist, in your study of Torsion in circular shafts, where you will remember the relationship:  $\tau/r = T/J$ , where:

$\tau$  = the shear stress at a distance  $r$  from the polar axis of the shaft,

$T$  = the twisting moment on the shaft,

$J$  = the polar second moment of cross-sectional area of the shaft.

We can use this relationship to calculate power transmitted by a torque in a shaft.

### Example 3.2.20

Calculate the power transmitted by a hollow circular shaft, with an external diameter of 250 mm, when rotating at 100 rpm. If the maximum shear stress is  $70 \text{ MN/m}^2$  and  $J = 220 \times 10^6 \text{ mm}^4$ .

Then from engineers theory of twist

$$T = \frac{(70)(220 \times 10^6)}{125} = 123.2 \times 10^6 \text{ Nm or}$$

$$T = 123.2 \times 10^3 \text{ Nmm}$$

Note that we used the relationship  $70 \text{ MN/m}^2 = 70 \text{ N/mm}^2$ , to keep the unit consistent!

Then we know that, the power transmitted by a torque,  $W = 2\pi nT/60$ . Therefore:

$$\text{Power} = \frac{(2\pi)(100)(123.2 \times 10^3)}{60} = \mathbf{1.2901 \text{ MW.}}$$

### Problems 3.2.2

1. A crane raises a load of 1640 N to a height of 10 m in 8 s. Calculate the average power developed.
2. The scale of a spring balance which indicates weights up to 20 N extends over a length of 10 cm. Calculate the work done in pulling the balance out until it indicates 12 N.
3. A train having a mass of 15 ton is brought to rest when striking the buffers of a terminus. The buffers consist of two springs in parallel, each having a spring constant of 120 kN/m and able to be compressed to a maximum of 0.75 m. Find:
  - (i) the strain energy gained by the buffers.
  - (ii) the velocity of the train at the instant it strikes the buffers.
4. Find the KE of a mass of 2000 kg moving with a velocity of 40 km/h.

5. A motor vehicle starting from rest, free wheels down a slope whose gradient is 1 in 8. Neglecting all resistances to motion, find its velocity after travelling a distance of 200 m down the slope.
6. A railway carriage of mass 9000 kg is travelling at 10 kph, when it is shunted by a railway engine of mass 12000 kg travelling at 15 kph in the same direction. After the shunt they move off together in the same direction.
  - (a) What is the velocity of engine and carriage, immediately after the shunt?
  - (b) What is the loss of KE, resulting from the shunt?
7. A rotating disc, has a mass of 20 kg and a radius of gyration of 45 mm. It is brought to rest from 1500 rpm in 150 revolutions by a braking torque. Determine:
  - (a) the angular retardation.
  - (b) the value of the braking torque.

## Oscillating mechanical systems

---

In this short section we will look at simple harmonic motion and see how this motion is applied to linear and transverse systems. The effects of forcing and damping on oscillating systems will also be discussed.

### Simple harmonic motion

When a body oscillates backwards and forwards so that every part of its motion recurs regularly, we say that it has *periodic motion*. For example a piston attached to a connecting rod and crankshaft, moving up and down inside the cylinder, has periodic motion.

If we study the motion of the piston  $P$  carefully (Figure 3.2.11), we note that when the piston moves towards  $A$ , its velocity  $v$  is from right to left. At  $A$  it comes instantaneously to rest and reverses direction. So before reaching  $A$  the piston must slow down, in other words the acceleration  $a$ , must act in the opposite direction to the velocity. This is also true (in the opposite sense) when the piston reverses its direction and reaches  $B$  at the other end of its stroke. At the times in between these two extremities the piston is being accelerated in the same direction as its velocity.

Neither the velocity nor the acceleration acting on the piston is uniform and so they cannot be modelled using the methods for uniform acceleration. However all is not lost, because although the periodic movement of the piston is complicated, its motion may be modelled using *simple harmonic motion* (s.h.m.).

We may define *s.h.m.* as a special periodic motion in which:

- (i) the acceleration of the body is always directed towards a fixed point in its path;
- (ii) the magnitude of the acceleration of the body is proportional to its distance from the fixed point.

Thus the motion is similar to the motion of the piston described, except that the acceleration has been defined in a particular way.



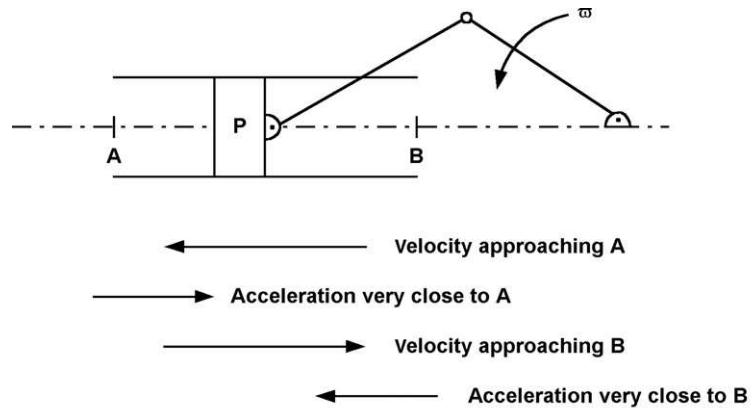


**Figure 3.2.11** Reciprocal motion of spring-mass system

Note: Acceleration opposes motion as piston decelerates close to ends A & B, but is in the same direction as velocity during mid-travel.

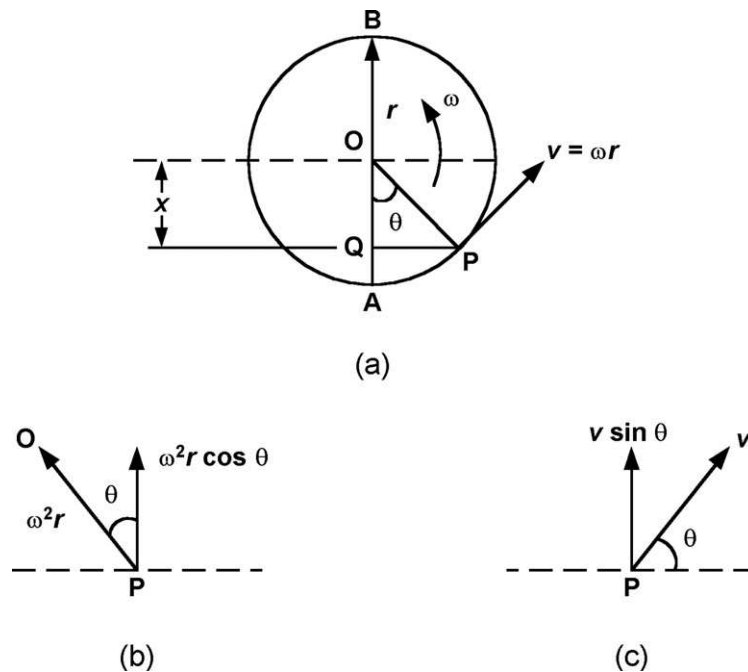


**Figure 3.2.12** Analysis of simple harmonic motion

## Mathematics in action

In order to analyse s.h.m. we need to make use of a technique that involves trigonometrical identities.

Consider the point  $P$  moving round in a circle of radius  $r$ , with angular velocity  $\omega$  (Figure 3.2.12(a)). We know from our study of the equations of motion that its tangential speed  $v = \omega r$ . Imagine that after leaving position  $t$ , the motion is frozen at a moment in time  $t$ , when the point is at position  $P$  as shown. Also note the point  $Q$  which is at the base of the perpendicular from the point  $P$  on the circle diameter

AOB. We will now show that this point Q moves with s.h.m. about the centre O.

The point Q moves as the point P moves, therefore the *acceleration of Q* is the component of the acceleration of P parallel to OB. Now the acceleration of P along PO is given by  $\omega^2 r$ . Then using trigonometry (Figure 3.2.12(b)), the component of this acceleration parallel to AB is  $\omega^2 r \cos \theta$ . So we may write that the acceleration  $a$  of Q is

$$a = -\omega^2 r \cos \theta \quad (3.2.26)$$

The negative sign results from our definition of s.h.m. given above. Where we state that the acceleration is always directed towards a fixed point. In our case towards the point O, which is represented mathematically by use of the negative sign.

Now we also know that  $x = r \cos \theta$

Therefore

$$a = -\omega^2 x \quad (3.2.27)$$

It is important to note how the acceleration of Q varies with different values of  $x$ . For example, the acceleration of Q will be zero, when  $x$  is at the fixed point of rotation O. Also the acceleration will be at a maximum, when  $x$  is at the limits of the oscillation, that is at points A and B.

The time required for the point Q to make one complete oscillation from A to B and back, is known as the *period T*. The time period can be determined using:

$$T = \frac{\text{Circumference of described circle}}{\text{Speed of Q}}$$

$$\text{then } T = \frac{2\pi r}{v} = \frac{2\pi}{\omega} \quad (3.2.28)$$

(since, from our transformation equations  $v = \omega r$ ).

Also note that the frequency  $f$  is the number of complete oscillations (cycles), back and forth, made in unit time. The frequency is therefore the reciprocal of the time period  $T$ . The unit of frequency is the hertz (Hz) which is one cycle per second. Thus:

$$f = \frac{1}{T} \text{ cycles per second} = \frac{\omega}{2\pi} \text{ Hz.} \quad (3.2.29)$$

The *velocity* of Q is the component of P's velocity parallel to AB (Figure 3.2.12c), which is:

$$\text{Velocity of Q} = -v \sin \theta$$

also

$$\text{Velocity of Q} = -\omega r \sin \theta \quad (\text{from } v = \omega r) \quad (3.2.30)$$

Note that  $\sin \theta$  may be positive or negative dependent on the value of  $\theta$ . A negative sign is added to the above equations so that, by convention, when the velocity acts upwards it is negative and when acting downwards it is positive.

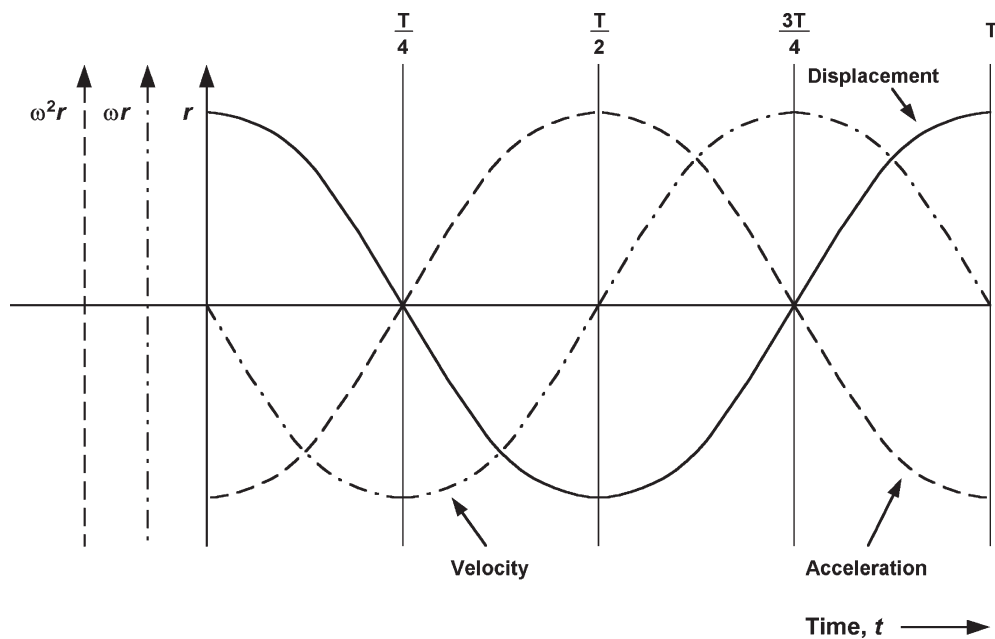


**Figure 3.2.13** Displacement, velocity and acceleration of a point subject to s.h.m.

The variation of the velocity of Q with time  $t$  assuming we start at time zero, is given by:

$$= -\omega r \sin \omega t \text{ (since } \theta = \omega t \text{)}.$$

The variation of the velocity of Q with displacement  $x$

$$= -\omega r \sin \theta$$

$$= \pm \omega r \sqrt{1 - \cos^2 \theta} \quad (\text{since } \sin^2 \theta + \cos^2 \theta = 1)$$

$$= \pm \omega r \sqrt{1 - (x/r)^2}$$

$$= \pm \omega \sqrt{r^2 - x^2} \quad (3.2.31)$$

In order to complete our analysis of the s.h.m. of Q we need to consider its *displacement*, which is given by

$$x = r \cos \theta$$

$$= r \cos \omega t \quad (3.2.32)$$

The displacement, like the velocity and acceleration of Q is sinusoidal (Figure 3.2.13). Note from the diagram that when the velocity is zero the acceleration is a maximum.

If you found the mathematical analysis of s.h.m. difficult then you should refer to Chapter 6.

**Example 3.2.21**

A piston performs reciprocal motion in a straight line which can be modelled as s.h.m. Its velocity is 15 m/s when the displacement is 80 mm from the mid-position, and 3 m/s when the displacement is 160 mm from the mid-position. Determine:

- (i) the frequency and amplitude of the motion;
- (ii) the acceleration when the displacement is 120 mm from the mid-position.

We first need to determine the distance the piston moves on either side of the fixed point, in our case the mid-point. This is known as the *amplitude*.

We use the equation  $V = \omega\sqrt{r^2 - x^2}$  and substitute the two sets of values for the velocity and displacement given in the question, then solve for  $r$ , using simultaneous equations.

From given data we have,  $x = 0.08$  m when  $v = 15$  m/s and  $x = 0.16$  m when  $v = 3$  m/s.

So substituting into above equation gives:

$$15 = \omega\sqrt{r^2 - 0.0064} \quad (3.2.33)$$

$$3 = \omega\sqrt{r^2 - 0.0256} \quad (3.2.34)$$

Dividing the top equation by the bottom eliminates  $\omega$  and squaring both sides gives:

$$25 = \frac{r^2 - 0.0064}{r^2 - 0.0256}$$

and so  $r = 162.5$  mm.

In order to determine the frequency we first need to find the angular velocity from Equation 3.2.33,

$$\text{then } 15 = \omega\sqrt{0.0264 - 0.0064}$$

$$\text{so } \omega = 106 \text{ rad/s}$$

and using equation  $f = \frac{\omega}{2\pi} = \frac{106}{2\pi}$  then frequency = 16.9 Hz.

To find the acceleration when  $x = 0.12$  m we use Equation 3.2.27, then

$$a = \omega^2 x$$

$$a = (106)^2(0.12)$$

$$a = 1348.3 \text{ m/s}^2.$$

**Motion of a spring**

You will be aware that for a spring which obeys Hooke's law, the extension of the spring is directly proportional to the force applied to it. So it follows that the extension of the spring is directly proportional to the tension in the spring resulting from the applied force.

Consider the spring with the mass ( $m$ ) attached to it hanging freely as shown in Figure 3.2.14. The mass exerts a downward tension ( $mg$ ) on the

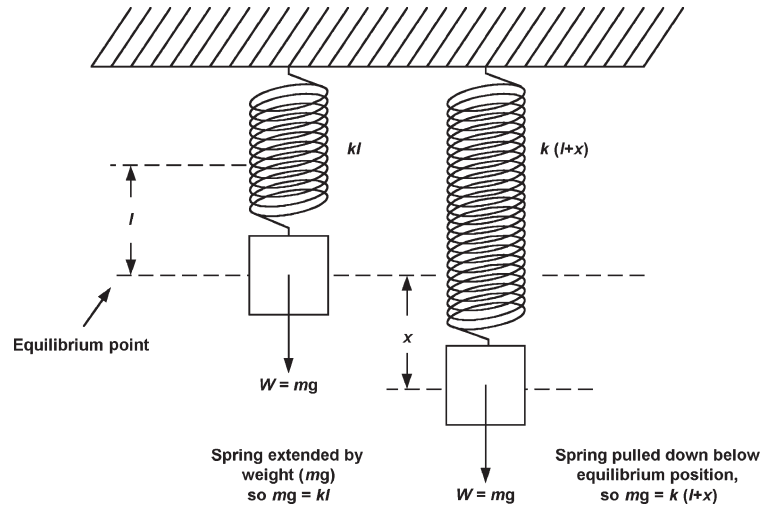


**Figure 3.2.14** Reciprocal motion of spring-mass system

spring, which stretches the spring by an amount ( $l$ ), until the internal resistance of the spring ( $kl$ ), balances the tension created by the mass. Then:

$$mg = kl \quad (3.2.35)$$

where  $k$  = the spring constant or force required to stretch the spring by unit length (N/m). The spring constant is an inherent property of the spring and depends on its material properties and dimensions.  $l$  = the length the spring stretches in reaching equilibrium (m),  $m$  = the mass (kg) and  $g$  = acceleration due to gravity ( $\text{m/s}^2$ ). Note that both sides of Equation 3.2.35 have units of force.

Suppose the mass is now pulled down a further distance  $x$  below its equilibrium position, the stretching tension acting downwards is now  $k(l+x)$  which is also the tension in the spring acting upwards. So we have

$$mg = k(l+x)$$

and so the resultant restoring force upwards acting on the mass will be

$$\begin{aligned} &= k(l+x) - mg \\ &= kl + kx - kl \quad (kl = mg \text{ from 3.2.35}) \\ &= kx. \end{aligned}$$

When we release the mass it will oscillate up and down. If it has acceleration ( $a$ ) at extension ( $x$ ) then by Newton's second law

$$-kx = ma \quad (3.2.36)$$

The negative sign results from the assumption that at the instant shown the acceleration acts upwards, or in a negative direction (using our sign convention). At the same time the displacement ( $x$ ) acts in the opposite direction, that is downwards and positive.

Now transposing Equation 3.2.36 for  $a$  gives:

$$a = -\frac{kx}{m}$$

also from Equation 3.2.27,  $a = -\omega^2 x$  and so  $k/m = \omega^2$  and since  $k$  and  $m$  are fixed for any system  $k/m$  is always positive. Thus the motion of the

spring is a simple harmonic about the equilibrium point, providing the spring system obeys Hooke's law.

Now since the motion is simple harmonic, the time period for the spring system is given by:

$$T = \frac{2\pi}{\omega} \quad \text{and from above } \sqrt{\frac{k}{m}} = \omega$$

$$\text{so } T = 2\pi\sqrt{\frac{m}{k}} \quad (3.2.37)$$

### Example 3.2.22

A spiral spring is loaded with a mass of 5 kg which extends it by 100 mm. Calculate the period of vertical oscillations.

The time period of the oscillations is given by Equation 3.2.37.

$$T = 2\pi\sqrt{\frac{m}{k}}$$

Now  $k$  = force per unit displacement and is given from the information in the question as:

$$k = \frac{5 \times 9.81 \text{ N}}{0.1 \text{ m}} = 490.5 \text{ N/m}$$

$$\text{then } T = 2\pi\sqrt{\frac{5}{490.5}} = 0.63 \text{ s}$$

### Motion of a simple pendulum

The simple pendulum consists of a small bob of mass  $m$  (which is assumed to act as a particle). A light inextensible cord of length  $l$  to which the bob is attached, is suspended from a fixed point O. Figure 3.2.15 illustrates the situation when the pendulum is drawn aside and

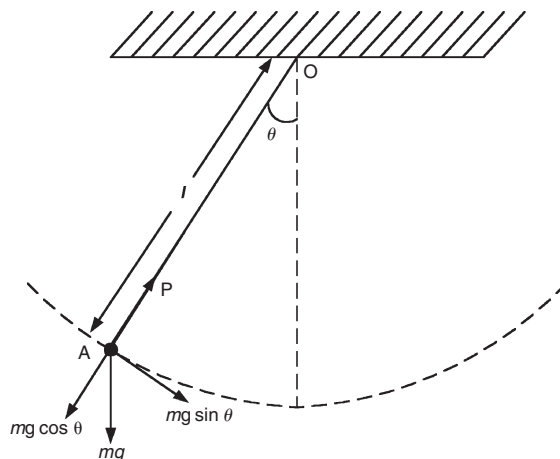


**Figure 3.2.15** Motion of simple pendulum

oscillates freely in the vertical plane along the arc of a circle, shown by the dotted line.

It can be shown that the pendulum for small angles  $\theta$  (rads) describes s.h.m. Figure 3.2.15 shows the forces resulting from the weight of the bob, together with its radial and tangential components. Note that the tangential component  $mg \sin \theta$  is the unbalanced restoring force acting towards the centre, while the radial component  $mg \cos \theta$  balances the force in the cord P. Then if  $a$  is the acceleration of the bob along the arc at A due to the unbalanced restoring force, the equation of motion of the bob is given by:

$$-mg \sin \theta = ma \quad (3.2.38)$$

Again using our sign convention the restoring force is towards Q and is, therefore, considered to be negative, while the displacement  $x$  is measured from point Q and is, therefore, positive.

Now when  $\theta$  is small  $\sin \theta = \theta$  (radi) using this fact it can be shown that:

$$a = \frac{g}{l} x = -\omega^2 x \quad (\text{where } \omega^2 = g/l) \quad (3.2.39)$$

which is s.h.m. and so the time period  $T$  for the motion is given by:

$$T = \frac{2\pi}{\omega} = 2\pi \sqrt{\frac{l}{g}} \quad (3.2.40)$$

### Example 3.2.23

A simple pendulum has a bob attached to an inextensible cord 50 cm long. Determine its frequency.

Knowing that the time period is equal to the reciprocal of the frequency, all that is needed is to find the time period  $T$  and then the frequency.

$$\begin{aligned} \text{Then } T &= 2\pi \sqrt{\frac{l}{g}} \\ &= 2\pi \sqrt{\frac{0.5}{9.81}} \\ &= 1.42 \text{ s.} \end{aligned}$$

$$\begin{aligned} \text{So the frequency in Hz is } &= \frac{1}{1.42} \\ &= \mathbf{0.704 \text{ Hz.}} \end{aligned}$$

### Damping, forcing and resonance

If we observe a simple pendulum which is allowed to swing freely, we note that after a time the amplitude of the oscillations of the bob decrease to zero. Its motion, is therefore, not perfect s.h.m., it has been acted upon

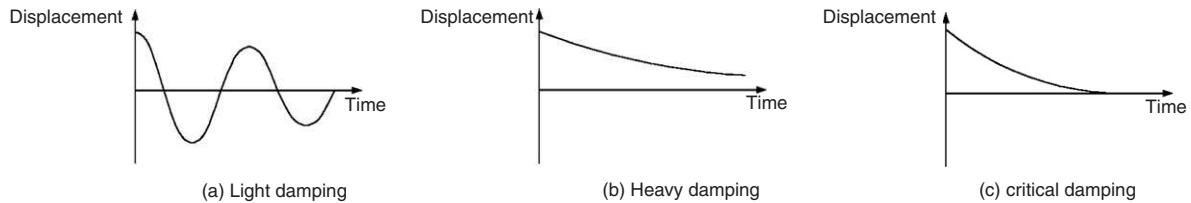


**Figure 3.2.16** Effects of damping on a body subject to oscillatory motion

by the air which offers a resistance to its motion, we say that the pendulum has been *damped* by air resistance.

The rate at which our pendulum, or any body subject to oscillatory motion is brought to rest depends on the *degree of damping*. For example, in the case of our pendulum the air provides *light damping*, because the number of oscillations that occur before the displacement of the motion is reduced to zero, is large (Figure 3.2.16a). Similarly, a body subject to heavy damping has its displacement reduced quickly (Figure 3.2.16b). *Critical damping* occurs when the time taken for the displacement to become zero is a minimum (Figure 3.2.16c, d).

Engineering examples of damped oscillatory motion are numerous. For example, on traditional motor vehicle suspension systems the motion of the springs are damped using oil shock absorbers. These prevent the onset of vibrations which are likely to make the control and handling of the vehicle difficult.

The design engineer needs to control the vibration when it is undesirable and to enhance the vibration when it is useful. So large machines may be placed on anti-vibration mountings to reduce unwanted vibration which, if left unchecked, would cause the loosening of parts leading to possible malfunction or failure. Delicate instrument systems in aircraft are insulated from the vibrations set up by the aircraft engines and atmospheric conditions, again by placing anti-vibration mountings between the instrument assembly and the aircraft structure. Occasionally vibration is considered useful, for example, shakers in foundries and vibrators for testing machines.

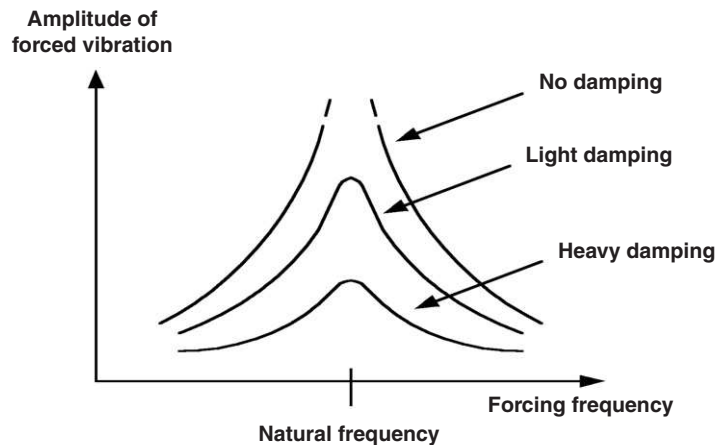
If we again return to the example of our simple pendulum the frequency of vibration is dependent on the length of the cord, so that each pendulum of given length, will vibrate at its *natural frequency*, when allowed to swing *freely*. If a body is subject to an external vibration force then it will oscillate at the frequency of this external vibration and we call this *forced vibration*, our foundry shaker subjects the melt to forced vibration to assist mixing and settlement.

When the frequency of the external driving force is equal to that of the natural frequency of the body or system, then we say *resonance* occurs. This is where the energy from the external driving force is most easily transferred to the body or system. Since no energy is required to maintain the vibrations of an undamped system at its natural frequency, then all the energy transferred from the driving force will be used to build up the amplitude of vibration which, at resonance, will increase without limit. Hence the need to ensure such a system is adequately damped, Figure 3.2.17 illustrates the relationship between amplitude, frequency and damping at resonance.

Resonance is generally considered troublesome, especially in mechanical systems. The classic example used to illustrate the effects of



**Figure 3.2.17** Relationship between amplitude, frequency and damping at resonance



unwanted resonance is the Tacoma Bridge disaster in America in 1940. The prevailing wind caused an oscillating force in resonance with the natural frequency of part of the bridge. Oscillations built up which were so large that they destroyed the bridge.

The rudder of an aircraft is subject to forced vibrations that result from the aircraft engines, flight turbulence and aerodynamic loads imposed on the aircraft. In order to prevent damage from possible resonant frequencies, the rudder often has some form of hydraulic or viscous damping fitted. The loss of a flying control resulting from resonant vibration could result in an aircraft accident, with subsequent loss of life.

So when might resonant frequencies be useful? Electrical resonance occurs when a radio circuit is tuned by making its natural frequency of oscillations equal to that of the incoming radio systems.

### Problems 3.2.3

- (1) A particle moves with simple harmonic motion. Given that its acceleration is  $12 \text{ m/s}^2$  when  $8 \text{ cm}$  from the mid-position. Find the time period for the motion. Also, if the amplitude of the motion is  $11 \text{ cm}$  find the velocity when the particle is  $8 \text{ cm}$  from the mid-position.
- (2) Write a short account of simple harmonic motion, explaining the terms amplitude, time period and frequency.
- (3) A simple pendulum is oscillating with amplitude of  $30 \text{ mm}$ . If the pendulum is  $80 \text{ cm}$  long, determine the velocity of the bob as it passes through the mid-point of its oscillation.
- (4) A body of mass  $2 \text{ kg}$  hangs from a spiral spring. When the mass is pulled down  $10 \text{ cm}$  and released, it vibrates with s.h.m. If the time period for this motion is  $0.4 \text{ s}$ , find the velocity when it passes through the mid-position of its oscillation and the acceleration when it is  $4 \text{ cm}$  from the same mid-position.
- (5) An engine consists of a simple crank and connecting rod mechanism, in which the crank is  $7.5 \text{ cm}$  long and the connecting rod is  $45 \text{ cm}$  long. When the crank is  $40^\circ$  from top dead centre, determine the velocity and acceleration of the piston when the crank rotates at  $360 \text{ rpm}$ .

### 3.3 DC AND AC THEORY

If you have not studied AC theory before, this introductory section has been designed to quickly get you up to speed. If, on the other hand, you have previously studied electrical and electronic principles at BTEC level N (or an equivalent GNVQ advanced level unit) you should move on to 'Using complex notation' (see page 222).

## DC electrical principles

### **Current, voltage, and resistance**

Current is simply the rate of flow of electric charge. Thus, if more charge moves in a given time, more current will be flowing. If no charge moves then no current is flowing. The unit of current is the ampere, A, and 1 ampere is equal to 1 Coulomb of charge flowing past a point in 1 s. Hence:

$$I = \frac{Q}{t}$$

where  $t$  = time in seconds and  $Q$  = charge in Coulombs.

So, for example, if a steady current of 3 A flows for 2 min, then the amount of charge transferred will be:

$$Q = I \times t = 3 \text{ A} \times 120 \text{ s} = 360 \text{ Coulombs.}$$

Alternatively, if 2400 Coulombs of charge are transferred in 1 min, the current flowing will be given by:

$$I = \frac{Q}{t} = \frac{2400 \text{ Coulombs}}{60 \text{ s}} = 40 \text{ A.}$$

Electric current arises from the flow of electrons (or negative charge carriers) in a metallic conductor. The ability of an energy source (e.g. a battery) to produce a current within a conductor may be expressed in terms of electromotive force (e.m.f.). Whenever an e.m.f. is applied to a circuit a potential difference (p.d.) exists. Both e.m.f. and p.d. are measured in volts (V). In many practical circuits there is *only one* e.m.f. present (the battery or supply) whereas a p.d. will be developed across *each* component present in the circuit.

The *conventional flow* of current in a circuit is from the point of more positive potential to the point of greatest negative potential (note that electrons move in the opposite direction!). *Direct current* result from the application of a direct e.m.f. (derived from batteries or a DC supply, such as a generator). An essential characteristic of such supplies is that the applied e.m.f. does not change its polarity (even though its value might be subject to some fluctuation).

For any conductor, the current flowing is directly proportional to the e.m.f. applied. The current flowing will also be dependent on the physical dimensions (length and cross-sectional area) and material of which the conductor is composed. The amount of current that will flow in a conductor when a given e.m.f. is applied is inversely proportional to its resistance. Resistance, therefore, may be thought of as an 'opposition to current

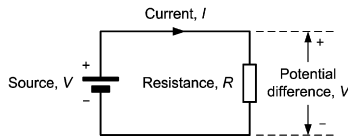


**Figure 3.3.1** Voltage, current, and resistance

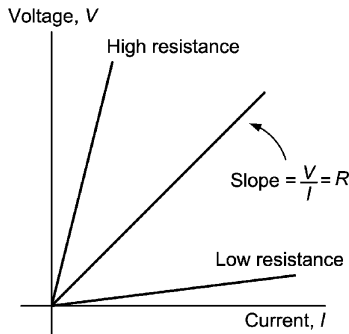


**Figure 3.3.2** Voltage plotted against resistance

flow'; the higher the resistance the lower the current that will flow (assuming that the applied e.m.f. remains constant).

### Ohm's law

Provided that temperature does not vary, the ratio of p.d. across the ends of a conductor to the current flowing in the conductor is a constant. This relationship is known as Ohm's law and it leads to the relationship:

$$\frac{V}{I} = \text{a constant} = R$$

where  $V$  is the p.d. (or voltage drop) in volts (V),  $I$  is the current in A, and  $R$  is the resistance in ohms ( $\Omega$ ) (see Figure 3.3.1). The relationship between  $V$  and  $I$  is linear, as shown in Figure 3.3.2. The slope of this relationship (when  $V$  is plotted against  $I$ ) gives the resistance. A steep slope corresponds to a high resistance whilst a shallow slope corresponds to a low resistance.

The relationship between  $V$  and  $I$  is linear, as shown in Figure 3.3.2. The slope of this relationship (when  $V$  is plotted against  $I$ ) gives the resistance. A steep slope corresponds to a high resistance whilst a shallow slope corresponds to a low resistance.

The relationship between  $V$ ,  $I$ , and  $R$  may be arranged to make  $V$ ,  $I$ , or  $R$  the subject, as follows:

$$V = I \times R \quad I = \frac{V}{R} \quad \text{and} \quad R = \frac{V}{I}.$$

#### Example 3.3.1

A current of 0.1 A flows in a  $56 \Omega$  resistor. What voltage drop will be developed across the resistor?

Here we must use  $V = I \times R$  (where  $I = 0.1 \text{ A}$  and  $R = 56 \Omega$ ). Hence:

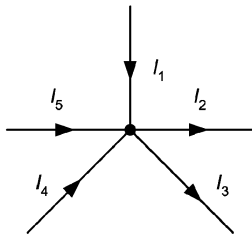
$$V = I \times R = 0.1 \times 56 = 5.6 \text{ V}.$$

#### Example 3.3.2

An  $18 \Omega$  resistor is connected to a 9 V battery. What current will flow in the resistor?

Here we must use  $I = \frac{V}{R}$  (where  $V = 9 \text{ V}$  and  $R = 18 \Omega$ ). Hence:

$$I = \frac{V}{R} = \frac{9}{18} = 0.5 \text{ A}.$$



$$I_1 - I_2 - I_3 + I_4 + I_5 = 0$$

Convention  
Current flowing towards the junction is positive (+)  
Current flowing away from the junction is negative (-)

**Figure 3.3.3** Kirchhoff's current law

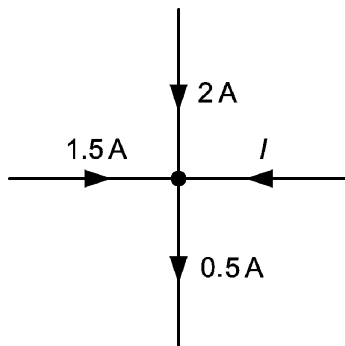
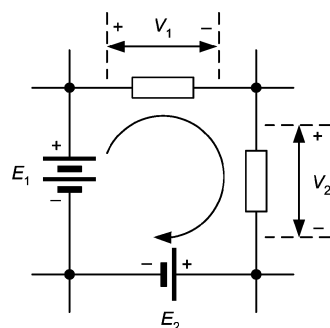


**Figure 3.3.4** See example 3.3.4



$$E_1 - V_1 - V_2 - E_2 = 0$$

Convention  
Move clockwise around the circuit starting with the positive terminal of the largest e.m.f.  
Voltages acting in the same sense are positive (+)  
Voltages acting in the opposite sense are negative (-)

**Figure 3.3.5** Kirchhoff's voltage law

### Example 3.3.3

A voltage drop of 15 V appears across a resistor in which a current of 1 mA flows. What is the value of the resistance?

Here we must use  $R = \frac{V}{I}$  where  $V = 15 \text{ V}$  and  $I = 1 \text{ mA} = 1 \times 10^{-3} \text{ A}$ . Hence:

$$R = \frac{V}{I} = \frac{15}{1 \times 10^{-3}} = 15 \times 10^3 \Omega = 15 \text{ k}\Omega.$$

### Kirchhoff's current law

Used on its own, Ohm's law is insufficient to determine the magnitude of the voltages and currents present in complex circuits. For these circuits we need to make use of two further laws: Kirchhoff's current law and Kirchhoff's voltage law.

*Kirchhoff's current law* states that the algebraic sum of the currents present at a junction (or *node*) in a circuit is zero (see Figure 3.3.3).

### Example 3.3.4

Determine the value of the missing current,  $I$ , shown in Figure 3.3.4.

By applying Kirchhoff's current law in Figure 3.3.4, and adopting the convention that currents flowing towards the junction are positive, we can say that:

$$+2 \text{ A} + 1.5 \text{ A} - 0.5 \text{ A} + I = 0.$$

Note that we have shown  $I$  as positive. In other words we have assumed that it is flowing towards the junction.

Re-arranging gives:

$$+3 \text{ A} + I = 0.$$

$$\text{thus } I = -3 \text{ A}.$$

The negative answer tells us that  $I$  is actually flowing in the other direction, that is away from the junction.

### Kirchhoff's voltage law

Kirchhoff's second, voltage law states that the algebraic sum of the potential drops present in a closed network (or *mesh*) is zero (see Figure 3.3.5).

**Example 3.3.5**

Determine the value of the missing voltage,  $V$ , shown in Figure 3.3.6.

By applying Kirchhoff's voltage law in Figure 3.3.6, starting at the positive terminal of the largest e.m.f. and moving clockwise around the closed network, we can say that:

$$+24\text{ V} - V + 6\text{ V} - 12\text{ V} = 0.$$

Note that we have shown  $V$  as positive. In other words we have assumed that the more positive terminal of the resistor is the one on the left.

Re-arranging gives:

$$+24\text{ V} - V + 6\text{ V} - 12\text{ V} = 0.$$

From which:

$$+18\text{ V} - V = 0$$

thus  $V = +18\text{ V}$ .

The positive answer tells us that we have made a correct assumption concerning the polarity of the voltage drop,  $V$ , that is the more positive terminal is on the left.

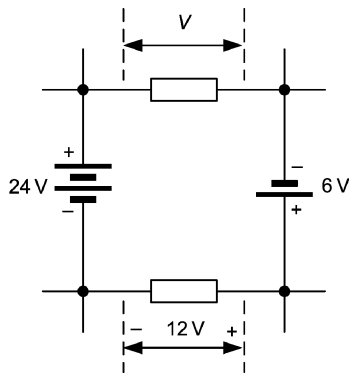


**Figure 3.3.6** See Example 3.3.5

**Series and parallel circuit calculations**

Ohm's law and Kirchhoff's laws can be combined to solve more complex series-parallel circuits. Before we show you how this is done, however, it's important to understand what we mean by 'series' and 'parallel' circuit!

Figure 3.3.7 shows three circuits, each containing three resistors,  $R_1$ ,  $R_2$ , and  $R_3$ .

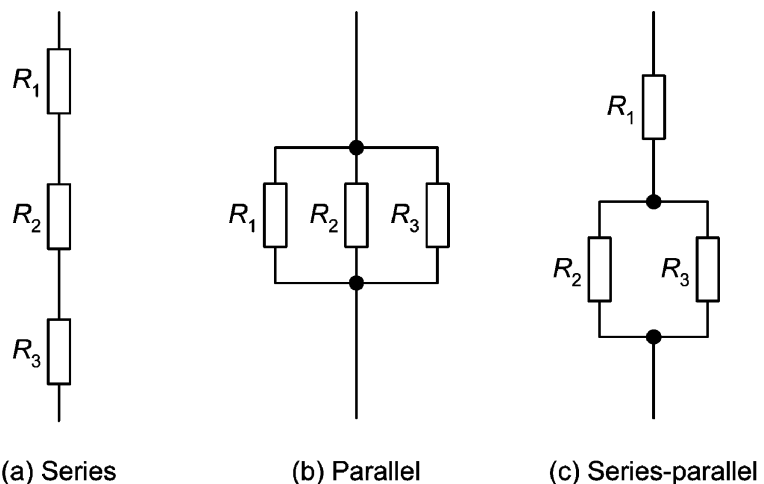


**Figure 3.3.7** Series and parallel circuits

In Figure 3.3.7(a), the three resistors are connected one after another. We refer to this as a *series circuit*. In other words the resistors are said to be connected *in series*. Its important to note that, in this arrangement, *the same current flows through each resistor*.

In Figure 3.3.7(b), the three resistors are all connected across one another. We refer to this as a *parallel circuit*. In other words the resistors are said to be connected *in parallel*. Its important to note that, in this arrangement, *the same voltage appears each resistor*.

In Figure 3.3.7(c), we have shown a mixture of these two types of connection. Here we can say that  $R_1$  is connected in series with the parallel combination of  $R_2$  and  $R_3$ . In other words,  $R_2$  and  $R_3$  are connected *in parallel* and  $R_2$  is connected *in series* with the parallel combination.

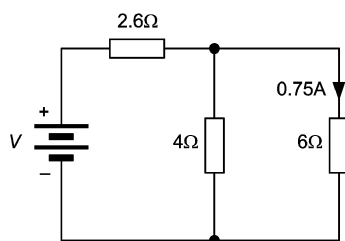


**Figure 3.3.8** See Example 3.3.6

### Example 3.3.6

For the circuit shown in Figure 3.3.8, determine:

- the voltage dropped across each resistor
- the current drawn from the supply
- the supply voltage.

We need to solve this problem in several small stages. Since we know the current flowing in the  $6\Omega$  resistor, we will start by finding the voltage dropped across it using Ohm's law:

$$V = I \times R = 0.75 \times 6 = 4.5 \text{ V.}$$

Now the  $4\Omega$  resistor is connected in parallel with the  $6\Omega$  resistor. Hence the voltage drop across the  $4\Omega$  resistor is also  $4.5 \text{ V}$ . We can now determine the current flowing in the  $4\Omega$  resistor using Ohm's law:

$$I = \frac{V}{R} = \frac{4.5}{4} = 1.125 \text{ A.}$$

Since we know now the current in both the  $4\Omega$  and  $6\Omega$  resistors we can use Kirchhoff's law to find the current,  $I$ , in the  $2.6\Omega$  resistor:

$$+I - 0.75 \text{ A} - 1.125 \text{ A} = 0.$$

From which:

$$I = 1.875 \text{ A.}$$

Since this current flows through the  $2.6\Omega$  resistor it will also be equal to the current taken from the supply.

Next we can find the voltage drop across the  $2.6\Omega$  resistor by applying Ohm's law:

$$V = I \times R = 1.875 \times 2.6 = 4.875 \text{ V.}$$

Finally, we can apply Kirchhoff's voltage law in order to determine the supply voltage,  $V$ :

$$+V - 4.875 - 4.5 = 0.$$

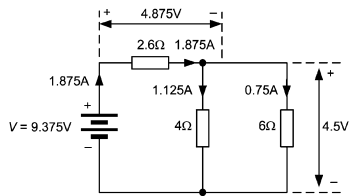
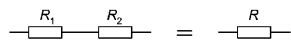
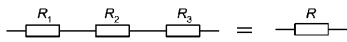


**Figure 3.3.9** See Example 3.3.6

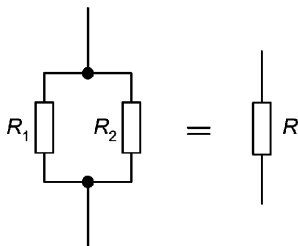


(a) Two resistors in series

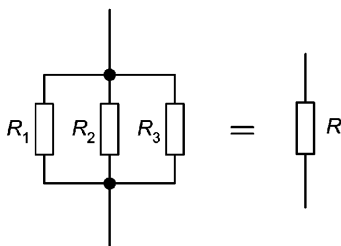


(b) Three resistors in series

**Figure 3.3.10** Series combination of resistors



(a) Two resistors in parallel



(b) Three resistors in parallel

**Figure 3.3.11** Parallel combination of resistors

From which:

$$V = +9.375 \text{ V.}$$

Hence the supply voltage is **9.375 V** (as shown in Figure 3.3.9).

### Series and parallel combinations of resistors

In order to obtain a particular value of resistance, fixed resistors may be arranged in either series or parallel as shown in Figures 3.3.10 and 3.3.11.

The equivalent resistance of each of the series circuits shown in Figure 3.3.10 is simply equal to the sum of the individual resistances.

Hence, for Figure 3.3.10(a):

$$R = R_1 + R_2$$

whilst for Figure 3.3.10(b):

$$R = R_1 + R_2 + R_3.$$

Turning to the parallel resistors shown in Figure 3.3.11, the reciprocal of the equivalent resistance of each circuit is equal to the sum of the reciprocals of the individual resistances. Hence, for Figure 3.3.11(a):

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

whilst for Figure 3.3.11(b):

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}.$$

In the former case (*for just two resistors connected in parallel*) the equation can be more conveniently re-arranged as follows:

$$R = \frac{R_1 \times R_2}{R_1 + R_2}.$$

Thus the equivalent resistance of two resistors connected in parallel can be found by taking the *product* of the two resistance values and *dividing* it by the *sum* of the two resistance values (in other words, *product over sum*).

### Example 3.3.7

Resistors of  $22 \Omega$ ,  $47 \Omega$  and  $33 \Omega$  are connected: (a) in series and (b) in parallel. Determine the effective resistance in each case.

(a) In the series circuit:  $R = R_1 + R_2 + R_3$   
 thus  $R = 22 \Omega + 47 \Omega + 33 \Omega = \mathbf{102 \Omega}.$

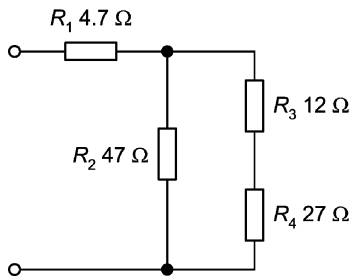


**Figure 3.3.12** See Example 3.3.8

(b) In the parallel circuit:

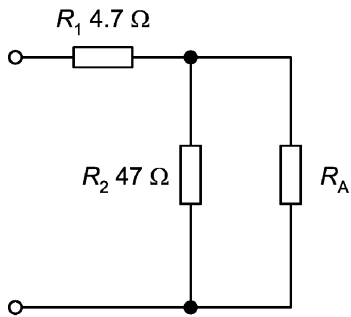
$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

thus

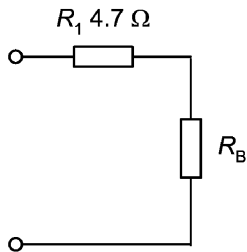
$$\frac{1}{R} = \frac{1}{22} + \frac{1}{47} + \frac{1}{33}$$

$$\text{or } \frac{1}{R} = 0.045 + 0.021 + 0.03 = 0.096$$

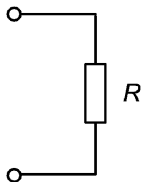
thus  $R = 10.42 \Omega$ .



(a)  $R_A$  replaces the series combination of  $R_3$  and  $R_4$



(b)  $R_B$  replaces the parallel combination of  $R_a$  and  $R_B$



(c)  $R$  replaces the series combination of  $R_1$  and  $R_B$

**Figure 3.3.13** See Example 3.3.8

### Example 3.3.8

Determine the effective resistance of the circuit shown in Figure 3.3.12.

The circuit can be progressively simplified as shown in Figure 3.3.13.

The stages in this simplification are:

- (a)  $R_3$  and  $R_4$  are in series and they can be replaced by a single resistance ( $R_A$ ) of  $12 + 27 = 39 \Omega$ .
- (b)  $R_A$  appears in parallel with  $R_2$ . These two resistors can be replaced by a single resistance ( $R_B$ ) of:

$$\frac{39 \times 47}{39 + 47} = 21.3 \Omega.$$

- (c)  $R_B$  appears in series with  $R_1$ . These two resistors can be replaced by a single resistance,  $R$ , of  $21.3 + 4.7 = 26 \Omega$ .

### The potential divider

The potential divider circuit (see Figure 3.3.14) is commonly used to reduce voltage levels in a circuit. The output voltage produced by the circuit is given by:

$$V_{\text{out}} = V_{\text{in}} \times \frac{R_2}{R_1 + R_2}.$$

It is, however, important to note that the output voltage ( $V_{\text{out}}$ ) will fall when current is drawn away from the arrangement.

### Example 3.3.9

Determine the output voltage of the circuit shown in Figure 3.3.15.

Here we can use the potential divider formula,

$$V_{\text{out}} = V_{\text{in}} \times \frac{R_2}{R_1 + R_2}, \quad \text{where } V_{\text{in}} = 5 \text{ V}$$



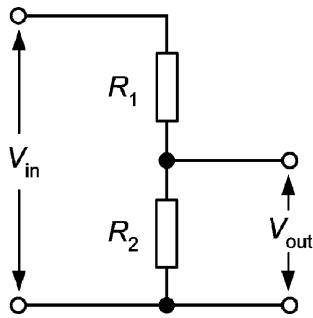


**Figure 3.3.14** The potential divider

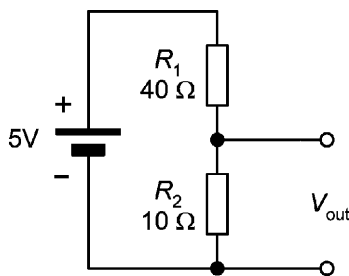


**Figure 3.3.15** See Example 3.3.9

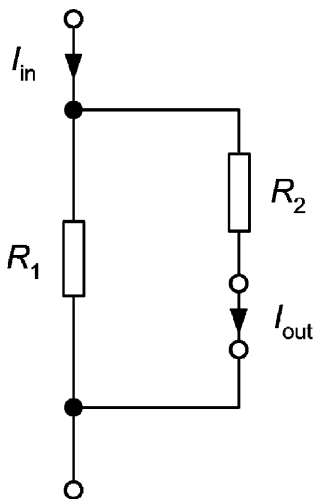


**Figure 3.3.16** The current divider

$R_1 = 40\ \Omega$  and  $R_2 = 10\ \Omega$ , thus:

$$V_{\text{out}} = V_{\text{in}} = \frac{R_2}{R_1 + R_2} = 5 \times \frac{10}{10 + 40} = 5 \times \frac{1}{5} = 1\text{ V}.$$

### The current divider

The current divider circuit (see Figure 3.3.16) is used to divert current from one branch of a circuit to another. The output current produced by the circuit is given by:

$$I_{\text{out}} = I_{\text{in}} \times \frac{R_1}{R_1 + R_2}.$$

It is important to note that the output current ( $I_{\text{out}}$ ) will fall when the load connected to the output terminals has any appreciable resistance.

### Example 3.3.10

A moving coil meter requires a current of 1 mA to provide full-scale deflection. If the meter coil has a resistance of  $100\ \Omega$  and is to be used as a milliammeter reading 5 mA full-scale, determine the value of parallel 'shunt' resistor required.

This problem may sound a little complicated so it is worth taking a look at the equivalent circuit of the meter (Figure 3.3.17) and comparing it with the current divider shown in Figure 3.3.16.

We can apply the current divider formula, replacing  $I_{\text{out}}$  with  $I_m$  (the meter full-scale deflection current) and  $R_2$  with  $R_m$  (the meter resistance).  $R_1$  is the required value of shunt resistor,  $R_s$ .

From  $I_{\text{out}} = I_{\text{in}} \times \frac{R_1}{R_1 + R_2}$  we can say that  $I_m = I_{\text{in}} \times \frac{R_s}{R_s + R_m}$  where  $I_m = 1\text{ mA}$ ,  $I_{\text{in}} = 5\text{ mA}$  and  $R_2 = 100\ \Omega$ .

Re-arranging the formula gives:

$$I_m (R_s + R_m) = I_{\text{in}} \times R_s$$

$$\text{thus } I_m R_s + I_m R_m = I_{\text{in}} \times R_s$$

$$\text{or } I_{\text{in}} \times R_s - I_m R_s = I_m R_m$$

$$\text{hence } R_s (I_{\text{in}} - I_m) = I_m R_m$$

$$\text{and } R_s = \frac{I_m R_m}{I_{\text{in}} - I_m}$$

Now  $I_m = 1\text{ mA}$ ,  $R_m = 100\ \Omega$ , and  $I_{\text{in}} = 5\text{ mA}$   
thus

$$R_s = \frac{I_m R_m}{I_{\text{in}} - I_m} = \frac{1 \times 10^{-3} \times 100}{5 \times 10^{-3} - 1 \times 10^{-3}} = \frac{100}{4} = 25\ \Omega.$$

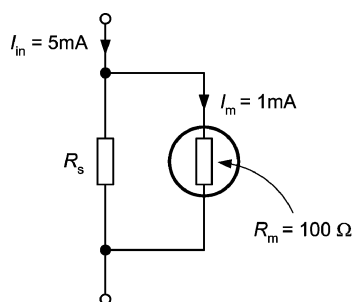


**Figure 3.3.17** Meter circuit

## Power, work, and energy

From your study of dynamic engineering systems you will recall that energy can exist in many forms including kinetic energy, potential energy, heat energy, light energy, etc. Kinetic energy is concerned with the movement of a body whilst potential energy is the energy that a body possesses due to its position. Energy can be defined as ‘the ability to do work’ whilst power can be defined as ‘the rate at which work is done’.

In electrical circuits, energy is supplied by batteries or generators. It may also be stored in components such as capacitors and inductors. Electrical energy is converted into various other forms of energy by components such as resistors (producing heat), loudspeakers (producing sound energy), light emitting diodes (producing light).

The unit of energy is the joule (J). Power is the rate of use of energy and it is measured in watts (W). A power of 1 W results from energy being used at the rate of 1 J/s. Thus:

$$P = \frac{W}{t}$$

where  $P$  is the power in watts (W),  $W$  is the energy in joules (J), and  $t$  is the time in seconds (s).

We can re-arrange the previous formula to make  $W$  the subject, as follows:

$$W = P \times t$$

The power in a circuit is equivalent to the product of voltage and current. Hence:

$$P = I \times V$$

where  $P$  is the power (in W),  $I$  is the current (in A), and  $V$  is the voltage (in V).

The formula may be arranged to make  $P$ ,  $I$ , or  $V$  the subject, as follows:

$$P = I \times V \quad I = \frac{P}{V} \quad \text{and} \quad V = \frac{P}{I}$$

When a resistor gets hot it is dissipating power. In effect, a resistor is a device that converts electrical energy into heat energy. The amount of power dissipated in a resistor depends on the current flowing in the resistor. The more current flowing in the resistor the more power will be dissipated and the more electrical energy will be converted into heat.

It is important to note that the relationship between the current applied and the power dissipated is not linear – in fact it obeys a *square law*. In other words, the power dissipated in a resistor is proportional to the square of the applied current.

The relationship,  $P = I \times V$ , may be combined with that which results from Ohm’s law (i.e.  $V = I \times R$ ), to produce two further relationships. Firstly, substituting for  $V$  gives:

$$P = I \times (I \times R) = I^2 R$$

Secondly, substituting for  $I$  gives:

$$P = \left( \frac{V}{R} \right) \times V = \frac{V^2}{R}$$

**Example 3.3.11**

A current of 1.5 A is drawn from a 3 V battery. What power is supplied?

Here we must use  $P = I \times V$  where  $I = 1.5 \text{ A}$  and  $V = 3 \text{ V}$ :

$$P = I \times V = 1.5 \times 3 = \mathbf{4.5 \text{ W}}.$$

**Example 3.3.12**

A voltage drop of 4 V appears across a resistor of  $100 \Omega$ . What power is dissipated in the resistor?

Here we must use  $P = \frac{V^2}{R}$  (where  $V = 4 \text{ V}$  and  $R = 100 \Omega$ ):

$$P = \frac{V^2}{R} = \frac{4^2}{100} = \frac{16}{100} = \mathbf{0.16 \text{ W}} \text{ (or } 160 \text{ mW)}.$$

**Example 3.3.13**

A current of 200 mA flows in a  $1 \text{ k}\Omega$  resistor. What power is dissipated in the resistor and what energy is used if the current flows for 10 min?

Here we must use  $P = I^2 R$  (where  $I = 200 \text{ mA}$  and  $R = 1000 \Omega$ ):

$$P = I^2 R = 0.2^2 \times 1000 = 0.04 \times 1000 = 40 \text{ W}.$$

To find the energy we need to use  $W = P \times t$  (where  $P = 40 \text{ W}$  and  $t = 10 \text{ min}$ ):

$$W = P \times t = 40 \times (10 \times 60) = 24\,000 \text{ J} = \mathbf{24 \text{ kJ}}.$$

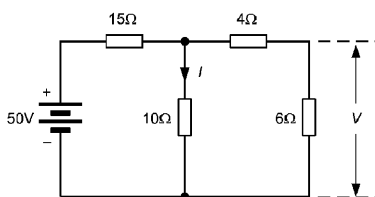


**Figure 3.3.18** See Problems 3.3.1

**Problems 3.3.1**

1. A charge of 60 mC is transferred between two points in AC circuit in a time interval of 15 ms. What current will be flowing?
2. A current of 27 mA flows in a resistance of  $3.3 \text{ k}\Omega$ . What voltage drop will appear across the resistor?
3. Determine the current,  $I$ , and voltage,  $V$ , in Figure 3.3.18.
4. Determine the equivalent resistance of each of the networks of resistors shown in Figure 3.3.19.
5. A resistor of  $22 \Omega$  carries a current of 0.5 A. What power is dissipated?
6. A 12 V power supply delivers a current of 70 mA for 6 h. What energy is delivered?

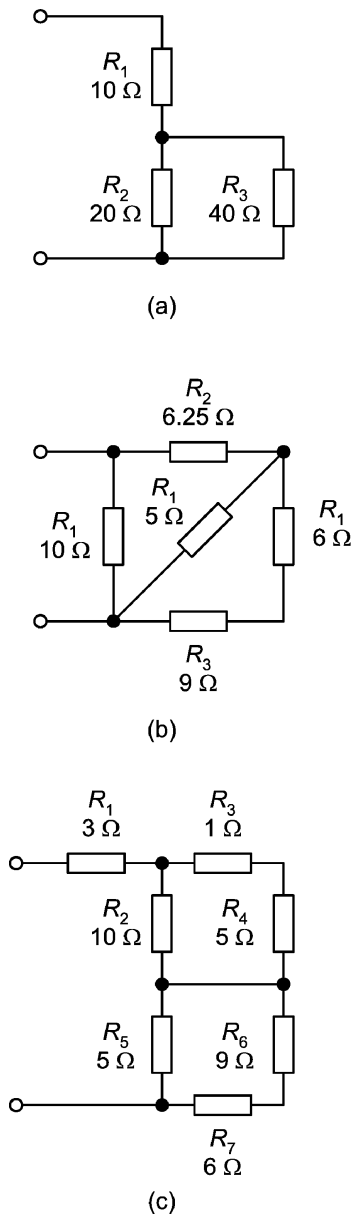


**Figure 3.3.19** See Problems 3.3.1

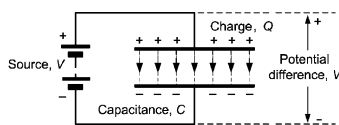


**Figure 3.3.20** Electric field between two charged parallel metal plates

### Charge, voltage, and capacitance

Figure 3.3.20 shows the electric field that exists between two charged metal plates. This arrangement forms a simple *capacitor*. The quantity of electricity,  $Q$ , that can be stored in the electric field between the capacitor plates is proportional to the applied voltage,  $V$ , and the capacitance,  $C$ , of the capacitor. Thus:

$$Q = CV$$

where  $Q$  is the charge (in Coulombs),  $C$  is the capacitance (in F), and  $V$  is the p.d. (in V).

The relationship between  $Q$  and  $V$  is linear, as shown in Figure 3.3.21. The slope of this relationship (when  $Q$  is plotted against  $V$ ) gives the resistance. A steep slope corresponds to a high capacitance whilst a shallow slope corresponds to a low capacitance.

The relationship between  $Q$ ,  $C$ , and  $V$  can be re-arranged to make voltage or capacitance the subject as shown below:

$$V = \frac{Q}{C} \quad \text{and} \quad C = \frac{Q}{V}.$$

#### Example 3.3.14

A  $10 \mu\text{F}$  capacitor is charged to a potential of 250 V. Determine the charge stored.

The charge stored will be given by:

$$Q = CV = 10 \times 10^{-6} \times 250 = 2500 \times 10^{-6} = 2.5 \times 10^{-3} = \mathbf{2.5 \text{ mC}}.$$

#### Example 3.3.15

A charge of  $11 \mu\text{C}$  is held in a  $220 \text{ nF}$  capacitor. What voltage appears across the plates of the capacitor?

To find the voltage across the plates of the capacitor we need to re-arrange the equation to make  $V$  the subject, as follows:

$$V = \frac{Q}{C} = \frac{11 \times 10^{-6}}{220 \times 10^{-9}} = \mathbf{50 \text{ V}}.$$

### Capacitance

The capacitance of a capacitor depends upon the physical dimensions of the capacitor (i.e. the size of the plates and the separation between them) and the dielectric material between the plates. The capacitance of a conventional parallel plate capacitor is given by:

$$C = \frac{\epsilon_0 \epsilon_r A}{d}$$

where  $C$  is the capacitance (in F),  $\epsilon_0$  is the *permittivity* of free space,  $\epsilon_r$  is the *relative permittivity* (or *dielectric constant*) of the dielectric medium

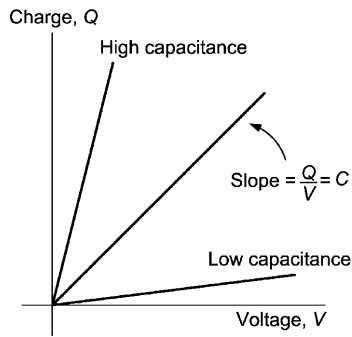


**Figure 3.3.21** Charge plotted against voltage

between the plates),  $A$  is the area of the plates (in  $\text{m}^2$ ), and  $d$  is the separation between the plates (in m). The *permittivity of free space*,  $\epsilon_0$ , is  $8.854 \times 10^{-12} \text{ F/m}$ .

Some typical capacitor dielectric materials and relative permittivity are given in the table below:

| Dielectric material | Relative permittivity (free space = 1) |
|---------------------|----------------------------------------|
| Vacuum              | 1                                      |
| Air                 | 1.0006 (i.e. 1!)                       |
| Polythene           | 2.2                                    |
| Paper               | 2 to 2.5                               |
| Epoxy resin         | 4.0                                    |
| Mica                | 3 to 7                                 |
| Glass               | 5 to 10                                |
| Porcelain           | 6 to 7                                 |
| Aluminium oxide     | 7                                      |
| Ceramic materials   | 15 to 500                              |

### Example 3.3.16

Two parallel metal plates, each of area  $0.2 \text{ m}^2$  are separated by an air gap of  $1 \text{ mm}$ . Determine the capacitance of this arrangement.

Here we must use the formula:

$$C = \frac{\epsilon_0 \epsilon_r A}{d}$$

where  $A = 0.2 \text{ m}^2$ ,  $d = 1 \times 10^{-3} \text{ m}$ ,  $\epsilon_r = 1$ , and  $\epsilon_0 = 8.854 \times 10^{-12} \text{ F/m}$ .

Hence:

$$C = \frac{8.854 \times 10^{-12} \times 1 \times 0.2}{1 \times 10^{-3}} = \frac{1.7708 \times 10^{-12}}{1 \times 10^{-3}} = 1.7708 \times 10^{-9} \text{ F} = \mathbf{1.7708 \text{ nF}}$$

### Example 3.3.17

A capacitor of  $1 \text{ nF}$  is required. If a dielectric material of thickness  $0.5 \text{ mm}$  and relative permittivity  $5.4$  is available, determine the required plate area.

Re-arranging the formula  $C = \frac{\epsilon_0 \epsilon_r A}{d}$  to make  $A$  the subject gives:

$$A = \frac{Cd}{\epsilon_0 \epsilon_r} = \frac{1 \times 10^{-9} \times 0.5 \times 10^{-3}}{8.854 \times 10^{-12} \times 5.4} = \frac{0.5 \times 10^{-12}}{47.811 \times 10^{-12}} = 0.0105 \text{ m}^2$$

thus  $A = \mathbf{0.0105 \text{ m}^2}$  or  $105 \text{ cm}^2$ .

**Figure 3.3.22** Typical construction of a tubular capacitor

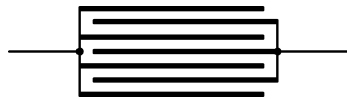
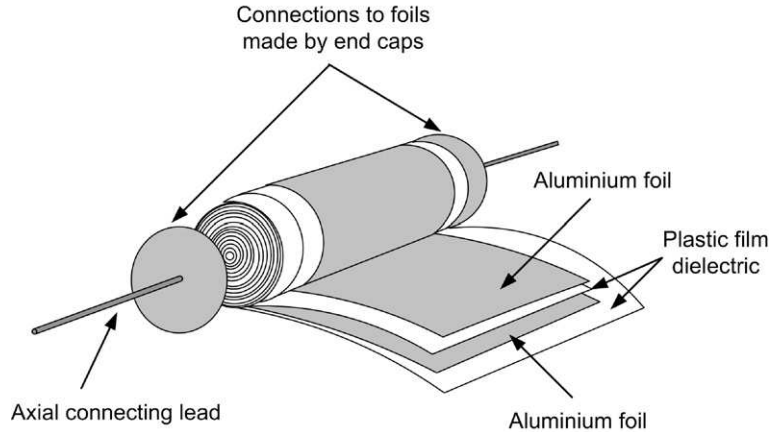


**Figure 3.3.23** A multiple-plate capacitor

In order to increase the capacitance of a capacitor, many practical components employ multiple plates (see Figures 3.3.22 and 3.3.23) in which case the capacitance is then given by:

$$C = \frac{\epsilon_0 \epsilon_r (n - 1) A}{d}$$

where  $C$  is the capacitance (in F),  $\epsilon_0$  is the permittivity of free space,  $\epsilon_r$  is the relative permittivity of the dielectric medium between the plates),  $n$  is the number of plates,  $A$  is the area of the plates (in  $\text{m}^2$ ), and  $d$  is the separation between the plates (in m).

### Example 3.3.18

A capacitor consists of six plates each of area  $20 \text{ cm}^2$  separated by a dielectric of relative permittivity 4.5 and thickness  $0.2 \text{ mm}$ . Determine the capacitance of the capacitor.

Using  $C = \frac{\epsilon_0 \epsilon_r (n - 1) A}{d}$  gives:

$$\begin{aligned} C &= \frac{8.854 \times 10^{-12} \times 4.5 \times (6 - 1) \times (20 \times 10^{-4})}{0.2 \times 10^{-3}} \\ &= \frac{3984.3 \times 10^{-16}}{2 \times 10^{-4}} \\ &= 1992.15 \times 10^{-12} \text{ F} = 1.992 \text{ nF}. \end{aligned}$$

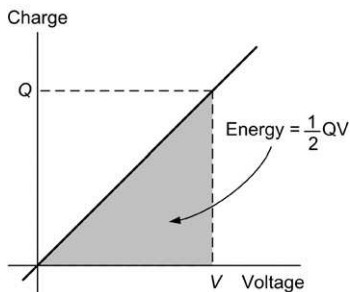


**Figure 3.3.24** Energy stored in a capacitor

### Energy storage

The area under the linear relationship between  $Q$  and  $V$  that we met earlier in Figure 3.3.21 gives the energy stored in the capacitor. This area is shown shaded in Figure 3.3.24. By virtue of its triangular shape the area under the line is  $\frac{1}{2}QV$  thus:

Energy stored,  $W = \frac{1}{2}QV$

Combining this with the earlier relationship,  $Q = CV$ , gives:

$$W = \frac{1}{2}(CV) V = \frac{1}{2}CV^2$$

where  $W$  is the energy (in J),  $C$  is the capacitance (in F), and  $V$  is the p.d. (in V).

This new relationship shows us that the energy stored in a capacitor is proportional to the product of the capacitance and the square of the p.d. between its plates.

### Example 3.3.19

A  $100\ \mu\text{F}$  capacitor is charged from a charge of  $20\text{ V}$  supply. How much energy is stored in the capacitor?

The energy stored in the capacitor will be given by:

$$\begin{aligned} W &= \frac{1}{2} CV^2 = \frac{1}{2} \times 100 \times 10^{-6} \times (20)^2 \\ &= 50 \times 400 \times 10^{-6} \\ &= 20000 \times 10^{-6} = \mathbf{0.02\text{ J}}. \end{aligned}$$

### Example 3.3.20

A capacitor of  $47\ \mu\text{F}$  is required to store energy of  $40\text{ J}$ . Determine the p.d. required to do this.

To find the p.d. (voltage) across the plates of the capacitor we need to re-arrange the equation to make  $V$  the subject, as follows:

$$\begin{aligned} V &= \sqrt{\frac{2W}{C}} = \sqrt{\frac{2 \times 40}{47 \times 10^{-6}}} = \sqrt{\frac{80}{47} \times 10^6} \\ &= \sqrt{1.702 \times 10^6} = 1.3 \times 10^3 = \mathbf{1.3\text{ kV}}. \end{aligned}$$

## Capacitors in series and parallel

In order to obtain a particular value of capacitance, fixed capacitors may be arranged in either series or parallel. Consider Figure 3.3.25 where  $C$  is the equivalent capacitance of the three capacitors ( $C_1$ ,  $C_2$  and  $C_3$ ) connected in series.

The applied voltage,  $V$ , will be the sum of the voltages that appear across each capacitor. Thus:

$$V = V_1 + V_2 + V_3.$$

Now, for each capacitor, the p.d.,  $V$ , across its plates will be given by the ratio of charge,  $Q$ , to capacitance,  $C$ . Hence:

$$V = \frac{Q}{C}, \quad V_1 = \frac{Q_1}{C_1}, \quad V_2 = \frac{Q_2}{C_2}, \quad \text{and} \quad V_3 = \frac{Q_3}{C_3}.$$

Combining these equations gives:

In the series circuit the same charge,  $Q$ , appears across each capacitor, thus:

$$Q = Q_1 = Q_2 = Q_3.$$

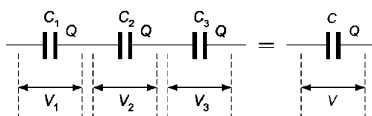


**Figure 3.3.25** Three capacitors in series

Hence:

$$\frac{Q}{C} = \frac{Q}{C_1} + \frac{Q}{C_2} + \frac{Q}{C_3}.$$

From which:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}.$$

When two capacitors are connected in series the equation becomes:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}.$$

This can be arranged to give the slightly more convenient expression:

$$C = \frac{C_1 \times C_2}{C_1 + C_2}.$$

Hence the equivalent capacitance of two capacitors connected in series can be found by taking the *product* of the two capacitance values and *dividing* it by the *sum* of the two capacitance values (in other words, *product over sum*).

Now consider Figure 3.3.26 where  $C$  is the equivalent capacitance of three capacitors ( $C_1$ ,  $C_2$  and  $C_3$ ) connected in parallel.

The total charge present,  $Q$ , will be the sum of the charges that appear in each capacitor. Thus:

$$Q = Q_1 + Q_2 + Q_3.$$

Now, for each capacitor, the charge present,  $Q$ , will be given by the product of the capacitance,  $C$ , and p.d.,  $V$ . Hence:

$$Q = CV, \quad Q_1 = C_1V_1, \quad Q_2 = C_2V_2, \quad \text{and} \quad Q_3 = C_3V_3.$$

Combining these equations gives:

$$CV = C_1V_1 + C_2V_2 + C_3V_3.$$

In the parallel circuit the same voltage,  $V$ , appears across each capacitor, thus:

$$V = V_1 = V_2 = V_3.$$

Hence:

$$CV = C_1V + C_2V + C_3V.$$

From which:

$$C = C_1 + C_2 + C_3.$$

When two capacitors are connected in parallel the equation becomes:

$$C = C_1 + C_2.$$

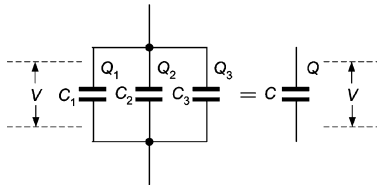


**Figure 3.3.26** Three capacitors in parallel



**Example 3.3.21**

Capacitors of  $2.2\ \mu\text{F}$  and  $6.8\ \mu\text{F}$  are connected (a) in series and (b) in parallel. Determine the equivalent value of capacitance in each case.

- (a) Here we can use the simplified equation for *just two* capacitors connected in series:

$$C = \frac{C_1 \times C_2}{C_1 + C_2} = \frac{2.2 \times 6.8}{2.2 + 6.8} = \frac{14.96}{9} = \mathbf{1.66\ \mu\text{F}}.$$

- (b) Here we use the formula for two capacitors connected in parallel:

$$C = C_1 + C_2 = 2.2 + 6.8 = \mathbf{9\ \mu\text{F}}.$$

**Example 3.3.22**

Capacitors of  $2\ \mu\text{F}$  and  $5\ \mu\text{F}$  are connected in series across a  $100\ \text{V}$  DC supply. Determine:

- (a) the charge on each capacitor and (b) the voltage dropped across each capacitor.

- (a) First we need to find the equivalent value of capacitance,  $C$ , using the simplified equation for two capacitors in series:

$$C = \frac{C_1 \times C_2}{C_1 + C_2} = \frac{2 \times 5}{2 + 5} = \frac{10}{7} = 1.43\ \mu\text{F}.$$

Next we can determine the charge (note that, since the capacitors are connected in series, the *same* charge will appear in each capacitor):

$$Q = CV = 1.43 \times 10^{-6} \times 100 = \mathbf{143\ \mu\text{C}}.$$

- (b) In order to determine the voltage dropped across each capacitor we can use:

$$V = \frac{Q}{C}.$$

Hence, for the  $2\ \mu\text{F}$  capacitor:

$$V_1 = \frac{Q}{C_1} = \frac{143 \times 10^{-6}}{2 \times 10^{-6}} = 71.5\ \text{V}.$$

Similarly, for the  $5\ \mu\text{F}$  capacitor:

$$V_2 = \frac{Q}{C_2} = \frac{143 \times 10^{-6}}{5 \times 10^{-6}} = 28.5\ \text{V}.$$

We should now find that the total voltage ( $100\ \text{V}$ ) applied to the series circuit is the sum of the two capacitor voltages, that is:

$$V = V_1 + V_2 = 71.5 + 28.5 = \mathbf{100\ \text{V}}.$$

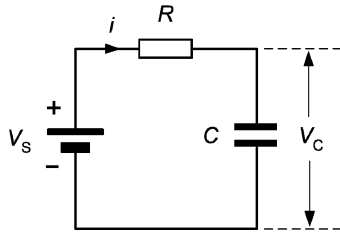


**Figure 3.3.27** *C–R circuit with C charging through R*

### C–R circuits

Networks of capacitors and resistors (known as *C–R networks*) form the basis of many simple timing circuits. When the *C–R* network is connected to a constant voltage source ( $V_s$ ), as shown in Figure 3.3.27, the voltage ( $v_C$ ) across the (initially uncharged) capacitor will rise exponentially as shown in Figure 3.3.28. At the same time, the current in the circuit ( $i$ ) will fall, also as shown in Figure 3.3.28.

The rate of growth of voltage with time and decay of current with time will be dependent upon the product of capacitance and resistance. This value is known as the *time constant* of the circuit. Hence:

$$\text{Time constant, } \tau = C \times R$$

where  $C$  is the value of capacitance (in F),  $R$  is the resistance (in  $\Omega$ ), and  $\tau$  is the time constant (in s).

The voltage developed across the charging capacitor ( $v_C$ ) varies with time ( $t$ ) according to the relationship:

$$v_C = V_s \left( 1 - e^{-t/CR} \right)$$

where  $v_C$  is the capacitor voltage (in V),  $V_s$  is the DC supply voltage (in V),  $t$  is the time (in s), and  $CR$  is the time constant of the circuit (equal to the product of capacitance,  $C$ , and resistance,  $R$ , in s).

The capacitor voltage will rise to approximately 63% of the supply voltage in a time interval equal to the time constant. At the end of the next interval of time equal to the time constant (i.e. after an elapsed time equal to  $2CR$ ) the voltage will have risen by 63% of the remainder, and so on. In theory, the capacitor will never quite become fully charged. However, after a period of time equal to  $5CR$ , the capacitor voltage will to all intents and purposes be equal to the supply voltage. At this point

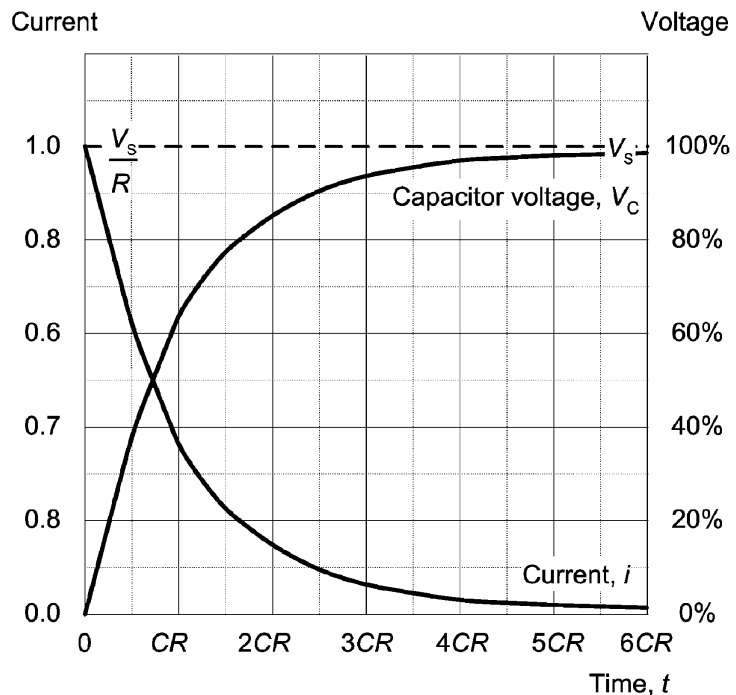


**Figure 3.3.28** *Exponential growth of capacitor voltage ( $v_C$ ) and corresponding decay of current ( $i$ ) in Figure 3.3.27*

the capacitor voltage will have risen to 99.3% of its final value and we can consider it to be fully charged.

During charging, the current in the capacitor, ( $i$ ) varies with time ( $t$ ) according to the relationship:

$$i = V_S e^{-t/CR}$$

where  $i$  is the current (in A),  $V_S$  is the DC supply voltage (in V),  $t$  is the time, and  $CR$  is the time constant of the circuit (equal to the product of capacitance,  $C$ , and resistance,  $R$ , in s).

The current will fall to approximately 37% of the initial current in a time equal to the time constant. At the end of the next interval of time equal to the time constant (i.e. after a total time of  $2CR$  has elapsed) the current will have fallen by a further 37% of the remainder, and so on.

### Example 3.3.23

An initially uncharged capacitor of  $1\ \mu\text{F}$  is charged from a  $9\ \text{V}$  DC supply via a  $3.3\ \text{M}\Omega$  resistor. Determine the capacitor voltage  $1\ \text{s}$  after connecting the supply.

The formula for exponential growth of voltage in the capacitor is:

$$v_C = V_S(1 - e^{-t/CR})$$

where  $V_S = 9\ \text{V}$ ,  $t = 1\ \text{s}$  and  $CR = 1\ \mu\text{F} \times 3.3\ \text{M}\Omega = 3.3\ \text{s}$

$$v_C = 9(1 - e^{-1/3.3}) = 9(1 - 0.738) = \mathbf{2.358\ \text{V}}.$$

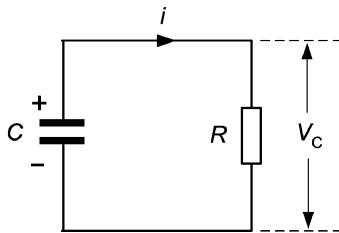


**Figure 3.3.29**  $C$ – $R$  circuit with  $C$  discharging through  $R$

A charged capacitor contains a reservoir of energy stored in the form of an electric field. When the fully charged capacitor from Figure 3.3.27 is connected as shown in Figure 3.3.29, the capacitor will discharge through the resistor, and the capacitor voltage ( $v_C$ ) will fall exponentially with time, as shown in Figure 3.3.30. The current in the circuit ( $i$ ) will also fall, as shown in Figure 3.3.30. The rate of discharge (i.e. the rate of decay of voltage with time) will once again be governed by the time constant of the circuit ( $C \times R$ ).

The voltage developed across the discharging capacitor ( $v_C$ ) varies with time ( $t$ ) according to the relationship:

$$v_C = V_S e^{-t/CR}$$

where  $v_C$  is the capacitor voltage (in V),  $V_S$  is the DC supply voltage (in V),  $t$  is the time (in s), and  $CR$  is the time constant of the circuit (equal to the product of capacitance,  $C$ , and resistance,  $R$ , in s) (Figure 3.3.30).

The capacitor voltage will fall to approximately 37% of the initial voltage in a time equal to the time constant. At the end of the next interval of time equal to the time constant (i.e. after an elapsed time equal to  $2CR$ ) the voltage will have fallen by 37% of the remainder, and so on. In theory, the capacitor will never quite become fully discharged. After a period of time equal to  $5CR$ , however, the capacitor voltage will to all

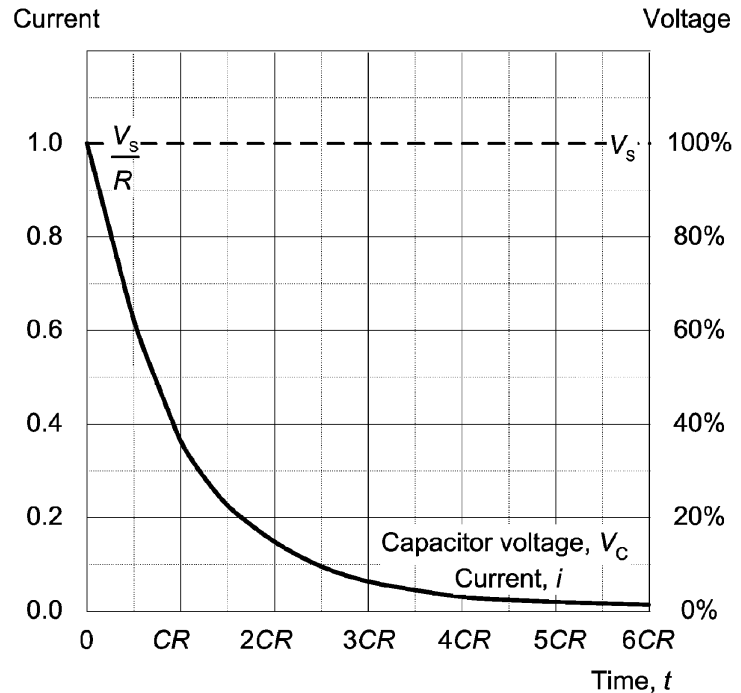


**Figure 3.3.30** Exponential decay of capacitor voltage ( $v_C$ ) and corresponding decay of current ( $i$ ) in Figure 3.3.29

intents and purposes be zero. At this point the capacitor voltage will have fallen below 1% of its initial value. At this point we can consider it to be fully discharged.

As with charging, the current in the capacitor,  $i$ , varies with time,  $t$ , according to the relationship:

$$i = V_S e^{-t/CR}$$

where  $i$  is the current (in A),  $V_S$  is the DC supply voltage (in V),  $t$  is the time, and  $CR$  is the time constant of the circuit (equal to the product of capacitance,  $C$ , and resistance,  $R$ , in s).

The current will fall to approximately 37% of the initial current in a time equal to the time constant. At the end of the next interval of time equal to the time constant (i.e. after a total time of  $2CR$  has elapsed) the current will have fallen by a further 37% of the remainder, and so on.

### Example 3.3.24

A  $10\ \mu\text{F}$  capacitor is charged to a potential of 20 V and then discharged through a  $47\ \text{k}\Omega$  resistor. Determine the time taken for the capacitor voltage to fall below 10 V.

The formula for exponential decay of voltage in the capacitor is:

$$v_C = V_S e^{-t/CR}$$

In this case,  $V_S = 20\ \text{V}$  and  $CR = 10\ \mu\text{F} \times 47\ \text{k}\Omega = 0.47\ \text{s}$ , and we need to find  $t$  when  $v_C = 10\ \text{V}$ .

Re-arranging the formula to make  $t$  the subject gives:

$$t = -CR \times \ln\left(\frac{V_c}{V_s}\right)$$

thus:

$$t = -0.47 \times \ln\left(\frac{10}{20}\right) = -0.47 \times -0.693 = \mathbf{0.325 \text{ s.}}$$

In order to simplify the mathematics of exponential growth and decay, the table below provides an alternative tabular method that may be used to determine the voltage and current in a  $C$ - $R$  circuit:

| $\frac{t}{CR}$ | $k$ (ratio of instantaneous value to final value) |                             |
|----------------|---------------------------------------------------|-----------------------------|
|                | Exponential growth                                | Exponential decay           |
| 0.0            | 0.0000                                            | 1.0000                      |
| 0.1            | 0.0951                                            | 0.9048                      |
| 0.2            | 0.1812                                            | 0.8187 (see Example 3.3.25) |
| 0.3            | 0.2591                                            | 0.7408                      |
| 0.4            | 0.3296                                            | 0.6703                      |
| 0.5            | 0.3935                                            | 0.6065                      |
| 0.6            | 0.4511                                            | 0.5488                      |
| 0.7            | 0.5034                                            | 0.4965                      |
| 0.8            | 0.5506                                            | 0.4493                      |
| 0.9            | 0.5934                                            | 0.4065                      |
| 1.0            | 0.6321                                            | 0.3679                      |
| 1.5            | 0.7769                                            | 0.2231                      |
| 2.0            | 0.8647                                            | 0.1353                      |
| 2.5            | 0.9179                                            | 0.0821                      |
| 3.0            | 0.9502                                            | 0.0498                      |
| 3.5            | 0.9698                                            | 0.0302                      |
| 4.0            | 0.9817                                            | 0.0183                      |
| 4.5            | 0.9889                                            | 0.0111                      |
| 5.0            | 0.9933                                            | 0.0067                      |

### Example 3.3.25

A  $150 \mu\text{F}$  capacitor is charged to a potential of  $150 \text{ V}$ . The capacitor is then removed from the charging source and connected to a  $2 \text{ M}\Omega$  resistor. Determine the capacitor voltage  $1 \text{ min}$  later.

We will solve this problem using the tabular method rather than using the exponential formula. First we need to find the time constant:

$$C \times R = 150 \mu\text{F} \times 2 \text{ M}\Omega = 300 \text{ s.}$$

Next we find the ratio of  $t$  to  $CR$ . After  $1 \text{ min}$ ,  $t = 60 \text{ s}$  therefore the ratio of  $t$  to  $CR$  is:

$$\frac{t}{CR} = \frac{60}{300} = 0.2.$$

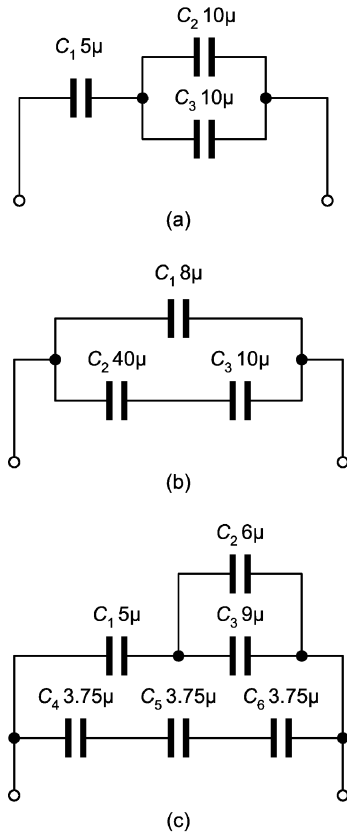


**Figure 3.3.31** See Problems 3.3.2

Referring to the table we find that when  $t/CR = 0.2$ , the ratio of instantaneous value to final value ( $k$ ) for decay is 0.8187.

Thus:

$$\frac{v_C}{V_S} = 0.8187$$

or

$$v_C = 0.8187 \times 150 = \mathbf{122.8 \text{ V.}}$$

### Problems 3.3.2

1. A capacitor holds a charge of  $60 \mu\text{C}$  when  $240 \text{ V}$  is applied to its plates. What is the value of the capacitor?
2. A  $220 \mu\text{F}$  capacitor is required to store  $11 \text{ mJ}$  of energy. What voltage needs to appear across plates of the capacitor?
3. Two parallel metal plates measuring  $10 \text{ cm}$  by  $10 \text{ cm}$  are separated by a dielectric material having a thickness of  $0.5 \text{ mm}$  and a relative permittivity of  $5.4$ . Determine the capacitance of this arrangement.
4. Determine the equivalent capacitance of each of the networks of capacitors shown in Figure 3.3.31.
5. A  $C$ - $R$  circuit consists of resistor of  $2 \text{ M}\Omega$  connected in series with a capacitor of  $10\,000 \mu\text{F}$ . If the capacitor is initially discharged, determine the capacitor voltage  $10 \text{ s}$  after the circuit is connected to a  $5 \text{ V}$  supply.

### Magnetic field strength and current

Figure 3.3.32 shows the magnetic field that exists around a conductor arranged to form a coil. The magnetic flux is concentrated in the centre of the coil and the magnetic field takes the same shape as that which would be produced by a bar magnet. This arrangement forms a simple *inductor*. The magnetic field strength,  $H$ , will be given by:

$$H = \frac{NI}{l}$$

where  $H$  is the magnetic field strength (in  $\text{A/m}$ ),  $N$  is the number of turns, and  $l$  is the mean length of the magnetic circuit (in  $\text{m}$ ).

With most coils and inductors we are not concerned so much the *total* flux produced but more with how dense or concentrated the flux is. To increase the flux density we introduce a ferromagnetic core material into the centre of the coil.

The density of the magnetic flux,  $B$ , produced by the coil is given by:

$$B = \frac{\Phi}{A}$$

where  $B$  is the magnetic flux density (in  $\text{T}$ ),  $\Phi$  is the flux (in  $\text{Wb}$ ), and  $A$  is the area (in  $\text{m}^2$ ) over which the flux acts.

For any given magnetic medium, the ratio of flux density,  $B$ , to magnetic field strength,  $H$ , is a constant. This constant is known as the *permeability* of the medium and it is represented by the symbol,  $\mu$ . Hence:

$$\mu = \frac{B}{H}.$$

Figure 3.3.33 shows the linear relationship that exists between  $B$  and  $H$  for some typical magnetic materials. It is important to note that there is a

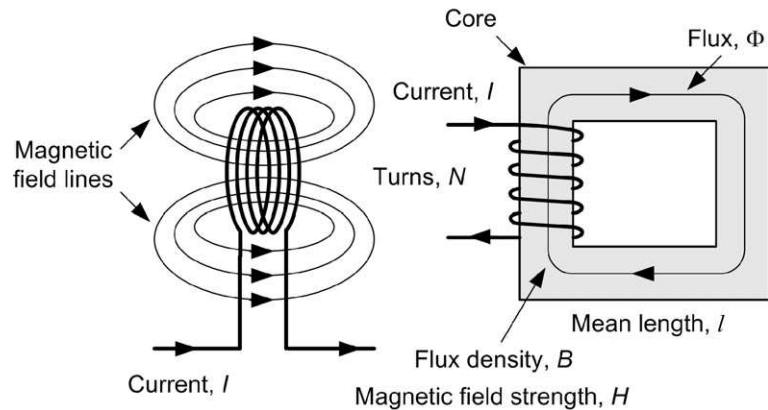


**Figure 3.3.32** Magnetic field around a coil

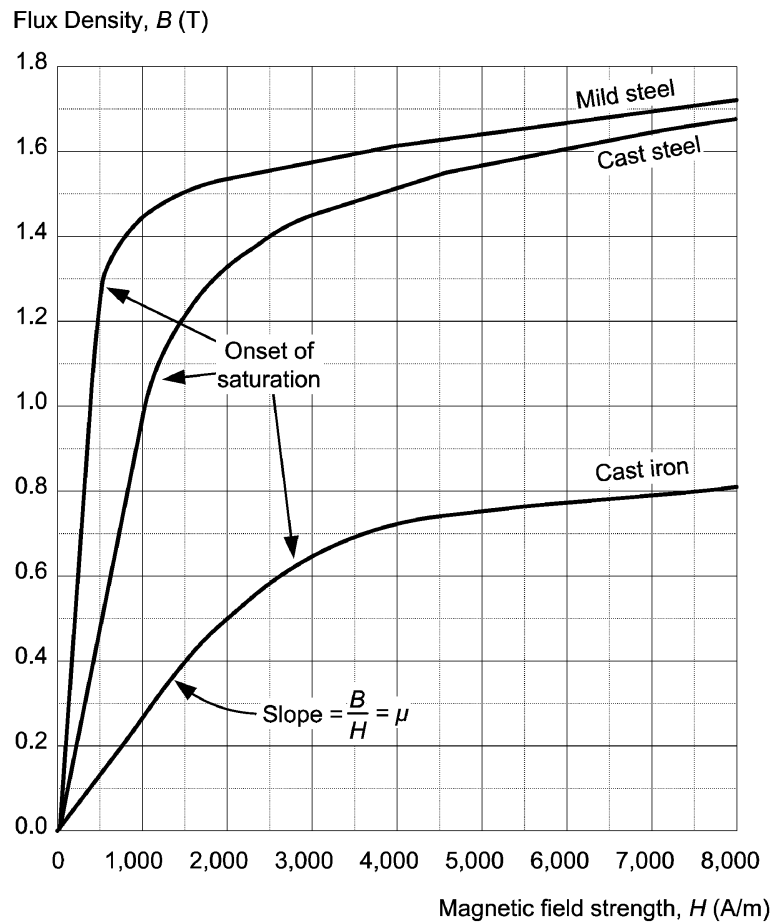


**Figure 3.3.33** Linear relationship between flux density and magnetising force for various magnetic materials

limit to the density of flux that these materials can support and we refer to this as saturation. When this occurs, any further increase in magnetic field strength (magnetising force) produces no significant further increase in flux density.

### Example 3.3.26

An inductor consists of 40 turns of wire wound on a toroidal core having a mean length of 0.2 m and cross-sectional area  $0.05 \text{ m}^2$ . If the core material has a permeability of  $4 \times 10^{-3} \text{ H/m}$ , determine the flux produced when a current of 6 A flows in the coil.

Now since  $\mu = \frac{B}{H}$  we can deduce that  $B = \mu H$ .

Combining this with  $H = \frac{NI}{l}$  gives  $B = \mu H = \mu \frac{NI}{l}$ .

Since  $B = \frac{\Phi}{A}$  we can obtain an expression for  $\Phi$  in terms of  $N$ ,  $I$ ,  $l$ , and  $A$ :

$$\Phi = B \times A = \mu \frac{NIA}{l} = 4 \times 10^{-3} \frac{40 \times 6 \times 0.05}{0.2} \text{ Wb}$$

$$\Phi = 4 \times 10^{-3} \frac{40 \times 6 \times 0.05}{0.2} = \mathbf{0.24 \text{ T.}}$$

### Inductance

The inductance of an inductor depends upon the physical dimensions of the inductor (e.g. the length and diameter of the winding), the number of turns, and the permeability of the material of the core. The inductance,  $L$ , of an inductor is given by:

$$L = \frac{\mu N^2 A}{l}$$

where  $L$  is the inductance (in H),  $\mu$  is the permeability of the core material,  $N$  is the number of turns,  $l$  is the length of the core (in m), and  $A$  is the cross-sectional area of the core (in  $\text{m}^2$ ).

It is often convenient to specify the permeability of a magnetic material,  $\mu$ , relative to those of air or free space, hence:

$$\mu = \mu_0 \times \mu_r$$

where  $\mu$  is the *absolute permeability* of the material (in H/m),  $\mu_0$  is the permeability of free space ( $12.57 \times 10^{-7} \text{ H/m}$ ) and,  $\mu_r$  is the relative permeability of the core material.

Combining the two foregoing relationships gives:

$$L = \frac{\mu_0 \mu_r N^2 A}{l}$$

where  $L$  is the inductance (in H),  $\mu_0$  is the *permeability of free space* ( $12.57 \times 10^{-7} \text{ H/m}$ ),  $\mu_r$  is the *relative permeability* of the core material,



$l$  is the length of the core (in m), and  $A$  is the cross-sectional area of the core (in m<sup>2</sup>).

Some typical core materials and relative permeability are given in the table below:

| Core material | Relative permeability (free space = 1) |
|---------------|----------------------------------------|
| Vacuum or air | 1                                      |
| Cast iron     | 100 to 250                             |
| Mild steel    | 200 to 800                             |
| Cast steel    | 300 to 900                             |
| Silicon iron  | 1000 to 5000                           |
| Mumetal       | 200 to 5000                            |
| Stalloy       | 500 to 6000                            |

### Example 3.3.27

An inductor of 100 mH is required. If a closed magnetic core of length 20 cm, cross-sectional area 15 cm<sup>2</sup> and relative permeability 500 is available, determine the number of turns required.

*Solution*

$$\text{Now } L = \frac{\mu_0 \mu_r N^2 A}{l} \text{ and hence } N = \sqrt{\frac{LI}{\mu_0 \mu_r A}}$$

$$\begin{aligned} \text{Thus } N &= \sqrt{\frac{LI}{\mu_0 \mu_r A}} = \sqrt{\frac{100 \times 10^{-3} \times 20 \times 10^{-2}}{12.57 \times 10^{-7} \times 500 \times 15 \times 10^{-4}}} \\ &= \sqrt{\frac{2 \times 10^{-2}}{94\,275 \times 10^{-11}}} = \sqrt{21\,215} = 146 \end{aligned}$$

Hence the inductor requires 146 turns of wire.

### Energy storage

The energy stored in an inductor is proportional to the product of the inductance and the square of the current flowing in it. Thus:

$$W = 0.5 LI^2$$

where  $W$  is the energy (in J),  $L$  is the inductance (in H), and  $I$  is the current (in A).

### Example 3.3.28

A current of 1.5 A flows in an inductor of 5 H. Determine the energy stored.

*Solution*

$$\text{Now } W = 0.5 LI^2 = 0.5 \times 5 \times 1.5^2 = \mathbf{5.625 \text{ J.}}$$

**Example 3.3.29**

An inductor of 20 mH is required to store 2.5 J of energy. Determine the current that must be applied to the inductor.

Now  $W = 0.5LI^2$  and hence:

$$I = \sqrt{\frac{W}{0.5L}} = \sqrt{\frac{2.5}{0.5 \times 20 \times 10^{-3}}} = \sqrt{\frac{2.5}{0.5 \times 20 \times 10^{-3}}} \\ = \sqrt{2.5 \times 10^2} = \mathbf{15.8 \text{ A.}}$$

**Self and mutual inductance**

An induced e.m.f. (often referred to as a *back e.m.f.*) is produced whenever a change of flux occurs in an inductor. This back e.m.f. is proportional to the rate of change of current (from Lenz's law) and this effect is known as *self-inductance* (though we often just refer to it as *inductance*). The e.m.f. induced,  $e$ , is given by:

$$e = -L \frac{di}{dt}$$

where  $L$  is the inductance,  $di/dt$  is the rate of change of current and the minus sign indicates that the polarity of the generated e.m.f. *opposes* the change.

A coil has an inductance of 1 H if a voltage of 1 V is induced across it when a current changing at the rate of 1 A/s is flowing in it.

**Example 3.3.30**

A coil has an inductance of 15 mH and is subject to a current that changes at a rate of 450 A/s. What e.m.f. is produced?

*Solution*

Now  $e = -L \frac{di}{dt}$  and hence:

$$e = -15 \times 10^{-3} \times 450 = -6.75 \text{ V.}$$

Note the minus sign! In other words, a *back e.m.f.* of 6.75 V is induced.

**Example 3.3.31**

A current increases at a uniform rate from 2 A to 6 A in a time of 250 ms. If this current is applied to an inductor determine the value of inductance if a back e.m.f. of 15 V is produced across its terminals.

*Solution*

Now  $e = -L \frac{di}{dt}$  and hence  $L = -e \frac{dt}{di}$ .

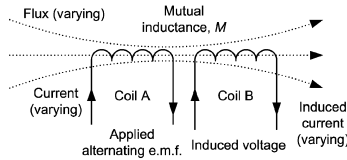


**Figure 3.3.34** Mutual inductance

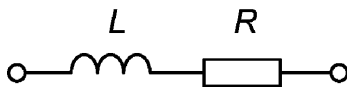


**Figure 3.3.35** A coil has both inductance and resistance

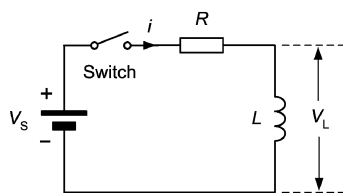


**Figure 3.3.36** Circuit in which a current is applied to an inductor

$$\begin{aligned}\text{Thus } L &= -(-15) \times \frac{250 \times 10^{-3}}{(6 - 2)} = 15 \times 62.5 \times 10^{-3} \\ &= 937.5 \times 10^{-3} = \mathbf{0.94 \text{ H}}.\end{aligned}$$

Finally, when two inductors are placed close to one another, the flux generated when a changing current flows in the first inductor will cut through the other inductor (see Figure 3.3.34). This changing flux will, in turn, induce a current in the second inductor. This effect is known as *mutual inductance* and it occurs whenever two inductors are *inductively coupled*. This is the principle of a very useful component, the *transformer*, which we shall meet later.

### L-R circuits

In practice every coil comprises both inductance and resistance and the circuit of Figure 3.3.35 shows these as two discrete components. In reality the inductance,  $L$ , and resistance,  $R$ , are both distributed throughout the component but it is convenient to treat the inductance and resistance as separate components in the analysis of the circuit.

Now let us consider what happens when a current is first applied to an inductor. If the switch in Figure 3.3.36 is left open, no current will flow and no magnetic flux will be produced by the inductor. If the switch is now closed, current will begin to flow as energy is taken from the supply in order to establish the magnetic field. However, the change in magnetic flux resulting from the appearance of current creates a voltage (an induced e.m.f.) across the coil which opposes the applied e.m.f. from the battery.

The induced e.m.f. results from the changing flux and it effectively prevents an instantaneous rise in current in the circuit. Instead, the current increases slowly to a maximum at a rate which depends upon the ratio of inductance,  $L$ , to resistance,  $R$ , present in the circuit (Figure 3.3.37).

After a while, a steady state condition will be reached in which the voltage across the inductor will have decayed to zero and the current will have reached a maximum value (determined by the ratio of  $V$  to  $R$ , i.e. using Ohm's law!).

If, after this steady state condition has been achieved, the switch is opened once again, the magnetic field will suddenly collapse and the energy will be returned to the circuit in the form of a *back e.m.f.* which will appear momentarily across the coil as the field collapses.

### Problems 3.3.3

1. An inductor is wound on a toroidal core having a mean length of 0.25 m and cross-sectional area 0.075 m<sup>2</sup>. If the core material has a permeability of 1200, determine the number of turns required on the coil if a flux of 0.5 T is to be produced by a current of 2 A flowing in the coil.
2. An inductor has a closed magnetic core of length 40 cm, cross-sectional area 10 cm<sup>2</sup> and relative permeability 450. Determine the value of inductance if the inductor has 250 turns of wire.
3. An inductor of 600 mH is required to store 400 mJ of energy. Determine the current that must be applied to the inductor.
5. A current increases at a uniform rate from 1.5 A to 4.5 A in a time of 50 ms. If this current is applied to a 2 H inductor determine the value of induced e.m.f.

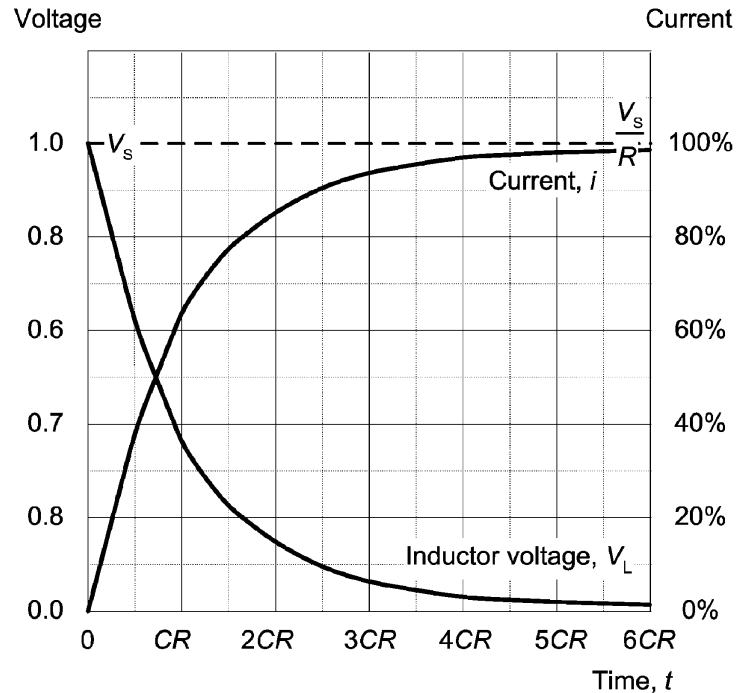


**Figure 3.3.37** Voltage and current in the circuit of Figure 3.3.36

### Basic generator principles

When a conductor is moved through a magnetic field, an e.m.f. will be induced across its ends. An induced e.m.f. will also be generated if the conductor remains stationary whilst the field moves. In either case, cutting at right angles through the lines of magnetic flux (see Figure 3.3.38) results in a generated e.m.f. the magnitude of which will be given by:

$$E = Blv$$

where  $B$  is the magnetic flux density (in T),  $l$  is the length of the conductor (in m), and  $v$  is the velocity of the field (in m/s).

If the field is cut at an angle,  $\theta$ , (rather than at right angles) the generated e.m.f. will be given by:

$$E = Blv \sin \theta$$

where  $\theta$  is the angle between the direction of motion of the conductor and the field lines.

### Example 3.3.32

A conductor of length 20 cm moves at 0.5 m/s through a uniform perpendicular field of 0.6 T. Determine the e.m.f. generated.

Since the field is perpendicular to the conductor, the angle is  $90^\circ$  ('perpendicular' means the same as 'at right angles') we can use the basic equation:

$$E = Blv$$

where  $B = 0.6$  T,  $l = 20$  cm = 0.02 m, and  $v = 0.5$  m/s. Thus:

$$E = Blv = 0.6 \times 0.02 \times 0.5 = 0.006 \text{ V} = \mathbf{6 \text{ mV.}}$$

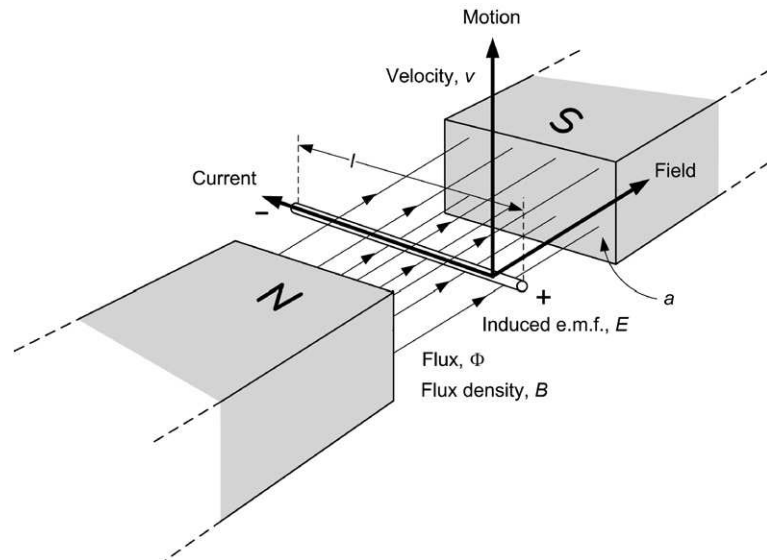


**Figure 3.3.38** Generating an e.m.f. by moving a conductor through a magnetic field

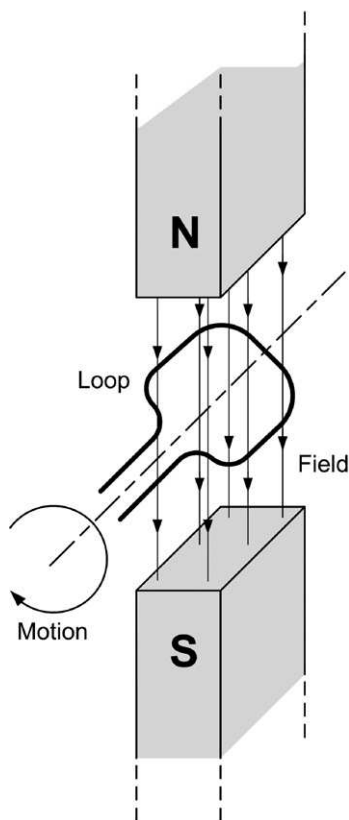


**Figure 3.3.39** A loop rotating within a magnetic field

### A simple AC generator

Being able to generate a voltage by moving a conductor through a magnetic field is extremely useful as it provides us with an easy way of generating electricity. Unfortunately, moving a wire at a constant linear velocity through a uniform magnetic field presents us with a practical problem simply because the mechanical power that can be derived from an aircraft engine is available in rotary (rather than linear) form!

The solution to this problem is that of using the rotary power available from the engine (via a suitable gearbox and transmission) to rotate a conductor shaped into the form of loop as shown in Figure 3.3.39. The loop is made to rotate inside a permanent magnetic field with opposite poles (N and S) on either side of the loop.

There now remains the problem of making contact with the loop as it rotates inside the magnetic field but this can be overcome by means of a pair of carbon brushes and copper slip rings. The brushes are spring loaded and held against the rotating slip rings so that, at any time, there is a path for current to flow from the loop to the load to which it is connected (Figure 3.3.40).

The opposite sides of the loop consist of conductors that move through the field. At  $0^\circ$  (with the loop vertical as shown in Figure 3.3.40(a)) the opposite sides of the loop will be moving in the same direction as the lines of flux. At that instant, the angle,  $\theta$ , at which the field is cut is  $0^\circ$  and since the sine of  $0^\circ$  is 0 the generated voltage (from  $E = Blv \sin \theta$ ) will consequently also be zero.

If the loop has rotated to a position which is  $90^\circ$  from that shown in Figure 3.3.40(b), the two conductors will effectively be moving at right angles to the field. At that instant, the generated e.m.f. will take a maximum value (since the sine of  $90^\circ$  is 1).

At  $180^\circ$  from the starting position the generated e.m.f. will have fallen back to zero since, once again, the conductors are moving along the flux lines (but in the direction opposite to that at  $0^\circ$ ).

At  $270^\circ$  the conductors will once again be moving in a direction which is perpendicular to the flux lines (but in the direction opposite to that at  $90^\circ$ ). At this point, a maximum generated e.m.f. will once again be produced. It is, however, important to note that the e.m.f. generated at this instant will be of opposite polarity to that which was generated at  $90^\circ$ . The reason for this is simply that the relative direction of motion (between the conductors and flux lines) has effectively been reversed.

Since  $E = Blv \sin \theta$ , the e.m.f. generated by the arrangement shown in Figure 3.3.40 will take a sinusoidal form, as shown in Figure 3.3.41. Note that the maximum values of e.m.f. occur at  $90^\circ$  and  $270^\circ$  and that the generated voltage is zero at  $0^\circ$ ,  $180^\circ$  and  $360^\circ$ .

In practice, the single loop shown in Figure 3.3.39 would comprise a coil of wire wound on a suitable non-magnetic former. This coil of wire effectively increases the length of the conductor within the magnetic field and the generated e.m.f. will then be directly proportional to the number of turns on the coil.

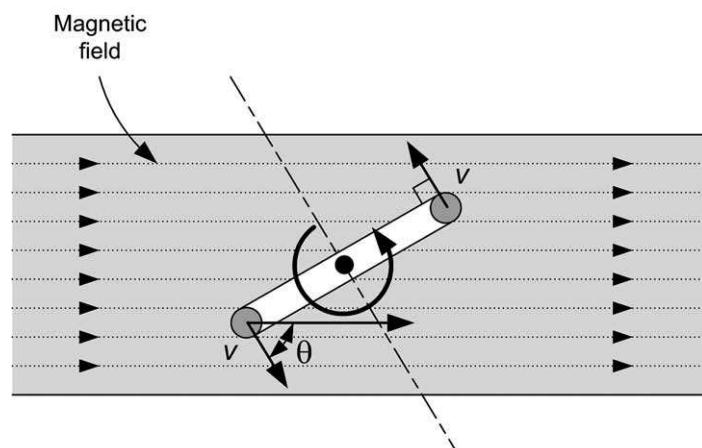
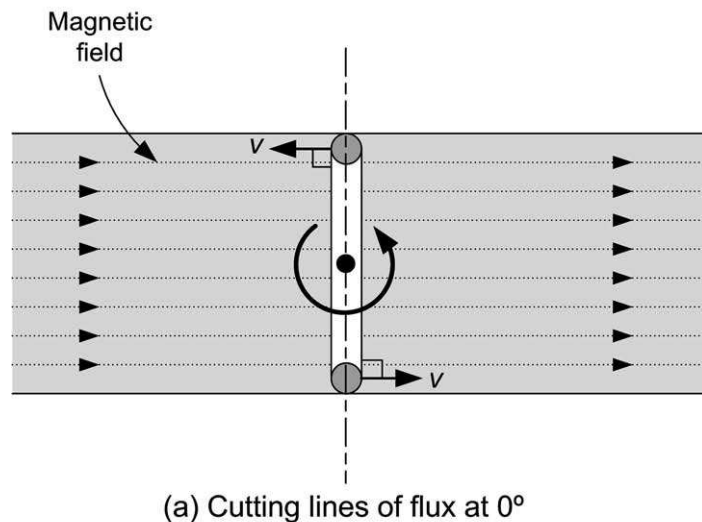


**Figure 3.3.40** Relative angle between loop and field

**Figure 3.3.41** Sinusoidal voltage produced by the rotating loop

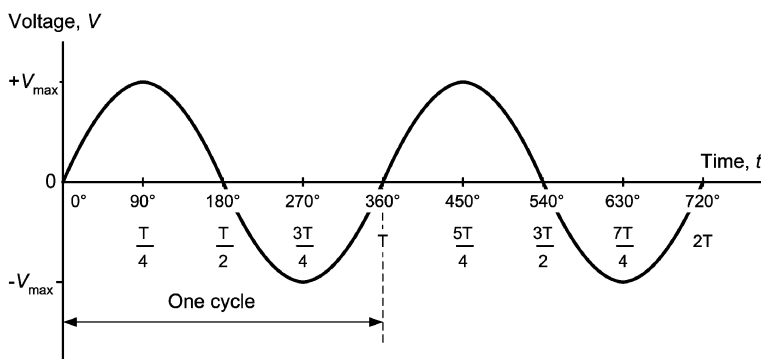
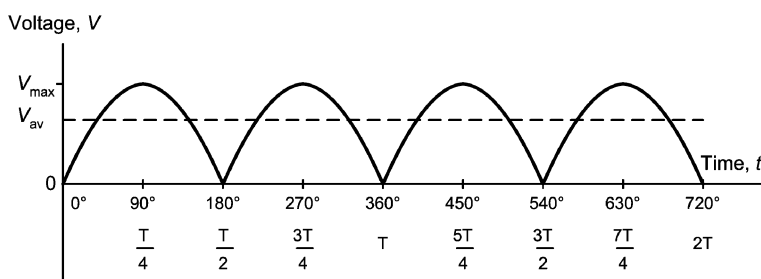


**Figure 3.3.42** E.m.f. generated (compare with Figure 3.3.41)



## DC generators

When connected to a load, the simple generator shown in Figure 3.3.39 produces a sinusoidal AC output. In many applications a steady DC output may be preferred. This can be achieved by means of a *commutator* arrangement which functions as a rotating reversing switch which ensures that the e.m.f. generated by the loop is reversed after rotating through  $180^\circ$ . The generated e.m.f. for this arrangement is shown in Figure 3.3.41. Its worth comparing this waveform with that shown in Figure 3.3.41.

The generated e.m.f. shown in Figure 3.3.42, whilst *unipolar* (i.e. all positive or all negative), is clearly far from ideal since a DC power source should provide a constant voltage output rather than a series of half-sine pulses.

In real generators, a coil comprising a large number of turns of conducting wire replaces the single-turn rotating loop. This arrangement effectively increases the total length of the conductor within the magnetic field and, as a result, also increases the generated output voltage. The output voltage also depends on the density of the magnetic flux through which the current carrying conductor passes. The denser the field the greater the output voltage will be.

## DC motors

A simple DC motor consists of a very similar arrangement to that of the DC generator that we met earlier. A loop of wire that is free to rotate is placed inside a permanent magnetic field. When a DC current is applied to the loop of wire, two equal and opposite forces are set up which act on the conductor and cause it to rotate.

The direction of the forces acting on each arm of the conductor can be established by again using the right-hand grip rule and Fleming's left-hand rule. Now because the conductors are equidistant from their pivot

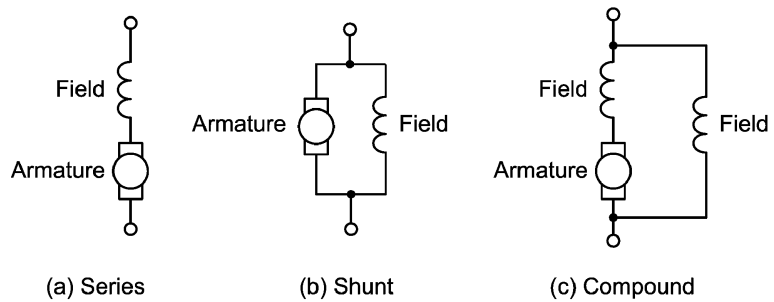


**Figure 3.3.43** Armature and field connections used for DC motors

point and the forces acting on them are *equal and opposite*, then they form a *couple*. The *moment* of this couple is equal to the magnitude of a single force multiplied by the distance between them and this moment is known as *torque*,  $T$ . Now,

$$T = Fd$$

where  $T$  is the torque (in Newton-metres, Nm),  $F$  is the force (N), and  $d$  is the distance (m).

We already know that the magnitude of the force  $F$  is given by  $F = BIl$ , therefore the torque produced by the current carrying thus the torque expression can be written:

$$T = BIl d$$

where  $T$  is the torque (Nm),  $B$  is the flux density (T),  $I$  is the current (A),  $l$  is the length of conductor in the magnetic field (m), and  $d$  is the distance (m).

The torque produces a *turning moment* such that the coil or loop rotates within the magnetic field and this rotation continues for as long as a current is applied.

In real motors, this rotating coil is known as the *armature* and consists of many hundreds of turns of conducting wire. This arrangement is needed in order to maximise the force imposed on the conductor by introducing the *longest possible* conductor into the magnetic field. Also from the relationship  $F = BIl$  it can be seen that the force used to provide the torque in a motor is directly proportional to the size of the magnetic flux,  $B$ . In a real motor an electromagnet is used to produce this flux and this coil is referred to as a *field winding*.

The field winding of a DC motor can be connected in various different ways according to the application envisaged for the motor in question. The following configurations are possible (see Figure 3.3.43):

- series-wound;
- shunt-wound;
- compound wound (where both series and shunt windings are present).

### Transformer principles

The principle of the transformer is illustrated in Figure 3.3.44. The primary and secondary windings are wound on a common low-reluctance magnetic core consisting of a number of steel laminations. All of the alternating flux generated by the primary winding is therefore coupled into the secondary winding (very little flux escapes due to leakage). A sinusoidal current flowing in the primary winding produces a sinusoidal flux within the transformer core.



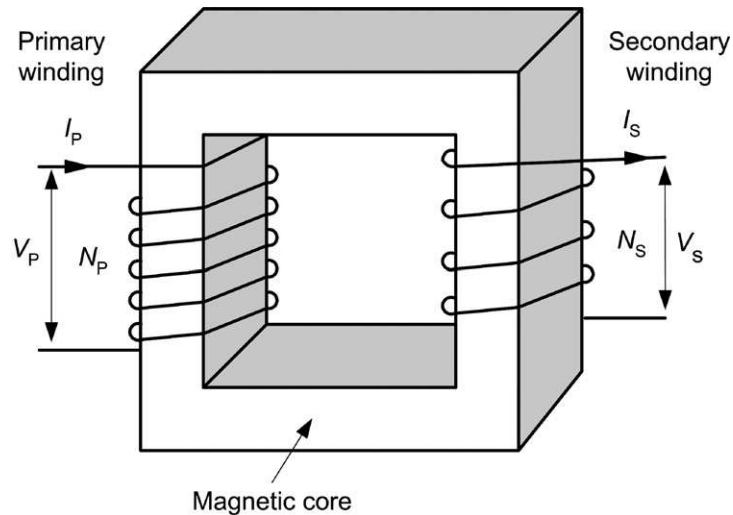


**Figure 3.3.44** The principle of the transformer

At any instant the flux,  $\Phi$ , in the transformer core is given by the equation:

$$\Phi = \Phi_{\max} \sin(\omega t)$$

where  $\Phi_{\max}$  is the maximum value of flux (in Wb),  $f$  is the frequency of the applied current, and  $t$  is the time in s.

The rms value of the primary voltage ( $V_P$ ) is given by:

$$V_P = 4.44 f N_P \Phi_{\max}$$

Similarly, the rms value of the secondary voltage ( $V_S$ ) is given by:

$$V_S = 4.44 f N_S \Phi_{\max}$$

From these two relationships (and since the same magnetic flux appears in both the primary and secondary windings) we can infer that:

$$\frac{V_P}{V_S} = \frac{N_P}{N_S}$$

where  $V_P$  and  $V_S$  are the primary and secondary voltages respectively and  $N_P$  and  $N_S$  are the primary and secondary turns, respectively (Figure 3.3.45).

Furthermore, assuming that no power is lost in the transformer (i.e. as long as the primary and secondary powers are the same) we can conclude that:

$$\frac{I_P}{I_S} = \frac{N_S}{N_P}$$

where  $I_S$  and  $I_P$  are the secondary and primary currents respectively and  $N_P$  and  $N_S$  are the primary and secondary turns respectively.

The ratio of primary turns to secondary turns ( $N_P/N_S$ ) is known as the *turns ratio*.

Furthermore, since ratio of primary voltage to primary turns is the same as the ratio of secondary turns to secondary voltage, we can conclude that, for a particular transformer:

$$\text{Turns-per-volt (t.p.v.)} = \frac{V_P}{N_P} = \frac{V_S}{N_S}.$$

The turns-per-volt rating can be quite useful when it comes to designing transformers with multiple secondary windings.

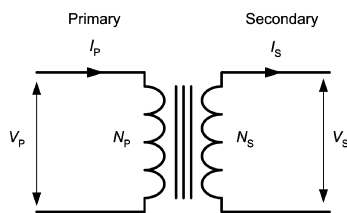


**Figure 3.3.45** Transformer turns and voltages

**Example 3.3.33**

A transformer has 2000 primary turns and 120 secondary turns. If the primary is connected to a 220 V AC mains supply, determine the secondary voltage.

Since  $\frac{V_P}{V_S} = \frac{N_P}{N_S}$  we can conclude that:

$$V_S = \frac{V_P N_S}{N_P} = \frac{220 \times 120}{2000} = \mathbf{13.2 \text{ V.}}$$

**Example 3.3.34**

A transformer has 1200 primary turns and is designed to operated with a 110 V AC supply. If the transformer is required to produce an output of 10 V, determine the number of secondary turns required.

Since  $\frac{V_P}{V_S} = \frac{N_P}{N_S}$  we can conclude that:

$$N_S = \frac{N_P V_S}{V_P} = \frac{1200 \times 10}{110} = \mathbf{109.1.}$$

**Example 3.3.35**

A transformer has a turns-per-volt rating of 1.5. How many turns are required to produce secondary outputs of (a) 50 V and (b) 350 V?

Here we will use  $N_S = \text{turns-per-volt} \times V_S$

(a) In the case of a 50 V secondary winding:

$$N_S = 1.5 \times 50 = 75 \text{ turns.}$$

(b) In the case of a 350 V secondary winding:

$$N_S = 1.5 \times 350 = \mathbf{525 \text{ turns.}}$$

**Example 3.3.36**

A transformer has 1200 primary turns and 60 secondary turns. Assuming that the transformer is loss-free, determine the primary current when a load current of 20 A is taken from the secondary.

Since  $\frac{I_S}{I_P} = \frac{N_P}{N_S}$  we can conclude that:

$$I_P = \frac{I_S N_S}{N_P} = \frac{20 \times 60}{1200} = \mathbf{1 \text{ A.}}$$

### Transformer applications

Transformers provide us with a means of coupling AC power from one circuit to another without a direct connection between the two. A further advantage of transformers is that voltage may be stepped-up (secondary voltage greater than primary voltage) or stepped-down (secondary voltage less than primary voltage). Since no increase in power is possible (like resistors capacitors and inductors, transformers are *passive* components) an increase in secondary voltage can only be achieved at the expense of a corresponding reduction in secondary current, and vice versa (in fact, the secondary power will be very slightly less than the primary power due to losses within the transformer).

Typical applications for transformers include stepping-up or stepping-down voltages in power supplies, coupling signals in AF amplifiers to achieve impedance matching and to isolate the DC potentials that may be present in certain types of circuit. The electrical characteristics of a transformer are determined by a number of factors including the core material and physical dimensions of the component. High frequency transformers (i.e. those designed for operation above about 10 kHz, or so) normally use ferrite as the core material whereas low frequency transformers (those designed for operation at frequencies below 10 kHz) use laminated steel cores.

The specifications for a transformer usually include the rated primary and secondary voltages and currents the required power rating (i.e. the rated power, usually expressed in volt-amperes, VA) which can be continuously delivered by the transformer under a given set of conditions), the frequency range for the component (usually stated as upper and lower working frequency limits), and the per-unit regulation of a transformer. As we shall see, this last specification is a measure of the ability of a transformer to maintain its rated output voltage under load.

### Transformer regulation and efficiency

The output voltage produced at the secondary of a real transformer the falls progressively as the load imposed on the transformer increases (i.e. as the secondary current increases from its no-load value). The *voltage regulation* of a transformer is a measure of its ability to keep the secondary output voltage constant over the full range of output load currents (i.e. from *unloaded* to *full-load*) at the same power factor. This change, when divided by the no-load output voltage, is referred to as the *per unit regulation* for the transformer. This can be best illustrated by the use of an example:

#### Example 3.3.37

A transformer produces an output voltage of 110 V under no-load conditions and an output voltage of 101 V when a full-load is applied. Determine the per-unit regulation.

The per-unit regulation can be determined for:

$$\begin{aligned}\text{Per-unit regulation} &= \frac{V_{S(\text{no-load})} - V_{S(\text{full-load})}}{V_{S(\text{no-load})}} \\ &= \frac{110 - 101}{110} = \mathbf{0.081} \text{ (or } 8.1\%).\end{aligned}$$

Most transformers operate with very high values of efficiency (typically between 85% and 95%). Despite this, in high power applications the losses in a transformer cannot be completely neglected. Transformer losses can be divided into two types of loss:

- losses in the magnetic core (often referred to as *iron loss*);
- losses due to the resistance of the coil windings (often referred to as *copper loss*).

Iron loss can be further divided into *hysteresis loss* (energy lost in repeatedly cycling the magnet flux in the core backwards and forwards) and *eddy current loss* (energy lost due to current circulating in the steel core).

Hysteresis loss can be reduced by using material for the magnetic core that is easily magnetised and has a very high permeability (see Figure 3.3.46, note that energy loss is proportional to the area *inside* the  $B$ – $H$  curve). Eddy current loss can be reduced by laminating the core (using E- and I-laminations, for example) and also ensuring that a small gap is present. These laminations and gaps in the core help to ensure that there is no closed path for current to flow. Copper loss results from the resistance of the coil windings and it can be reduced by using wire of large diameter and low resistivity.

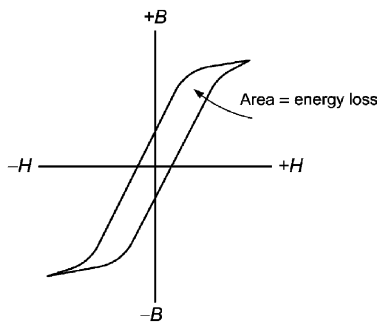


**Figure 3.3.46** Hysteresis curves and energy loss

It is important to note that, since the flux within a transformer varies only slightly between the no-load and full-load conditions, iron loss is substantially constant regardless of the load actually imposed on a transformer. Copper loss, on the other hand, is zero when a transformer is under no-load conditions and rises to a maximum at full-load.

The efficiency of a transformer is given by:

$$\text{Efficiency} = \frac{\text{output power}}{\text{input power}} \times 100\%$$

from which

$$\text{Efficiency} = \frac{\text{input power} - \text{losses}}{\text{input power}} \times 100\%$$

and

$$\text{Efficiency} = 1 - \frac{\text{losses}}{\text{input power}} \times 100\%.$$

As we have said, the losses present are attributable to iron loss and copper loss but the copper loss appears in both the primary and the secondary windings. Hence:

$$\text{Efficiency} = 1 - \frac{\text{iron loss} + \text{primary copper loss} + \text{secondary copper loss}}{\text{input power}} \times 100\%$$

### Example 3.3.38

A transformer rated at 500 VA has an iron loss of 3 W and a full-load copper loss (primary plus secondary) of 7 W. Calculate the efficiency of the transformer at 0.8 power factor.

The input power to the transformer will be given by the product of the apparent power (i.e. the transformer's VA rating) and the power factor.

Hence:

$$\text{Input power} = 0.8 \times 500 = 400 \text{ W.}$$

Now

$$\text{Efficiency} = 1 - \frac{7 + 3}{400} \times 100\% = \mathbf{97.5\%}.$$

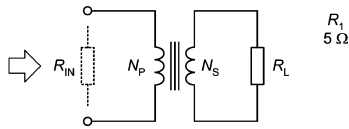


**Figure 3.3.47** Resistance seen looking into a transformer

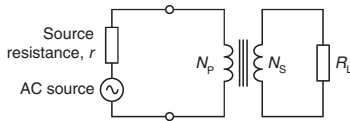


**Figure 3.3.48** Maximum power transfer

### Maximum power transfer

When a load is connected to the secondary winding of a transformer the resistance of the load is reflected by the transformer and appears as a resistance connected in the primary circuit, as shown in Figure 3.3.47.

The resistance,  $R_{IN}$ , seen looking into a loss-free transformer is given by:

$$R_{IN} = \left( \frac{N_P}{N_S} \right)^2 \times R_L$$

where  $N_P$  and  $N_S$  are the primary and secondary turns respectively and  $R_L$  is the load resistance.

For maximum power transfer between a source and a load the resistance of the source must be equal to the resistance of the load (this is known as the *Maximum Power Transfer Theorem*). Hence, in the case of the circuit shown in Figure 3.3.48, maximum power will be transferred to the load when:

$$r = \left( \frac{N_P}{N_S} \right)^2 \times R_L$$

where  $r$  is the resistance of the source.

### Example 3.3.39

A loss-free transformer has a turns ratio of 20:1 and is connected to a  $1.5 \Omega$  load. What resistance is seen at the input of the transformer?

$$\text{Now } R_{IN} = \left( \frac{N_P}{N_S} \right)^2 \times R_L = 20^2 \times 1.5 = 400 \times 1.5 = \mathbf{600 \Omega}.$$

### Example 3.3.40

A  $4 \Omega$  loudspeaker is to be matched to an amplifier that has an output impedance of  $5 \text{ k}\Omega$  using a transformer. What turns ratio is required to enable maximum power transfer from the amplifier to the speaker?

$$\text{Now } r = R_{IN} = \left( \frac{N_P}{N_S} \right)^2 R_L$$

from which:

$$\frac{N_P}{N_S} = \sqrt{\frac{r}{R_L}} = \sqrt{\frac{5 \times 10^3}{4}} = \mathbf{35.4}.$$

### Problems 3.3.4

1. A transformer has 480 primary turns and 120 secondary turns. If the primary is connected to a 110 V AC supply determine the secondary voltage.
2. A step-down transformer has a 220 V primary and a 24 V secondary. If the secondary winding has 60 turns, how many turns are there on the primary?
3. A transformer has 440 primary turns and 1800 secondary turns. If the secondary supplies a current of 250 mA, determine the primary current (assume that the transformer is loss-free).
4. A transformer produces an output voltage of 220 V under no-load conditions and an output voltage of 208 V when full-load is applied. Determine the per-unit regulation.
5. A 1 kVA transformer has an iron loss of 15 W and a full-load copper loss (primary plus secondary) of 20 W. Determine the efficiency of the transformer at 0.9 power factor.

### Alternating voltage and current

Unlike DC which have steady values and always flow in the same direction, AC flow alternately one way and then the other. The alternating p.d. (voltage) produced by an AC is thus partly positive and partly negative. An understanding of AC and voltages is important in a number of applications including AC power distribution, amplifiers and filters.

A graph showing the variation of voltage or current present in a circuit is known as a *waveform*. Some common types of waveform are shown in Figure 3.3.49. Note that, since the waveforms of speech and music comprise many components at different frequencies and of different amplitudes, these waveforms are referred to as *complex*.

The equation for the sinusoidal voltage shown in Figure 3.3.50, at a time  $t$ , is:

$$v = V_{\max} \sin(\omega t)$$

where  $v$  is the instantaneous voltage,  $V_{\max}$  is the maximum value of voltage (also known as the *amplitude* or *peak value* of voltage), and  $\omega$  is the angular velocity (in radians per second).

The frequency of a repetitive waveform is the number of cycles of the waveform which occur in unit time. Frequency is expressed in Hertz (Hz). A frequency of 1 Hz is equivalent to one cycle per second. Hence, if a voltage has a frequency of 400 Hz, 400 cycles will occur in every second.

Since there are  $2\pi$  radians in one complete revolution or *cycle*, a frequency of one cycle per second must be the same as  $2\pi$  radians per second. Hence a frequency,  $f$ , is equivalent to:

$$f = \omega/(2\pi)\text{Hz}.$$

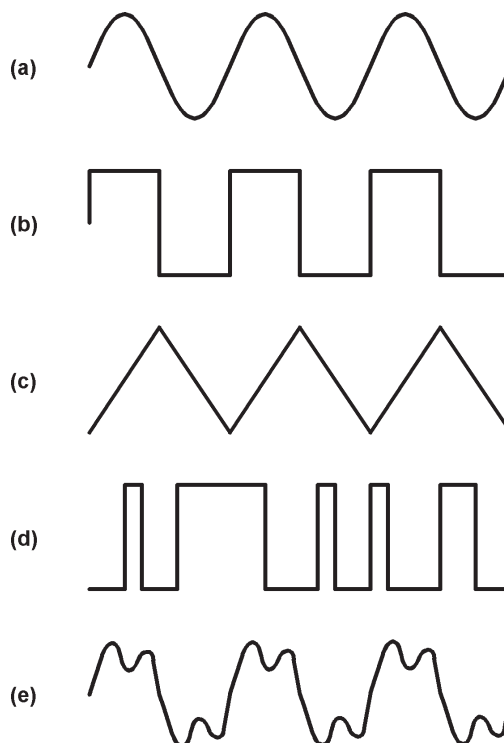


**Figure 3.3.49** Various waveforms: (a) sine, (b) square, (c) triangle, (d) pulse, and (e) complex

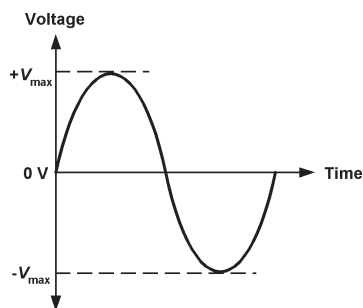


**Figure 3.3.50** One cycle of a sine wave

Alternatively, the angular velocity,  $\omega$ , is given by:

$$\omega = 2\pi f \text{ rad/s.}$$

We can thus express the instantaneous voltage in another way:

$$v = V_{\max} \sin(2\pi ft).$$

### Example 3.3.41

A sine wave voltage has a maximum value of 100 V and a frequency of 50 Hz. Determine the instantaneous voltage present (a) 2.5 ms and (b) 15 ms from the start of the cycle.

We can determine the voltage at any instant of time using:

$$v = V_{\max} \sin(2\pi ft)$$

where  $V_{\max} = 100\text{ V}$  and  $f = 50\text{ Hz}$ .

In (a),  $t = 2.5\text{ ms}$  hence:

$$\begin{aligned} v &= 100 \sin(2\pi \times 50 \times 0.0025) = 100 \sin(0.785) \\ &= 100 \times 0.707 = \mathbf{70.7\text{ V.}} \end{aligned}$$

In (b),  $t = 15\text{ ms}$  hence:

$$\begin{aligned} v &= 100 \sin(2\pi \times 50 \times 0.015) = 100 \sin(4.71) \\ &= 100 \times -1 = \mathbf{-100\text{ V.}} \end{aligned}$$

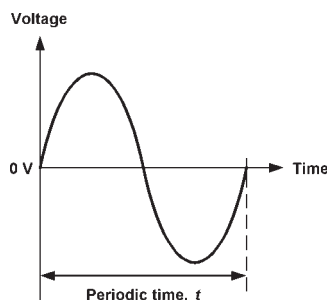


Figure 3.3.51 Periodic time

### Periodic time

The periodic time of a waveform is the time taken for one complete cycle of the wave (Figure 3.3.51). The relationship between periodic time and frequency is thus:

$$t = 1/f \quad \text{or} \quad f = 1/t$$

where  $t$  is the periodic time (in s) and  $f$  is the frequency (in Hz).

#### Example 3.3.42

A waveform has a frequency of 200 Hz. What is the periodic time of the waveform?

$$t = 1/f = 1/200 = \mathbf{0.005 \text{ s}} \text{ (or 5 ms).}$$

#### Example 3.3.43

A waveform has a periodic time of 2.5 ms. What is its frequency?

$$f = 1/t = 1/(2.5 \times 10^{-3}) = 1 \times 10^3/2.5 = \mathbf{400 \text{ Hz.}}$$

### Average, peak, peak-peak, and rms values

The *average* value of an AC which swings symmetrically above and below zero will obviously be zero when measured over a long period of time. Hence average values of currents and voltages are invariably taken over one complete half-cycle (either positive or negative) rather than over one complete full-cycle (which would result in an average value of zero).

The *peak* value (or *amplitude*) of a waveform is a measure of the extent of its voltage or current excursion from the resting value (usually zero). The *peak-to-peak* value for a wave which is symmetrical about its resting value is twice its peak value.

The *rms* (or *effective*) value of an alternating voltage or current is the value which would produce the same heat energy in a resistor as a direct voltage or current of the same magnitude. Since the rms value of a waveform is very much dependent upon its shape, values are only meaningful when dealing with a waveform of known shape. Where the shape of a waveform is not specified, rms values are normally assumed to refer to sinusoidal conditions (Figure 3.3.52).

The following formulae apply to a sine wave:

$$V_{\text{average}} = 0.636 \times V_{\text{peak}}$$

$$V_{\text{peak-peak}} = 2 \times V_{\text{peak}}$$

$$V_{\text{rms}} = 0.707 \times V_{\text{peak}}$$

Similar relationships apply to the corresponding AC, thus:

$$I_{\text{average}} = 0.636 \times I_{\text{peak}}$$

$$I_{\text{peak-peak}} = 2 \times I_{\text{peak}}$$

$$I_{\text{rms}} = 0.707 \times I_{\text{peak}}$$

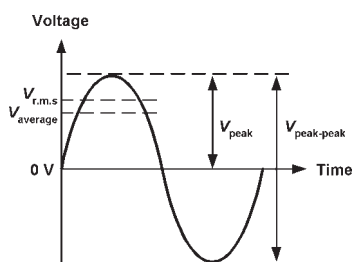


Figure 3.3.52 Relationship between average, peak, peak-peak, and rms values



**Example 3.3.44**

A sinusoidal voltage has an rms value of 220 V. What is the peak value of the voltage?

Since  $V_{\text{rms}} = 0.707 V_{\text{peak}}$ ,  $V_{\text{peak}} = V_{\text{rms}}/0.707 = 1.414 \times V_{\text{rms}}$ .

Thus  $V_{\text{peak}} = 1.414 \times 220 = \mathbf{311 \text{ V}}$ .

**Example 3.3.45**

A sinusoidal AC has a peak-peak value of 4 mA. What is its rms value?

First we must convert the peak-peak current into peak current:

Since  $I_{\text{peak-peak}} = 2 \times I_{\text{peak}}$ ,  $I_{\text{peak}} = 0.5 \times I_{\text{peak-peak}}$

Thus  $I_{\text{peak}} = 0.5 \times 4 = 2 \text{ mA}$ .

Now we can convert the peak current into rms current using:

$I_{\text{rms}} = 0.707 \times I_{\text{peak}}$ .

Thus  $I_{\text{rms}} = 0.707 \times 2 = \mathbf{1.414 \text{ mA}}$ .

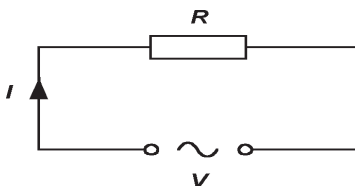


**Figure 3.3.53** Alternating current in a resistor

**AC in a resistor**

Ohm's law is obeyed in an AC circuit just as it is in a DC circuit. Thus, when a sinusoidal voltage,  $V$ , is applied to a resistor,  $R$ , (as shown in Figure 3.3.53) the current flowing in the resistor will be given by:

$$I = V/R.$$

This relationship must also hold true for the instantaneous values of current,  $i$ , and voltage,  $v$ , thus:

$$i = v/R$$

and since  $v = V_{\text{max}} \sin(\omega t)$

$$i = V_{\text{max}} \sin(\omega t)/R.$$



**Figure 3.3.54** Phasor diagram for the circuit in Figure 3.3.53

The current and voltage will both have a sinusoidal shape and since they rise and fall together, they are said to be *in-phase* with one another. We can represent this relationship by means of the *phasor diagram* shown in Figure 3.3.54. This diagram shows two rotating phasors (of magnitude  $I$  and  $V$ ) rotating at an angular velocity,  $\omega$ . The applied voltage ( $V$ ) is referred to as the *reference phasor* and this is aligned with the horizontal axis (i.e. it has a phase angle of  $0^\circ$ ).

## Another view

Phasor diagrams provide us with a quick way of illustrating the relationships that exist between sinusoidal voltages and currents in AC circuits without having to draw lots of time-related waveforms. Figure 3.3.55 will help you to understand how the previous phasor diagram relates to time-related waveforms for the voltage and current in a resistor.

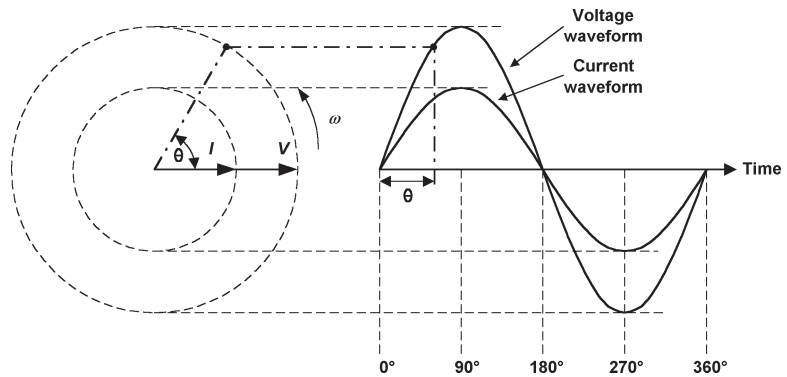


Figure 3.3.55 Phasor diagram and time-related waveforms

### Example 3.3.46

A sinusoidal voltage  $10 V_{\text{peak-peak}}$  is applied to a resistor of  $1 \text{ k}\Omega$ . What value of rms current will flow in the resistor?

This problem must be solved in several stages. First we will determine the peak-peak current in the resistor and then we will convert this value into a corresponding rms quantity.

Since:  $I = V/R$ ,  $I_{\text{peak-peak}} = V_{\text{peak-peak}}/R$ .

Thus:  $I_{\text{peak-peak}} = 10 V_{\text{peak-peak}}/1 \text{ k}\Omega = 10 \text{ mA}$ .

Next:  $I_{\text{peak}} = I_{\text{peak-peak}}/2 = 10/2 = 5 \text{ mA}$ .

Finally:  $I_{\text{rms}} = 0.707 \times I_{\text{peak}} = 0.707 \times 5 \text{ mA} = \mathbf{3.53 \text{ mA}}$ .

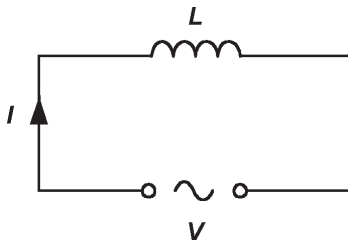


Figure 3.3.56 Alternating current in an inductor

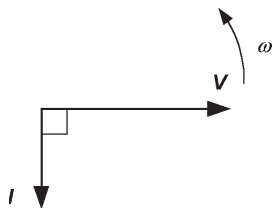


Figure 3.3.57 Phasor diagram for the circuit in Figure 3.3.56

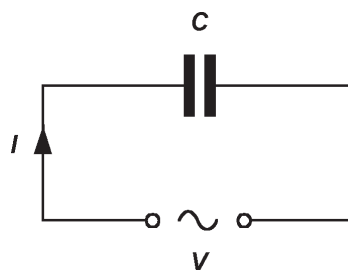


Figure 3.3.58 Alternating current in a capacitor

### AC in an inductor

When a sinusoidal voltage,  $V$ , is applied to an inductor,  $L$ , (as shown in Figure 3.3.56) the current flowing in the inductor will be given by:

$$I = V/X$$

where  $X$  is the *reactance* of the inductor. Like resistance, reactance is measured in ohms ( $\Omega$ ).

The current in an inductor lags behind the voltage by a *phase angle* of  $90^\circ$  ( $\pi/2$  rad) and since  $v = V_{\text{max}} \sin(\omega t)$

$$i = V_{\text{max}} \sin(\omega t - \pi/2)/X.$$

Once again, the current and voltage will both have a sinusoidal shape but they are  $90^\circ$  apart and this relationship has been illustrated by means of the phasor diagram shown in Figure 3.3.57. The applied voltage ( $V$ ) is the *reference phasor* (its phase angle is  $0^\circ$ ) whilst the current flowing ( $I$ ) has a *lagging phase angle* of  $90^\circ$ .

### AC in a capacitor

When a sinusoidal voltage,  $V$ , is applied to a capacitor,  $C$ , (as shown in Figure 3.3.58) the current flowing in the capacitor will be given by:

$$I = V/X$$

where  $X$  is the *reactance* of the capacitor. Like resistance, reactance is measured in ohms ( $\Omega$ ).



**Figure 3.3.59** Phasor diagram for the circuit in Figure 3.3.58

The current in a capacitor leads the applied voltage by a *phase angle* of  $90^\circ$  ( $\pi/2$  rad) and since  $v = V_{\max} \sin(\omega t)$

$$i = V_{\max} \sin(\omega t + \pi/2)/X.$$

As before, the current and voltage both have a sinusoidal shape  $90^\circ$  apart and this relationship is illustrated in the phasor diagram shown in Figure 3.3.59. The applied voltage ( $V$ ) is the *reference phasor* (its phase angle of  $0^\circ$ ) whilst the current flowing ( $I$ ) has a *leading phase angle* of  $90^\circ$ .

### Reactance

When alternating voltages are applied to capacitors or inductors the amount of current flowing will depend upon the value of capacitance or inductance and on the frequency of the voltage. In effect, capacitors and inductors oppose the flow of current in much the same way as a resistor. The important difference being that the effective resistance (or *reactance*) of the component varies with frequency (unlike the case of a pure resistance where the magnitude of the current *does not* change with frequency).

Reactance, like resistance, is simply the ratio of applied voltage to the current flowing. Thus:

$$X = V/I$$

where  $X$  is the reactance in ohms ( $\Omega$ ),  $V$  is the alternating p.d. in volts ( $V$ ) and  $I$  is the AC in A.

In the case of *capacitive reactance* (i.e. the reactance of a capacitor) we use the suffix,  $C$ , so that the reactance equation becomes:

$$X_C = V_C/I_C.$$

Similarly, in the case of *inductive reactance* (i.e. the reactance of an inductor) we use the suffix,  $L$ , so that the reactance equation becomes:

$$X_L = V_L/I_L.$$

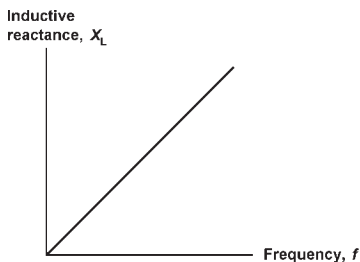


**Figure 3.3.60** Variation of inductive reactance with frequency

### Inductive reactance

Inductive reactance is directly proportional to the frequency of the applied AC and can be determined from the following formula:

$$X_L = 2\pi fL$$

where  $X_L$  is the reactance in  $\Omega$ ,  $f$  is the frequency in Hz, and  $L$  is the inductance in H.

Since inductive reactance is directly proportional to frequency ( $X_L \propto f$ ), the graph of inductive reactance plotted against frequency takes the form of a straight line (see Figure 3.3.60).

### Example 3.3.47

Determine the reactance of a 10 mH inductor at (a) 100 Hz and (b) at 10 kHz.

(a) At 100 Hz,  $X_L = 2\pi \times 100 \times 10 \times 10^{-3} = \mathbf{6.28 \Omega}$ .

(b) At 10 kHz,  $X_L = 2\pi \times 10\,000 \times 10 \times 10^{-3} = \mathbf{628 \Omega}$ .

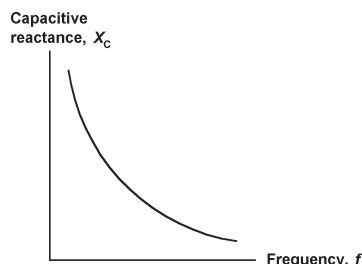


**Figure 3.3.61** Variation of capacitive reactance with frequency

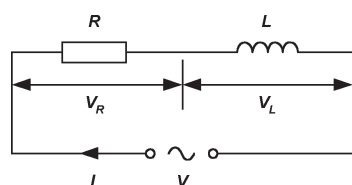


**Figure 3.3.62** Resistance and inductance in series

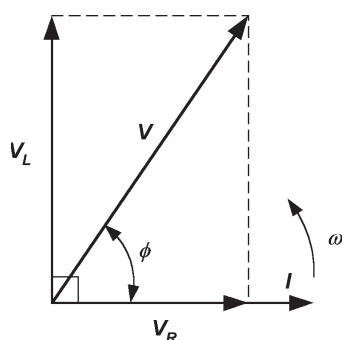


**Figure 3.3.63** Phasor diagram for the circuit in Figure 3.3.62

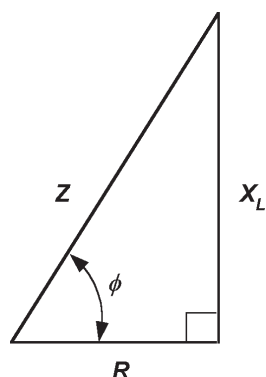


**Figure 3.3.64** Impedance triangle for the circuit in Figure 3.3.62

## Capacitive reactance

Capacitive reactance is inversely proportional to the frequency of the applied AC and can be determined from the following formula:

$$X_C = 1/2\pi fC$$

where  $X_C$  is the reactance in  $\Omega$ ,  $f$  is the frequency in Hz, and  $C$  is the capacitance in F.

Since capacitive reactance is inversely proportional to frequency ( $X_C \propto 1/f$ ), the graph of inductive reactance plotted against frequency takes the form of a rectangular hyperbola (see Figure 3.3.61).

### Example 3.3.48

Determine the reactance of a  $1\mu\text{F}$  capacitor at (a) 100 Hz and (b) 10 kHz.

- (a) At 100 Hz,  $X_C = 1/(2\pi \times 100 \times 1 \times 10^{-6})$  thus  $X_C = 0.159/10^{-4} = 0.159 \times 10^4 = \mathbf{1.59\text{ k}\Omega}$ .  
 (b) At 10 kHz,  $X_C = 1/(2\pi \times 10\,000 \times 1 \times 10^{-6})$  thus  $X_C = 0.159/10^{-2} = 0.159 \times 10^2 = \mathbf{15.9\ \Omega}$ .

## Resistance and inductance in series

When a sinusoidal voltage,  $V$ , is applied to a series circuit comprising resistance,  $R$ , and inductance,  $L$ , (as shown in Figure 3.3.62) the current flowing in the circuit will produce separate voltage drops across the resistor and inductor ( $V_R$  and  $V_L$ , respectively). These two voltage drops will be  $90^\circ$  apart – with  $V_L$  leading  $V_R$ . We can illustrate this relationship using the phasor diagram shown in Figure 3.3.63. Note that we have used current as the reference phasor in this series circuit for the simple reason that the same current flows through each component (recall that earlier we used the applied voltage as the reference).

From Figure 3.3.63 you should note that the supply voltage ( $V$ ) is simply the result of adding the two voltage phasors,  $V_R$  and  $V_L$ . Furthermore, the angle between the supply voltage ( $V$ ) and supply current ( $I$ ),  $\phi$ , is known as the *phase angle*.

Now  $\sin \phi = V_L/V$ ,  $\cos \phi = V_R/V_L$ , and  $\tan \phi = V_L/V_R$ .

Since  $X_L = V_L/I$ ,  $R = V_R/I$ , and  $Z = V/I$  (where  $Z$  is the *impedance* of the circuit), we can illustrate the relationship between  $X_L$ ,  $R$ , and  $Z$  using the *impedance triangle* shown in Figure 3.3.64.

Note that  $Z = \sqrt{R^2 + X_L^2}$  and  $\phi = \arctan(X_L/R)$ .

### Example 3.3.49

An inductor of 80 mH is connected in series with a  $100\ \Omega$  resistor. If a sinusoidal current of 20 mA at 50 Hz flows in the circuit, determine:

- the voltage dropped across the inductor;
- the voltage dropped across the resistor;
- the impedance of the circuit;

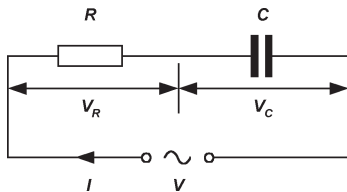


**Figure 3.3.65** Resistance and capacitance in series

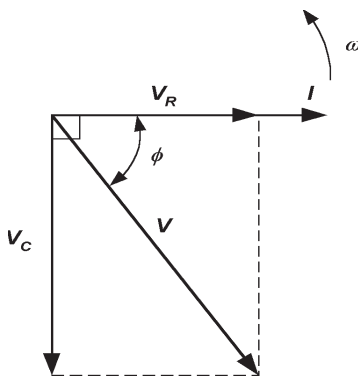


**Figure 3.3.66** Phasor diagram for the circuit in Figure 3.3.65

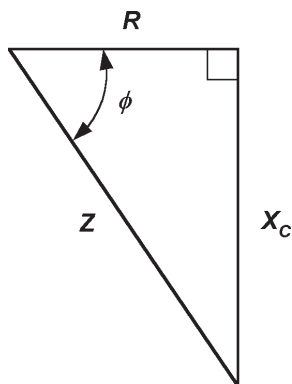


**Figure 3.3.67** Impedance triangle for the circuit in Figure 3.3.65

- (d) the supply voltage;
- (e) the phase angle.

(a)  $V_L = I X_L = I \times 2\pi f L = 0.02 \times 25.12 = 0.5 \text{ V}$ .

(b)  $V_R = I R = 0.02 \times 100 = 2 \text{ V}$ .

(c)  $Z = \sqrt{R^2 + X_L^2} = \sqrt{100^2 + 25.12^2}$   
 $= \sqrt{10\,631} = 103.1 \, \Omega$ .

(d)  $V = I \times Z = 0.02 \times 103.1 = 2.06 \text{ V}$ .

(e)  $\phi = \arctan(X_L/R) = \arctan(25.12/100) = \arctan(0.2512)$   
 $= 14.1^\circ$ .

### Resistance and capacitance in series

When a sinusoidal voltage,  $V$ , is applied to a series circuit comprising resistance,  $R$ , and inductance,  $L$ , (as shown in Figure 3.3.65) the current flowing in the circuit will produce separate voltage drops across the resistor and capacitor ( $V_R$  and  $V_C$ , respectively). These two voltage drops will be  $90^\circ$  apart – with  $V_C$  lagging  $V_R$ . We can illustrate this relationship using the phasor diagram shown in Figure 3.3.66. Note that once again we have used current as the reference phasor in this series circuit.

From Figure 3.3.66 you should note that the supply voltage ( $V$ ) is simply the result of adding the two voltage phasors,  $V_R$  and  $V_C$ . Furthermore, the angle between the supply voltage ( $V$ ) and supply current ( $I$ ),  $\phi$ , is known as the *phase angle*.

Now  $\sin \phi = V_C/V$ ,  $\cos \phi = V_R/V$ , and  $\tan \phi = V_C/V_R$ .

Since  $X_L = V_C/I$ ,  $R = V_R/I$ , and  $Z = V/I$  (where  $Z$  is the *impedance* of the circuit), we can illustrate the relationship between  $X_C$ ,  $R$ , and  $Z$  using the *impedance triangle* shown in Figure 3.3.67.

Note that  $Z = \sqrt{R^2 + X_C^2}$  and  $\phi = \arctan(X_C/R)$ .

### Example 3.3.50

A capacitor of  $22 \, \mu\text{F}$  is connected in series with a  $470 \, \Omega$  resistor. If a sinusoidal current of  $10 \text{ mA}$  at  $50 \text{ Hz}$  flows in the circuit, determine:

- (a) the voltage dropped across the capacitor;
- (b) the voltage dropped across the resistor;
- (c) the impedance of the circuit;
- (d) the supply voltage;
- (e) the phase angle.

(a)  $V_C = I X_C = I \times 1/(2\pi f C) = 0.01 \times 144.5 = 1.4 \text{ V}$ .

(b)  $V_R = I R = 0.01 \times 470 = 4.7 \text{ V}$ .

(c)  $Z = \sqrt{R^2 + X_C^2} = \sqrt{470^2 + 144.5^2}$   
 $= \sqrt{241\,780} = 491.7 \, \Omega$ .

(d)  $V = I \times Z = 0.01 \times 491.7 = 4.91 \text{ V}$ .

(e)  $\phi = \arctan(X_C/R) = \arctan(144.5/470)$   
 $= \arctan(0.3074) = 17.1^\circ$ .

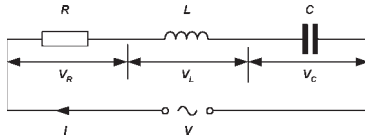


**Figure 3.3.68** Resistance, inductance, and capacitance in series

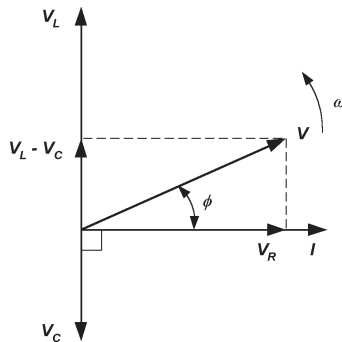


**Figure 3.3.69** Phasor diagram for the circuit in Figure 3.3.68 when  $X_L > X_C$

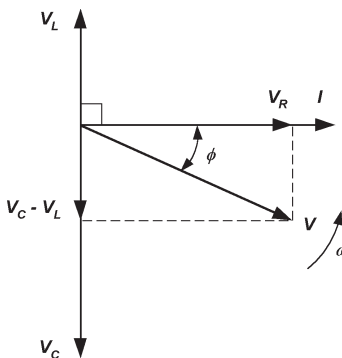


**Figure 3.3.70** Phasor diagram for the circuit in Figure 3.3.68 when  $X_L < X_C$

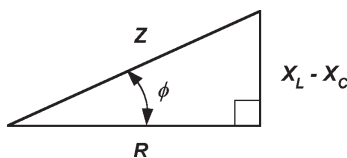


**Figure 3.3.71** Impedance triangle for the circuit in Figure 3.3.68 when  $X_L > X_C$

## Resistance, inductance, and capacitance in series

When a sinusoidal voltage,  $V$ , is applied to a series circuit comprising resistance,  $R$ , inductance,  $L$ , and capacitance,  $C$ , (as shown in Figure 3.3.68) the current flowing in the circuit will produce separate voltage drops across the resistor, inductor and capacitor ( $V_R$ ,  $V_L$ , and  $V_C$ , respectively). The voltage drop across the inductor will lead the applied current (and voltage dropped across  $V_R$ ) by  $90^\circ$  whilst the voltage drop across the capacitor will lag the applied current (and voltage dropped across  $V_R$ ) by  $90^\circ$ .

When the inductive reactance ( $X_L$ ) is greater than the capacitive reactance ( $X_C$ ),  $V_L$  will be greater than  $V_C$  and the resulting phasor diagram is shown in Figure 3.3.69.

Conversely, when the capacitive reactance ( $X_C$ ) is greater than the inductive reactance ( $X_L$ ),  $V_C$  will be greater than  $V_L$  and the resulting phasor diagram is shown in Figure 3.3.70.

Note that once again we have used current as the *reference phasor* in this series circuit.

From Figures 3.3.69 and 3.3.70, you should note that the supply voltage ( $V$ ) is simply the result of adding the three voltage phasors,  $V_L$ ,  $V_C$ , and  $V_R$  and that the first stage in simplifying the diagram is that of resolving  $V_L$  and  $V_C$  into a single voltage ( $V_L - V_C$  or  $V_C - V_L$  depending upon whichever is the greater). Once again, the phase angle,  $\phi$ , is the angle between the supply voltage and current.

Figures 3.3.71 and 3.3.72 show the impedance triangle for the circuit, for the cases when  $X_L > X_C$  and  $X_C > X_L$ , respectively.

Note that, when  $X_L > X_C$ ,  $Z = \sqrt{R^2 + (X_L - X_C)^2}$  and  $\phi = \arctan(X_L - X_C)/R$ ; similarly, when  $X_C > X_L$ ,  $Z = \sqrt{R^2 + (X_C - X_L)^2}$  and  $\phi = \arctan(X_C - X_L)/R$ .

### Example 3.3.51

A series circuit comprises an inductor of 80 mH, a resistor of  $200\ \Omega$  and a capacitor of  $22\ \mu\text{F}$ . If a sinusoidal current of 40 mA at 50 Hz flows in this circuit, determine:

- the voltage developed across the inductor;
- the voltage dropped across the capacitor;
- the voltage dropped across the resistor;
- the impedance of the circuit;
- the supply voltage;
- the phase angle.

(a)  $V_L = IX_L = I \times 2\pi fL = 0.04 \times 25.12 = 1\text{ V}$ .

(b)  $V_C = IX_C = I \times 1/(2\pi fC) = 0.04 \times 144.5 = 5.8\text{ V}$ .

(c)  $V_R = IR = 0.04 \times 200 = 8\text{ V}$ .

(d)  $Z = \sqrt{R^2 + (X_C - X_L)^2} = \sqrt{200^2 + (144.5 - 25.12)^2}$   
 $= \sqrt{54\ 252}$   
 $= 232.9\ \Omega$ .

(e)  $V = I \times Z = 0.04 \times 232.9 = 9.32\text{ V}$ .

(f)  $\phi = \arctan(X_C - X_L)/R = \arctan(119.38/200)$   
 $= \arctan(0.597) = 30.8^\circ$ .

### Impedance

Circuits that contain a mixture of both resistance and reactance (of either or both types) are said to exhibit *impedance*. Impedance, like resistance and reactance, is simply the ratio of applied voltage to the current flowing. Thus:

$$Z = V/I$$

where  $Z$  is the impedance in ohms ( $\Omega$ ),  $V$  is the alternating p.d. in V and  $I$  is the AC in A.

The impedance of a series circuit ( $R$  in series with  $X$  – see Figure 3.3.73) is given by:

$$Z = \sqrt{(R^2 + X^2)}$$

where  $Z$  is the impedance (in  $\Omega$ ),  $X$  is the reactance, either capacitive or inductive (expressed in  $\Omega$ ), and  $R$  is the resistance (in  $\Omega$ ).

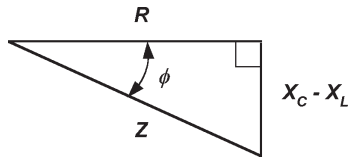


**Figure 3.3.72** Impedance triangle for the circuit in Figure 3.3.68 when  $X_L < X_C$

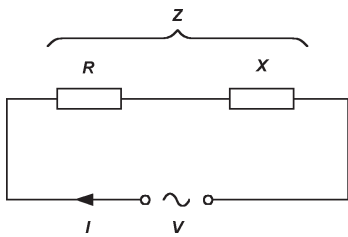


**Figure 3.3.73** A series circuit containing reactance and resistance

#### Example 3.3.52

A resistor of  $30\ \Omega$  is connected in series with a capacitive reactance of  $40\ \Omega$ . Determine the impedance of the circuit and the current flowing when the circuit is connected to a 115 V supply.

First we must find the impedance of the  $C$ – $R$  series circuit:

$$Z = \sqrt{(R^2 + X_C^2)} = \sqrt{(30^2 + 40^2)} = \sqrt{2500} = 50.$$

The current taken from the supply can now be found:

$$I = V/Z = 115/50 = \mathbf{2.3\ A}.$$

#### Example 3.3.53

A coil is connected to a 50 V AC supply at 400 Hz. If the current supplied to the coil is 200 mA and the coil has a resistance of  $60\ \Omega$ , determine the value of inductance.

We can find the impedance of the coil from:

$$Z = V/I = 50/0.2 = 250\ \Omega$$

Since

$$Z = \sqrt{(R^2 + X_L^2)}, Z^2 = (R^2 + X_L^2) \text{ and } X_L^2 = Z^2 - R^2$$

$$X_L^2 = Z^2 - R^2 = 250^2 - 60^2 = 62500 - 3600 = 58900.$$

$$\text{Thus } X_L = \sqrt{58900} = 243\ \Omega.$$

$$\text{Since } X_L = 2\pi fL, L = X_L/(2\pi f) = 243/(6.28 \times 400) = 0.097\ \text{H}.$$

Hence  $L = \mathbf{97\ \text{mH}}$ .

## Problems 3.3.5

- (1) A sinusoidal AC is specified by the equation,  $i = 250 \sin(628 t)$  mA. Determine:
  - (a) the peak value of the current;
  - (b) the rms value of the current;
  - (c) the frequency of the current;
  - (d) the periodic time of the current;
  - (e) the instantaneous value of the current at  $t = 2$  ms.
- (2) A sinusoidal alternating voltage has a peak value of 160 V and a frequency of 60 Hz. Write down an expression for the instantaneous voltage and use it to determine the value of voltage at 3 ms from the start of a cycle.
- (3) A resistance of  $45 \Omega$  is connected to a 150 V 400 Hz AC supply. Determine:
  - (a) an expression for the instantaneous current flowing in the resistor;
  - (b) the rms value of the current flowing in the resistor.
- (4) Determine the reactance of a 680 nF capacitor at (a) 400 Hz and (b) 20 kHz.
- (5) Determine the reactance of a 60 mH inductor at (a) 20 Hz and (b) at 4 kHz.
- (6) A resistor of  $120 \Omega$  is connected in series with a capacitive reactance of  $200 \Omega$ . Determine the impedance of the circuit and the current flowing when the circuit is connected to a 200 V AC supply.
- (7) A coil has an inductance of 200 mH and a resistance of  $40 \Omega$ . Determine:
  - (a) the impedance of the coil at a frequency of 60 Hz;
  - (b) the current that will flow in the coil when it is connected to a 110 V 60 Hz supply.
- (8) A capacitor of  $2 \mu\text{F}$  is connected in series with a  $100 \Omega$  resistor across a 24 V 400 Hz AC supply. Determine the current that will be supplied to the circuit and the voltage that will be dropped across each component.
- (9) A capacitor of 100 nF is used in a power line filter. Determine the frequency at which the reactance of the capacitor is equal to  $10 \Omega$ .
- (10) An inductor is used in an aerial tuning unit. Determine the value of inductance required if the inductor is to have a reactance of  $300 \Omega$  at a frequency of 3.5 MHz.

### Questions 3.3.1

A voltage,  $v = 17 \sin(314 t)$  volts, is applied to a pure resistive load of  $68 \Omega$ .

- (a) Sketch a fully labelled graph showing how the current flowing in the resistor varies over the period from 0 to 30 ms.
- (b) Mark the following on your graph and values for each:
  - (i) the peak value of current;
  - (ii) the periodic time.
- (c) Determine the instantaneous values of voltage and current at  $t = 3.5$  ms.



### Another view

From studying complex numbers in mathematics, you will recall that every complex number consists of a *real part* and an *imaginary part*. In an electrical context, the *real part* is that part of the complex quantity that is in-phase with the reference quantity (current for series circuits and voltage for parallel circuits). The *imaginary part*, on the other hand, is that part of the complex quantity that is at  $90^\circ$  to the reference.

### Questions 3.3.2

A series circuit comprises an inductor of 60 mH, a resistor of  $33\ \Omega$  and a capacitor of  $47\ \mu\text{F}$ . If a sinusoidal current of 50 mA at 50 Hz flows in this circuit, determine:

- the voltage developed across the inductor;
- the voltage dropped across the capacitor;
- the voltage dropped across the resistor;
- the impedance of the circuit;
- the supply voltage;
- the phase angle.

Sketch a phasor diagram showing the voltages and current present in the circuit. Label your drawing clearly and indicate values.

## Using complex notation

Complex notation provides us with a simple yet powerful method of solving even the most complex of AC circuits. Complex notation allows us to represent electrical quantities that have both *magnitude* and *direction* (you will already know that in other contexts we call these *vectors*). The *magnitude* is simply the amount of resistance, reactance, voltage or current, etc. In order to specify the *direction* of the quantity, we use an *operator* to denote the phase shift relative to the *reference* quantity (this is usually current for a series circuit and voltage for a parallel circuit). We call this operator 'j'.

## Mathematics in action

### The j-operator

You can think of the j-operator as a device that allows us to indicate a phase shift of  $90^\circ$ . A phase shift of  $+90^\circ$  is represented by  $+j$  whilst a phase shift of  $-90^\circ$  is represented by  $-j$ .

A *phasor* is simply an electrical vector. A *vector*, as you will doubtless recall, has magnitude (size) and direction (angle relative to some reference direction). The j-operator can be used to rotate a phasor. Each successive multiplication by j has the effect of rotating the phasor through a further  $90^\circ$ .

The j-operator has a value which is equal to  $\sqrt{-1}$ . Thus we can conclude that:  $j = \sqrt{-1}$

multiplying by j gives,  $j^2 = \sqrt{-1} \times \sqrt{-1} = -1$

multiplying again by j gives,  $j^3 = \sqrt{-1} \times \sqrt{-1} \times \sqrt{-1} = -1 \times j = -j$

multiplying again by j gives,  $j^4 = \sqrt{-1} \times \sqrt{-1} \times \sqrt{-1} \times \sqrt{-1}$   
 $= -1 \times -1 = +1$

multiplying again by j gives,  $j^5 = \sqrt{-1} \times \sqrt{-1} \times \sqrt{-1} \times \sqrt{-1} \times \sqrt{-1}$   
 $= -1 \times -1 \times j = +j$

and so on.

## Mathematics in action

### *The Argand diagram*

The Argand diagram provides a useful method of illustrating complex quantities and allowing us to solve problems graphically. In common with any ordinary 'x-y' graph, the Argand diagram has two sets of axes at right angles. The horizontal axis is known as the *real axis* whilst the vertical axis is known as the *imaginary axis* (do not panic – the imaginary axis is not really imaginary we simply use the term to indicate that we are using this axis to plot values that are multiples of the j-operator).

### Complex impedances

The j-operator and the Argand diagram provide us with a useful way of representing impedances. Any complex impedance can be represented by the relationship:

$$Z = (R \pm jX)$$

where  $Z$  represents impedance,  $R$  represents resistance, and  $X$  represents reactance.

All three quantities are, of course, measured in ohms ( $\Omega$ ).

The  $\pm j$  term simply allows us to indicate whether the reactance is due to inductance, in which case the j term is positive (i.e.  $+j$ ) or whether it is due to capacitance, in which case the j term is negative (i.e.  $-j$ ).

Consider, for example, the following impedances:

1.  $Z_1 = 20 + j10$  this impedance comprises a resistance of  $20\ \Omega$  connected in series with an inductive reactance (note the positive sign before the j term) of  $10\ \Omega$ .
2.  $Z_2 = 15 - j25$  this impedance comprises a resistance of  $15\ \Omega$  connected in series with a capacitive reactance (note the negative sign before the j term) of  $25\ \Omega$ .
3.  $Z_3 = 30 + j0$  this impedance comprises a pure resistance of  $30\ \Omega$  (there is no reactive component).

These three impedances are shown plotted on an Argand diagram in Figure 3.3.74.

Voltages and currents can also take complex values. Consider the following:

1.  $I_1 = 2 + j0.5$  this current is the result of an in-phase component of 2 A and a reactive component (at  $+90^\circ$ ) of 0.5 A.
2.  $I_2 = 1 - j1.5$  this current is the result of an in-phase component of 1 A and a reactive component (at  $-90^\circ$ ) of 1.5 A.
3.  $I_3 = 3 + j0$  this current in-phase and has a value of 3 A.

These three currents are shown plotted on an Argand diagram in Figure 3.3.75.

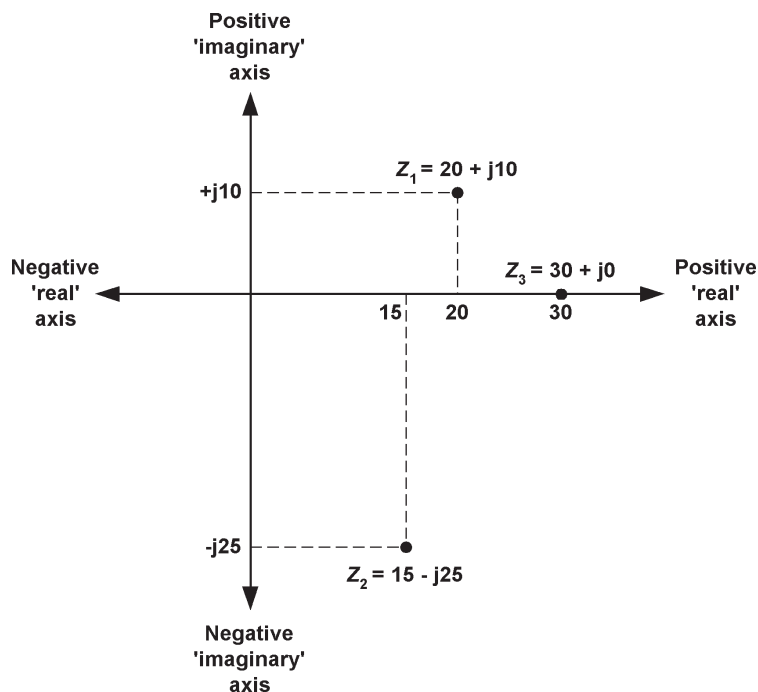


**Figure 3.3.74** Argand diagram showing complex impedances

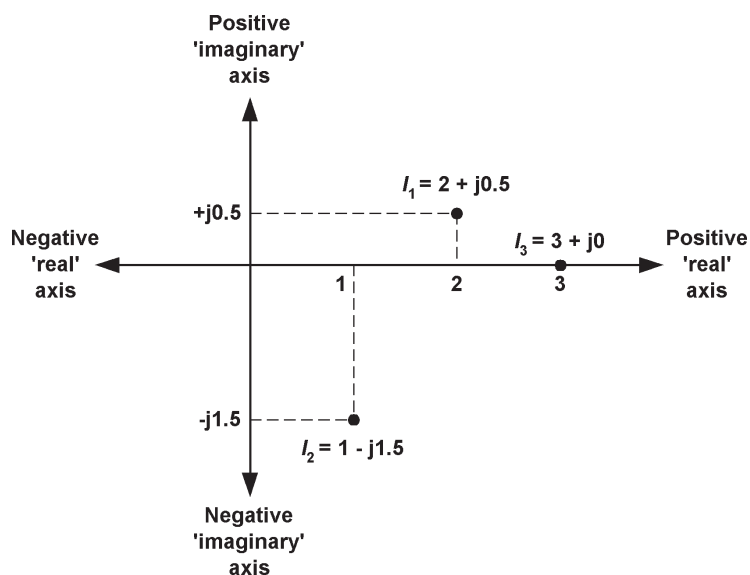


**Figure 3.3.75** Argand diagram showing complex currents

### Example 3.3.54

A current of 2 A flows in an impedance of  $(100 + j120) \Omega$ . Derive an expression, in complex form, for the voltage that will appear across the impedance.

Since  $V = I \times Z$

$$V = 2 \times (100 + j120) = (200 + j240) V.$$

Note that, in this example we have assumed that the supply current is the *reference*. In other words, it could be expressed in complex form as  $(2 + j0) A$ .

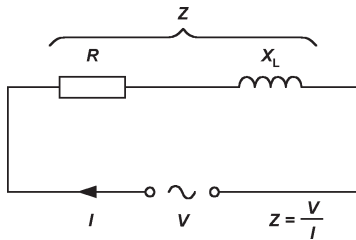


**Figure 3.3.76** Inductance and resistance in series

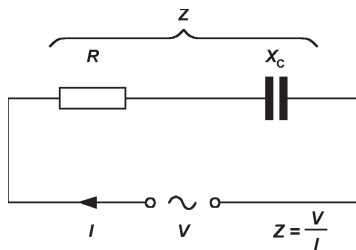


**Figure 3.3.77** Capacitance and resistance in series

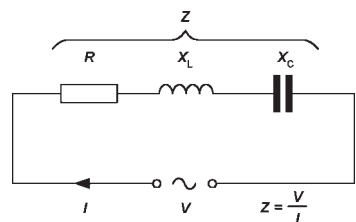


**Figure 3.3.78** Inductance, resistance, and capacitance in series

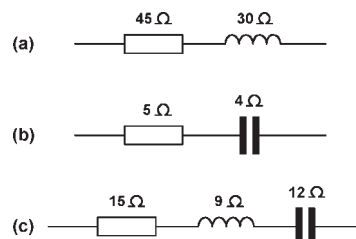


**Figure 3.3.79** See Example 3.3.56

### Example 3.3.55

An impedance of  $(200 + j100) \Omega$  is connected to a 100 V AC supply. Determine the current flowing and express your answer in complex form.

Since  $I = V/Z$

$$I = 100/(200 + j100) = 100 \times (200 - j100)/(200 + j100) \times (200 - j100)$$

(here we have multiplied the top and bottom by the *complex conjugate*)

$$I = 100 \times (200 - j100)/(200^2 + 100^2) \\ = (20\,000 - j10\,000)/(40\,000 + 10\,000).$$

Thus  $I = (2 - j)/5 = (0.4 - j0.2) \text{ A}$ .

Note that, in this example we have assumed that the supply voltage is the *reference*. In other words, it could be expressed in complex form as  $(100 + j0) \text{ V}$ .

### Inductance and resistance in series

A series circuit comprising resistance and inductance (see Figure 3.3.76) can be represented by:

$$Z = R + jX_L \quad \text{or} \quad Z = R + j\omega L$$

where  $\omega = 2\pi f$ .

### Capacitance and resistance in series

A series circuit comprising resistance and capacitance (see Figure 3.3.77) can be represented by:

$$Z = R - jX_C \quad \text{or} \quad Z = R - j/\omega C$$

where  $\omega = 2\pi f$ .

### Inductance, resistance, and capacitance in series

A series circuit comprising inductance, resistance and capacitance in series (see Figure 3.3.78) can be represented by:

$$Z = R + jX_L - jX_C = R + j(X_L - X_C)$$

$$\text{or } Z = R + j\omega L - j/\omega C = R + j(\omega L - 1/\omega C)$$

where  $\omega = 2\pi f$ .

### Example 3.3.56

Write down the impedance (in complex form) of each of the circuits shown in Figure 3.3.79.

(a)  $Z_a = 45 + j30$ .

(b)  $Z_b = 5 - j4$ .

(c)  $Z_c = 15 - j9 - j12 = 15 - j3$ .

### Complex admittances

When dealing with parallel circuits it is easier to work in terms of *admittance* ( $Y$ ) rather than *impedance* ( $Z$ ). Note that:

$$Y = 1/Z.$$

The impedance of a circuit comprising resistance connected in parallel with reactance is given by:

$$Y = (G \pm jB)$$

where  $Y$  represents *admittance*,  $G$  represents *conductance*, and  $B$  represents *susceptance*.

Similarly  $G = 1/R$  and  $B = 1/X$ .

All three quantities are measured in Siemens (S).

The  $\pm j$  term simply allows us to indicate whether the susceptance is due to capacitance (in which case the  $j$  term is positive, i.e.  $+j$ ) or whether it is due to inductance (in which case the  $j$  term is negative, i.e.  $-j$ ).

Consider, for example, the following admittances:

1.  $Y_1 = 0.05 - j0.1$  this admittance comprises a conductance of 0.05 S connected in parallel with a negative susceptance of 0.1 S. The value of resistance can be found from  $R = 1/G = 1/0.05 = 20 \Omega$  whilst the value of inductive reactance (note the minus sign before the  $j$  term) can be found from  $X = 1/B = 1/0.1 = 10 \Omega$ .
2.  $Y_2 = 0.2 + j0.05$  this admittance comprises a conductance of 0.2 S connected in parallel with a positive susceptance of 0.05 S. The value of resistance can be found from  $R = 1/G = 1/0.2 = 5 \Omega$  whilst the value of capacitive reactance (note the plus sign before the  $j$  term) can be found from  $X = 1/B = 1/0.05 = 20 \Omega$ .

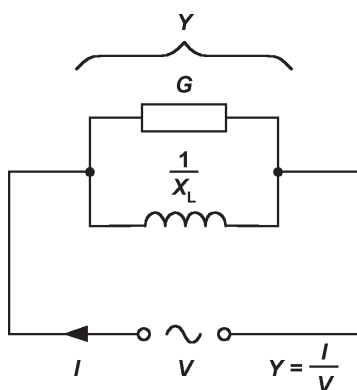


**Figure 3.3.80** Inductance and resistance in parallel

### Example 3.3.57

A voltage of 20 V appears across an admittance of  $(0.1 + j0.25) \Omega$ . Determine the current flowing and express your answer in complex form.

Since  $I = V/Z$  and  $Z = 1/Y$

$$I = V \times Y = 20 \times (0.1 + j0.25) = 2 + j5 \text{ A.}$$

Note that, in this example we have assumed that the supply voltage is the *reference*. In other words, it could be expressed in complex form as  $(20 + j0) \text{ A}$ .

### Inductance and resistance in parallel

A parallel circuit comprising resistance and inductance (see Figure 3.3.80) can be represented by:

$$Y = G - j(1/X_L) \quad \text{or} \quad Y = G - j/\omega L$$

where  $\omega = 2\pi f$ .

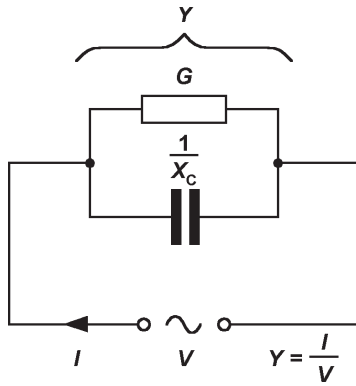


**Figure 3.3.81** Capacitance and resistance in parallel

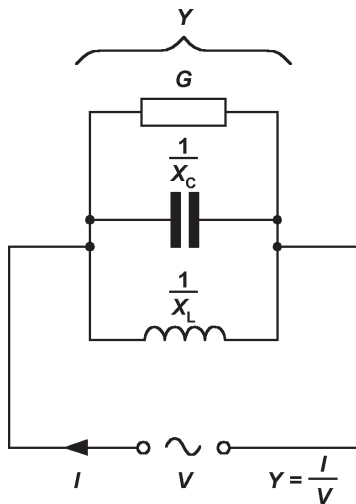


**Figure 3.3.82** Inductance, resistance, and capacitance in parallel

### Capacitance and resistance in parallel

A series circuit comprising resistance and capacitance (see Figure 3.3.81) can be represented by:

$$Y = G + j(1/X_C) \quad \text{or} \quad Y = G + j\omega C$$

where  $\omega = 2\pi f$ .

### Inductance, resistance, and capacitance in parallel

A series circuit comprising capacitance, inductance, and resistance in parallel (see Figure 3.3.82) can be represented by:

$$Y = G + j(1/X_C) - j(1/X_L) = G + j(1/X_C - 1/X_L)$$

$$Y = G + j(\omega C - 1/\omega L)$$

where  $\omega = 2\pi f$ .

#### Example 3.3.58

Write down the admittance (in complex form) of each of the circuits shown in Figure 3.3.83.

- (a)  $Y_a = 0.5 + j0.2$ .
- (b)  $Y_b = 0.25 - j0.125$ .
- (c)  $Y_c = 0.1 - j0.04 + j0.05 = 0.1 + j0.01$ .

## Resonant circuits

Thus far, we have considered circuits that contain either a combination of resistance and capacitance or a combination of resistance and inductance. These circuits can be described as 'non-resonant' since the voltage and current will not be in-phase at any frequency. More complex circuits, containing all three types of component are described as 'resonant' since there will be one frequency at which the two reactive components will be equal but opposite. At this particular frequency (known as the *resonant frequency*) the effective reactance in the circuit will be zero and the voltage and current will be in-phase. This can be a fairly difficult concept to grasp at first sight so we will explain it in terms of the way in which a resonant circuit behaves at different frequencies.

### Series resonance

A series resonant circuit comprises inductance, resistance, and capacitance connected in series, as shown in Figure 3.3.84.

At the resonant frequency, the inductive reactance,  $X_L$ , will be equal to the capacitive reactance,  $X_C$ . In other words,  $X_L = X_C$ . In this condition, the supply voltage will be in-phase with the supply current. Furthermore, since  $X_L = X_C$ , the reactive components will cancel (recall that they are  $180^\circ$  out of phase with one another) the impedance of the circuit will take a minimum value, equal to the resistance,  $R$ .

At a frequency which is less than the resonant frequency (i.e. below resonance) the inductive reactance,  $X_L$ , will be smaller than the capacitive

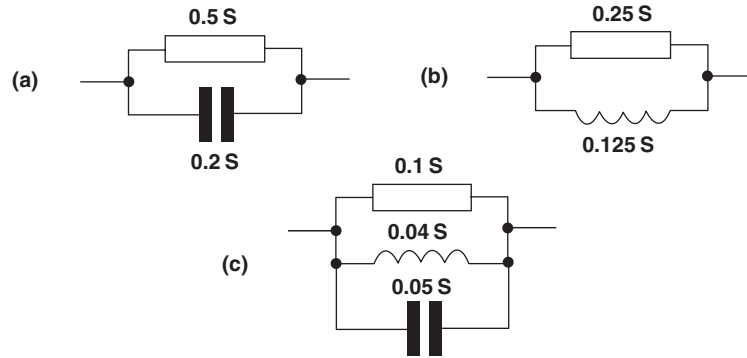


**Figure 3.3.83** See Example 3.3.58

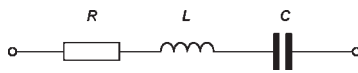


**Figure 3.3.84** Series resonant circuit

## Another view

Using  $j$  notation, the impedance of the circuit shown in Figure 3.3.84 will be given by  $Z = R + jX_L - jX_C = R + j(X_L - X_C)$ . At resonance,  $X_L = X_C$ , thus  $Z = R + j(X_L - X_C) = R + j0 = R$ . In this condition, the circuit behaves like a pure resistor and thus the supply current and voltage will be in-phase.

reactance,  $X_C$ . In other words,  $X_L < X_C$ . In this condition, the supply voltage will lag the supply current by  $90^\circ$ .

At a frequency which is greater than the resonant frequency (i.e. above resonance) the capacitive reactance,  $X_C$ , will be smaller than the inductive reactance,  $X_L$ . In other words,  $X_C < X_L$ . In this condition, the supply voltage will lag the supply current by  $90^\circ$ .

## Resonant frequency

The frequency of resonance can be determined by equating the two reactive components, as follows:

$$X_L = X_C$$

$$\text{thus } 2\pi f_o L = 1/(2\pi f_o C).$$

We need to make  $f_o$  the subject of this equation (where  $f_o$  is the frequency of resonance):

$$\text{Now } f_o^2 = 1/(4\pi^2 LC)$$

$$\text{or } f_o = \frac{1}{2\pi \sqrt{LC}}.$$

## Example 3.3.59

A series circuit consists of  $L = 60 \text{ mH}$ ,  $R = 15 \Omega$  and  $C = 15 \text{ nF}$ . Determine the frequency of resonance and the voltage dropped across each component at resonance if the circuit is connected to a 300 mV AC supply.

$$f_o = 1/2\pi \sqrt{LC} = 1/2\pi \sqrt{(60 \times 10^{-3} \times 15 \times 10^{-9})} \text{ Hz}$$

$$f_o = 1/2\pi \sqrt{(900 \times 10^{-12})} = 1/(2\pi \times 30 \times 10^{-6}) \text{ Hz}$$

$$f_o = 1 \times 10^6 / (2\pi \times 30) = 5.3 \times 10^3 \text{ Hz} = 5.3 \text{ kHz}.$$

At resonance, the reactive components ( $X_L$  and  $X_C$ ) will be equal but of opposite sign. The impedance at resonance will thus be  $R$  alone. We can determine the supply current from:

$$I = V/Z = V/R = 300 \text{ mV}/15 \Omega = 20 \text{ mA}.$$

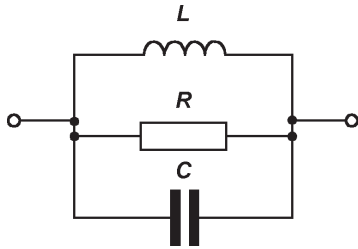


**Figure 3.3.85** Parallel resonant circuit

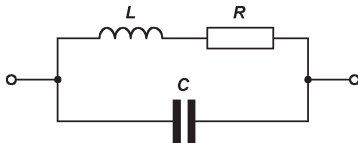


**Figure 3.3.86** Alternative form of parallel resonant circuit

The reactance of the inductor is equal to:

$$X_L = 2\pi f_o L = 2\pi \times 5.3 \times 10^3 \times 60 \times 10^{-3} \Omega \\ = 2\pi \times 5.3 \times 60 \Omega = 2 \text{ k}\Omega.$$

The voltage developed across the inductor will be given by:

$$V_L = I \times X_L = 20 \text{ mA} \times 2 \text{ k}\Omega = 40 \text{ V}.$$

Note that this voltage leads the supply current by  $90^\circ$ .

Since the reactance of the capacitor will be the same as that of the inductor, the voltage developed across the capacitor will be identical (but lagging the supply current by  $90^\circ$ ). Thus:

$$V_C = V_L = 40 \text{ V}.$$

### Parallel resonance

A parallel resonant circuit comprises inductance, resistance, and capacitance connected in parallel, as shown in Figure 3.3.85. A variation on this is when a capacitor is connected in parallel with a series combination of inductance and resistance, as shown in Figure 3.3.86.

In the case of the circuit shown in Figure 3.3.85, the frequency of resonance can once again be determined by simply equating the two reactive components, as follows:

$$X_L = X_C$$

$$\text{thus } 2\pi f_o L = 1/(2\pi f_o C)$$

$$\text{or } f_o = \frac{1}{2\pi \sqrt{LC}}.$$

### Example 3.3.60

A parallel circuit consists of  $L = 40 \text{ mH}$ ,  $R = 1 \text{ k}\Omega$  and  $C = 10 \text{ nF}$ . Determine the frequency of resonance and the current in each component at resonance if the circuit is connected to a  $2 \text{ V}$  AC supply.

$$f_o = 1/2\pi \sqrt{LC} = 1/2\pi \sqrt{(40 \times 10^{-3} \times 10 \times 10^{-9})} \text{ Hz}$$

$$f_o = 1/2\pi \sqrt{(400 \times 10^{-12})} = 1/(2\pi \times 20 \times 10^{-6}) \text{ Hz}$$

$$f_o = 1 \times 10^6 / (2\pi \times 20) = 7.96 \times 10^3 \text{ Hz} = 7.96 \text{ kHz}.$$

At resonance, the reactive components ( $X_L$  and  $X_C$ ) will be equal but of opposite sign. The impedance at resonance will thus be  $R$  alone. We can determine the supply current from:

$$I = V/Z = V/R = 2 \text{ V}/1 \text{ k}\Omega = 2 \text{ mA}.$$

The reactance of the inductor is equal to:

$$X_L = 2\pi f_o L = 2\pi \times 7.96 \times 10^3 \times 40 \times 10^{-3} \Omega \\ = 2\pi \times 7.96 \times 40 \Omega = 2 \text{ k}\Omega.$$



The current in the inductor will be given by:

$$I = V_L / X_L = 2 \text{ V} / 2 \text{ k}\Omega = 1 \text{ mA}.$$

Note that this current lags the supply voltage by  $90^\circ$ .

Since the reactance of the capacitor will be the same as that of the inductor, the current in the capacitor will be identical (but leading the supply current by  $90^\circ$ ). Thus:

$$I_C = I_L = \mathbf{2 \text{ mA}}.$$

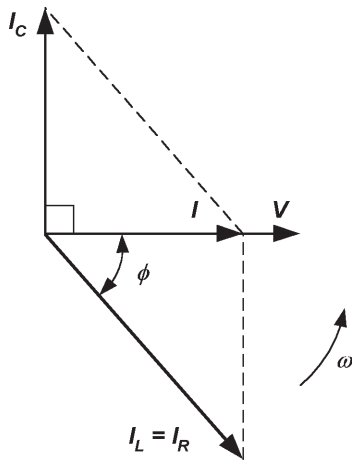


**Figure 3.3.87** Phasor diagram for the circuit of Figure 3.3.68

In the case of the circuit shown in Figure 3.3.86, the frequency of resonance is once again the frequency at which the supply voltage and current are in-phase. Note that the current in the inductor,  $I_L$ , lags the supply voltage by an angle,  $\phi$  (see Figure 3.3.87).

$$\text{Now } I_L = V/Z_L = V/\sqrt{(R^2 + X_L^2)}$$

$$\text{and } I_C = V/X_C.$$

$$\text{At resonance, } I_C = I_L \sin \phi \text{ and } \sin \phi = X_L/Z_L.$$

$$\text{Thus } V/X_C = V/\sqrt{(R^2 + X_L^2)} \times X_L/Z_L$$

$$\text{or } V\omega C = V/\sqrt{(R^2 + (\omega L)^2)} \times \omega L/\sqrt{(R^2 + X_L^2)}$$

$$C = L/(R^2 + (\omega L)^2)$$

$$L/C = (R^2 + (\omega L)^2)$$

$$(\omega L)^2 = (L/C) - R^2$$

$$\omega^2 = 1/LC - R^2/L^2$$

$$\omega = \sqrt{(1/LC - R^2/L^2)} \text{ but, } \omega = 2\pi f_0$$

$$\text{thus } f_0 = \frac{1}{2\pi} \sqrt{\left(\frac{1}{LC} - \frac{R^2}{L^2}\right)}.$$

At resonance, the impedance of the circuit shown in Figure 3.3.86 is called its *dynamic impedance*. This impedance is given by:

$$Z_d = V/I.$$

From the phasor diagram of Figure 3.3.87:

$$I = I_C/\tan \phi.$$

$$\text{Thus, } Z_d = V \times \tan \phi / I_C = (V/I_C) \times \tan \phi = X_C \times \tan \phi.$$

$$\text{Since } \tan \phi = X_L/R$$

$$Z_d = X_C \times X_L/R = 1/(2\pi f_0 C) \times (2\pi f_0 L/R) = L/CR.$$

This value of impedance is known as the *dynamic impedance* of the circuit. You should note that the dynamic impedance increases as the ratio of  $L/C$  increases.

**Example 3.3.61**

A coil having an inductance of 1 mH and a resistance of 100  $\Omega$  is connected in parallel with a capacitor of 10 nF. Determine the frequency of resonance and the dynamic impedance of the circuit.

$$f_o = \frac{1}{2\pi} \sqrt{\left(\frac{1}{LC} - \frac{R^2}{L^2}\right)}$$

$$f_o = 1/2\pi \sqrt{(1/(1 \times 10^{-3} \times 10 \times 10^{-9}) - (100)^2/(1 \times 10^{-3})^2)}$$

$$f_o = 1/2\pi \sqrt{(1/(1 \times 10^{-11}) - (10^4/10^{-6}))}$$

$$f_o = 1/2\pi \sqrt{(10^{11} - 10^{10})}$$

$$f_o = 1/2\pi \sqrt{(9 \times 10^{10})} = (3 \times 10^5)/2\pi = 47.7 \text{ kHz.}$$

The dynamic impedance is given by:

$$Z_d = L/(CR) = 1 \times 10^{-3}/(10 \times 10^{-9} \times 100) = 1 \times 10^3 = 1 \text{ k}\Omega.$$

**Q-factor**

The Q-factor (or *quality factor*) is a measure of the ‘goodness’ of a tuned circuit. Q-factor is sometimes also referred to as ‘voltage magnification factor’. Consider the circuit shown in Figure 3.3.88. The Q-factor of the circuit tells you how many times greater the inductor or capacitor voltage is than the supply voltage. The better the circuit, the higher the ‘voltage magnification’ and the greater the Q-factor.

In Figure 3.3.88,  $Q = V_L/V = V_C/V$ .

Since  $V_L = I \times X_L = I \times \omega L$  and  $V = I \times Z = I \times R$  (at resonance):

$$Q = V_L/V = I \times \omega L / I \times R = \omega L / R = \frac{2\pi f_o L}{R}.$$

Similarly, since  $V_C = I \times X_C = I/\omega C$  and  $V = I \times Z = I \times R$  (at resonance):

$$Q = V_C/V = I/\omega C \times I \times R = 1/\omega CR = \frac{1}{2\pi f_o CR}.$$

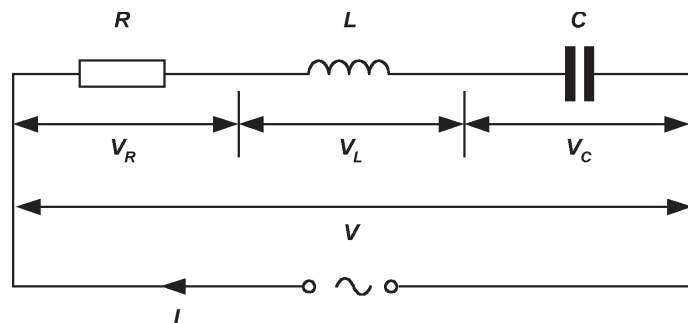


**Figure 3.3.88** Voltages in a resonant circuit

### Bandwidth

As the Q-factor of a tuned circuit increases, the frequency response curve becomes sharper. As a consequence, the bandwidth of the tuned circuit is reduced. This relationship is illustrated in Figure 3.3.89.

We normally specify the range of frequencies that will be accepted by a tuned circuit by referring to the two *half-power frequencies*. These are the frequencies at which the power in a tuned circuit falls to 50% of its maximum value. At these frequencies:

- (1) The tuned circuits resistive and reactive components ( $R$  and  $X$ ) will be equal.
- (2) The phase angle (between current and voltage) will be  $\pi/4$  (or  $45^\circ$ ).
- (3) Both current and voltage will have fallen to  $1/\sqrt{2}$  or 0.707 of their maximum value.

At the upper of the two cut-off frequencies,  $f_2$

$$R = X_{\text{effective}} \text{ (from 1 on page 231)}$$

where  $X_{\text{effective}} = X_L - X_C$  (since  $X_L > X_C$  above  $f_0$ )

$$\text{thus: } R = X_L - X_C = 2\pi f_2 L - 1/2\pi f_2 C$$

rearranging gives:

$$R - 1/2\pi f_2 L = -1/2\pi f_2 C$$

$$\text{thus } 2\pi f_2 R - (2\pi f_2)^2 L = -1/C$$

$$\text{or } 1/C = -2\pi f_2 R + (2\pi f_2)^2 L \dots (i).$$

Whereas, at the lower of the two cut-off frequencies,  $f_1$ :

$$R = X_{\text{effective}} \text{ (from 1 on page 231)}$$

where  $X_{\text{effective}} = X_C - X_L$  (since  $X_C > X_L$  below  $f_0$ )

$$\text{thus: } R = X_C - X_L = 1/2\pi f_1 C - 2\pi f_1 L$$

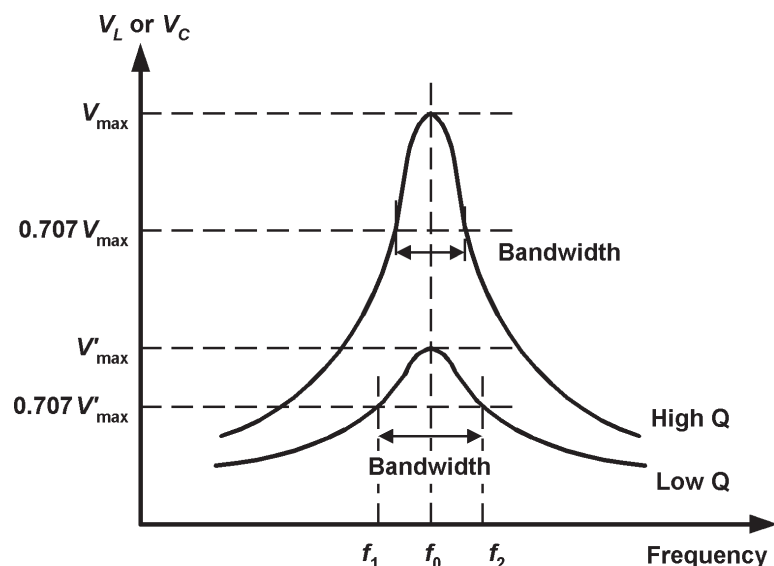


**Figure 3.3.89** Relationship between Q-factor and bandwidth

re-arranging gives:

$$R + 2\pi f_1 L = 1/2\pi f_1 C$$

$$\text{thus } 2\pi f_1 R + (2\pi f_1)^2 L = 1/C$$

$$\text{or } 1/C = 2\pi f_1 R + (2\pi f_1)^2 L \dots (ii).$$

Equating (i) and (ii) gives:

$$-2\pi f_2 R + (2\pi f_2)^2 L = 2\pi f_1 R + (2\pi f_1)^2 L$$

$$(2\pi f_2)^2 L - (2\pi f_1)^2 L = 2\pi f_1 R + 2\pi f_2 R$$

$$(2\pi f_2)^2 - (2\pi f_1)^2 = R/L \times (2\pi f_1 + 2\pi f_2)$$

$$(2\pi f_1 + 2\pi f_2)(2\pi f_1 - 2\pi f_2) = R/L \times (2\pi f_1 + 2\pi f_2)$$

$$(2\pi f_1 - 2\pi f_2) = R/L$$

$$f_1 - f_2 = R/2\pi L.$$

$$\text{But bandwidth, } f_w = f_1 - f_2$$

$$\text{thus } f_w = R/2\pi L.$$

Dividing both sides by  $f_o$  gives:

$$f_w/f_o = R/2\pi L f_o.$$

Earlier, we defined Q-factor as:

$$Q = 2\pi L f_o / R.$$

$$\text{Thus } f_w/f_o = 1/Q$$

$$\text{or } Q = f_o/f_w.$$

### Example 3.3.62

A tuned circuit comprises a  $400 \mu\text{H}$  inductor connected in series with a  $100 \text{ pF}$  capacitor and a resistor of  $10 \Omega$ . Determine the resonant frequency of the tuned circuit, its Q-factor and bandwidth.

$$f_o = 1/2\pi \sqrt{LC} = 1/2\pi \sqrt{(400 \times 10^{-6} \times 100 \times 10^{-12})}$$

$$f_o = 1/2\pi \sqrt{(4 \times 10^{-14})} = 1/2\pi \times 2 \times 10^{-7}$$

$$f_o = 0.159/2 \times 10^{-7} = 0.0795 \times 10^7$$

$$f_o = 0.795 \times 10^6 \text{ Hz or } 795 \text{ kHz.}$$

The Q-factor is determined from:

$$Q = 2\pi f_o L / R = 6.28 \times 795 \times 10^3 \times 400 \times 10^{-6}$$

$$Q = 199.7 \times 10^4 \times 10^3 \times 10^{-6} = 19.97.$$

The bandwidth is determined from:

$$f_w = f_o / Q = 795 \times 10^3 / 19.97 = 39.8 \times 10^3 \text{ Hz.}$$

$$\text{Thus } f_w = \mathbf{39.8 \text{ kHz.}}$$

**Example 3.3.63**

A series tuned circuit filter is to be centred on a frequency of 73 kHz. If the filter is to use a capacitor of 1 nF, determine the value of inductance required and the maximum value of series loss resistance that will produce a bandwidth of less than 1 kHz.

$$\text{Now } f_o = 1/2\pi\sqrt{LC}.$$

$$\text{Thus } L = 1/(4\pi^2 Cf_o^2)$$

$$L = 1/(4\pi^2 \times 1 \times 10^{-9} \times (73 \times 10^3)^2)$$

$$L = 1/(4\pi^2 \times 1 \times 10^{-9} \times 5329 \times 10^6)$$

$$L = 10^3/(4\pi^2 \times 5329) = 10^3/210\,435 = 0.00475 \text{ H}$$

$$\text{or } L = 4.75 \text{ mH.}$$

$$\text{Now } f_w = f_o/Q \text{ and } Q = 2\pi f_o L/R$$

$$\text{Thus } f_w = f_o/(2\pi f_o L/R) \text{ or } f_w = f_o \times R/(2\pi f_o L).$$

$$\text{Hence } R = f_w \times (2\pi f_o L)/f_o = f_w \times 2\pi L$$

$$\text{thus } R = 1 \times 10^3 \times 2\pi \times 4.75 \times 10^{-3} = \mathbf{29.8 \, \Omega}.$$

Thus, to achieve a bandwidth of 1 kHz or less the total series loss resistance must not be greater than 29.8  $\Omega$ .

**Problems 3.3.6**

- (1) A sinusoidal current of 2 A flows in an impedance of  $(100 + j120) \, \Omega$ . Derive an expression, in complex form, for the voltage that will appear across the impedance.
- (2) A sinusoidal voltage of 20 V appears across an admittance of  $(0.1 + j0.25) \, \Omega$ . Determine the current flowing and express your answer in complex form.
- (3) A coil has an inductance of 60 mH and a resistance of 20  $\Omega$ . If a sinusoidal current of 20 mA at 100 Hz flows in the coil, derive an expression for the voltage that will appear across it. Express your answer in complex form.
- (4) A capacitor of 2  $\mu\text{F}$  is connected in parallel with a resistor of 500  $\Omega$ . If the parallel circuit is connected to a 20 V 50 Hz sinusoidal supply, derive an expression for the current that will be supplied. Express your answer in complex form.
- (5) A series circuit consists of  $L = 60 \text{ mH}$ ,  $R = 15 \, \Omega$ , and  $C = 15 \text{ nF}$ . Determine the frequency of resonance and the voltage dropped across each component at resonance if the circuit is connected to a 300 mV AC supply.
- (6) A coil having an inductance of 1 mH and a resistance of 100  $\Omega$  is connected in parallel with a capacitor of 10 nF. Determine the frequency of resonance and the dynamic impedance of the circuit.
- (7) The aerial tuned circuit of a long wave receiver is to tune from 150 to 300 kHz. Determine the required maximum and minimum values of tuning capacitor if the inductance of the aerial coil is 900  $\mu\text{H}$ .

- (8) The parallel tuned circuit to be used in a variable frequency oscillator comprises a coil of inductance  $500\ \mu\text{H}$  and resistance  $45\ \Omega$  and a variable tuning capacitor having maximum and minimum values of  $1000\ \text{pF}$  and  $100\ \text{pF}$ , respectively. Determine the tuning range for the oscillator and the Q-factor of the oscillator tuned circuit at each end of the tuning range.

## Power in an AC circuit

We have already shown how power in an AC circuit is only dissipated in resistors and not in inductors or capacitors. This merits further consideration as we need to be aware of the implications, particularly in applications where appreciable amounts of power may be present.

### Power in a pure resistance

The voltage dropped across a pure resistance rises and falls in sympathy with the current. The voltage and current are said to be *in-phase*. Since the power at any instant is equal to the product of the *instantaneous voltage* and the *instantaneous current*, we can determine how the power supplied varies over a complete cycle of the supply current. This relationship is illustrated in Figure 3.3.90. From this you should note that:

- the power curve represents a cosine law at twice the frequency of the supply current,
- all points on the power curve are positive throughout a complete cycle of the supply current.

From the foregoing, the power at any instant of time (known as the *instantaneous power*) is given by:

$$p = i \times v$$

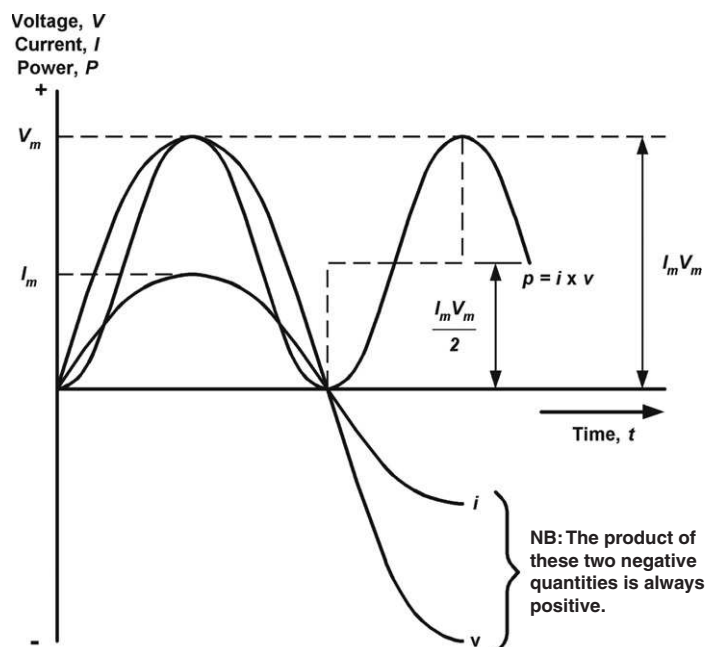


**Figure 3.3.90** Voltage, current, and power in pure resistance

where  $i$  is the *instantaneous current* and  $v$  is the *instantaneous voltage*.

since  $i = I_{\max} \sin(\omega t)$  and  $v = V_{\max} \sin(\omega t)$

$$p = i \times v = I_{\max} \sin(\omega t) \times V_{\max} \sin(\omega t) = I_{\max} V_{\max} \sin^2(\omega t).$$

$$\text{Thus } p = \frac{I_{\max} V_{\max}}{2} (1 - \cos(2\omega t)).$$

## Mathematics in action

### The double angle formula

The foregoing relationship is derived from the ‘double angle formula’,  $\sin^2 \theta = 1/2(1 - \cos(2\theta))$ . There are two important results to note:

- (1) The power function is *always positive* (i.e. a graph of  $p$  plotted against  $t$  will always be above the  $x$  axis – see Figure 3.3.90).
- (2) The power function has *twice the frequency* of the current or voltage.
- (3) The average value of power over a complete cycle is equal to

$$\frac{I_{\max} V_{\max}}{2}.$$

We normally express power in terms of the rms values of current and voltage. In this event, the *average power* over one complete cycle of current is given by:

$$P = I \times (I \times R) = I^2 \times R \quad \text{or} \quad P = V \times V/R = V^2/R$$

where  $I$  and  $V$  are rms values of voltage and current.

### Question 3.3.3

Show that the instantaneous power in a resistive load connected to sinusoidal AC supply is given by:

$$p = \frac{I_{\max}^2 R}{2} \times (1 - \cos(2\omega t)).$$

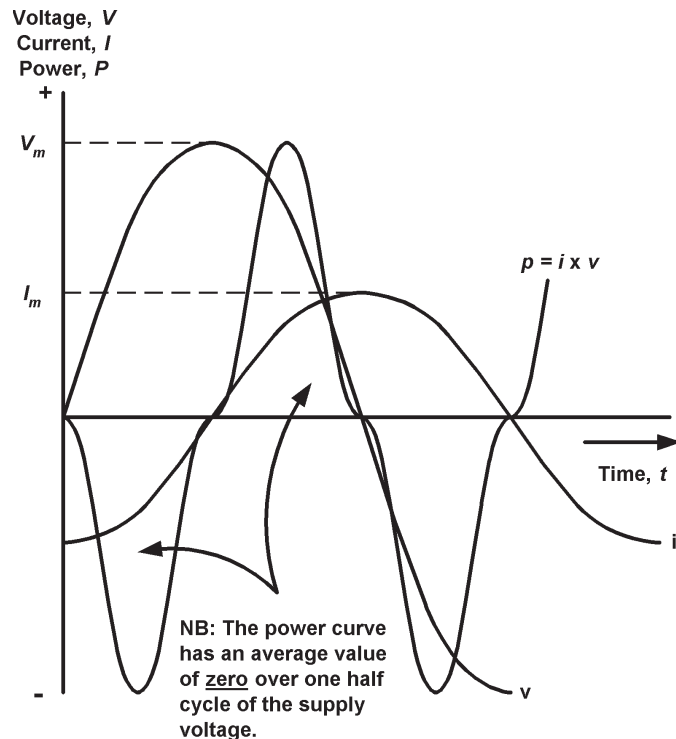
Using a graph of the power function, show that the average power over a complete cycle of the supply is given by:

$$P = I_{\max}^2 R/2.$$

### Power in a pure reactance

We have already shown how the voltage dropped across a pure inductive reactance leads the current flowing in the inductor by an angle of  $90^\circ$ . Since the power at any instant is equal to the product of the instantaneous voltage and current, we can once again determine how the power supplied varies over a complete cycle of the supply current, as shown in Figure 3.3.91.

**Figure 3.3.91** Voltage, current, and power in a pure inductive reactance



From this you should note that:

- (a) the power curve represents a sine law at twice the frequency of the supply current, and
- (b) the power curve is partly positive and partly negative and the average value of power over a complete cycle of the supply current is zero.

Similarly, we have already shown how the voltage dropped across a pure capacitive reactance lags the current flowing in the capacitor by an angle of  $90^\circ$ . Since the power at any instant is equal to the product of the instantaneous voltage and current, we can once again determine how the power supplied varies over a complete cycle of the supply current, as shown in Figure 3.3.92.

From this you should note that, once again:

- (a) the power curve represents a sine law at twice the frequency of the supply current,
- (b) the power curve is partly positive and partly negative and the average value of power over a complete cycle of the supply current is zero.

## Power factor

The *true power* in an AC circuit is simply the average power that it consumes. Thus:

$$\text{True power} = I^2 \times R.$$

True power is measured in watts (W).

The *apparent power* in an AC circuit is simply the product of the supply voltage and the supply current. Thus:

$$\text{Apparent power} = V \times I$$



**Figure 3.3.92** Voltage, current, and power in a pure capacitive reactance

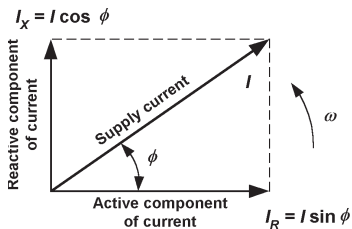
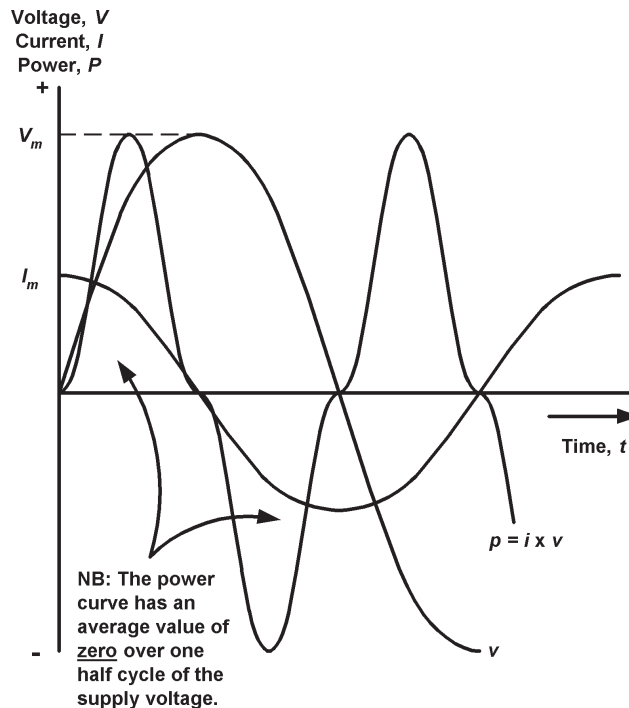


**Figure 3.3.93** Relationship between active and reactive components of supply current

### Another view

In circuits that contain only inductance or capacitance, power is taken from the supply on each cycle when the magnetic or electric fields are created and then returned to the supply when the fields later collapse. If the components are 'loss-free' (i.e. if they have no internal resistance) the same amount of power is returned to the supply as is originally taken from it. The average power must, therefore, be zero.

Apparent power is measured in volt-amperes (VA).

The *power factor* of an AC circuit is simply the ratio of the *true power* to the *apparent power*. Thus:

$$\text{Power factor} = \text{True power} / \text{Apparent power}.$$

Power factor provides us with an indication of how much of the power supplied to an AC circuit is converted into useful energy. A high power factor (i.e. a value close to 1) in an  $L$ - $R$  or  $C$ - $R$  circuit indicates that most of the energy taken from the supply is dissipated as heat produced by the resistor. A low power factor (i.e. a value close to zero) indicates that, despite the fact that current and voltage is supplied to the circuit, most of the energy is returned to the supply and very little of it is dissipated as heat.

In terms of units, power factor = watts/VA.

The current supplied to a reactive load can be considered to have two components acting at right angles. One of these components is in-phase with the supply current and it is known as the *active component of current*. The other component is  $90^\circ$  out of phase with the supply current ( $+90^\circ$  in the case of a capacitive load and  $-90^\circ$  in the case of an inductive load) and this component is known as the *reactive component of current*. The relationship between the active and reactive components of the supply current is illustrated in Figure 3.3.93.

From Figure 3.3.93, you should note that the active component of current is given by  $(I \times \cos \phi)$  whilst the reactive component of current is given by  $(I \times \sin \phi)$ .

The true power is given by the product of supply voltage and the active component of current. Thus:

$$\text{True power} = V \times (I \times \cos \phi) = VI \times \cos \phi \dots \quad (\text{i})$$

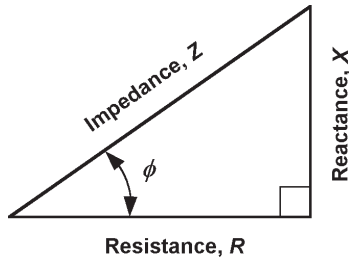


**Figure 3.3.94** Impedance triangle

Similarly, the reactive power is given by the product of supply voltage and the reactive component of current. Thus:

$$\text{Reactive power} = V \times (I \times \sin \phi) = VI \times \sin \phi \dots \quad (\text{ii})$$

Also from Figure 3.3.93, the product of supply current and voltage (i.e. the *apparent power*) is given by:

$$\text{Apparent power} = V \times I \dots \quad (\text{iii})$$

Combining formula (i) and formula (iii) allows us to express power factor in a different way:

$$\text{Power factor} = \text{True power} / \text{Apparent power} = (VI \cos \phi) / (VI) = \cos \phi$$

$$\text{Hence power factor} = \cos \phi.$$

(Note that power factor can have values of between 0 and 1).

We can also express power factor in terms of the ratio of resistance,  $R$ , and reactance,  $X$ , present in the load. This is done by referring to the impedance triangle (Figure 3.3.94), where:

$$\text{Power factor} = \cos \phi = R/Z.$$

The ratio of reactive power to true power can be determined by dividing Equation (ii) by Equation (i). Thus:

$$\text{Reactive power} / \text{True power} = VI \sin \phi / VI \cos \phi = \tan \phi.$$

### Example 3.3.64

A power of 1 kW is supplied to an AC load which operates from a 220 V 50 Hz supply. If the supply current is 6 A, determine the power factor of the load.

$$\begin{aligned} \text{Power factor} &= \text{True power} / \text{Apparent power} \\ &= 1 \text{ kW} / (220 \text{ V} \times 6 \text{ A}). \end{aligned}$$

$$\text{Thus power factor} = 1 \text{ kW} / 1.32 \text{ kVA} = \mathbf{0.758}.$$

### Example 3.3.65

The phase angle between the voltage and current supplied to an AC load is  $25^\circ$ . If the supply current is 3 A and the supply voltage is 220 V, determine the true power and reactive power in the load.

$$\text{Power factor} = \cos \phi = \cos 25^\circ = 0.906.$$

$$\begin{aligned} \text{True power} &= 220 \times 3 \times \cos 25^\circ = 220 \times 3 \times 0.906 \\ &= \mathbf{598 \text{ W}}. \end{aligned}$$

$$\begin{aligned} \text{Reactive power} &= 220 \times 3 \times \sin 25^\circ = 220 \times 3 \times 0.423 \\ &= \mathbf{279 \text{ W}}. \end{aligned}$$

### Another view

It might help to explain the relationship between power factor and supply current by taking a few illustrative values. Let us assume that a motor produces a power of 1 kW and it operates with an efficiency 60%. The input power (in other words, the product of the voltage and current taken from the supply) will be (1 kW/0.6) or 1.667 kW. The supply current that would be required at various different power factors (assuming a 220 V AC supply) is given below:

| Power factor | Supply current |
|--------------|----------------|
| 0.2          | 38 A           |
| 0.4          | 19 A           |
| 0.6          | 12.6 A         |
| 0.8          | 9.5 A          |
| 1.0          | 7.6 A          |

High values of power factor clearly result in the lowest values of supply current!

### Example 3.3.66

An AC load comprises an inductance of 0.5 H connected in series with a resistance of 85  $\Omega$ . If the load is to be used in conjunction with a 220 V 50 Hz AC supply, determine the power factor of the load and the supply current.

The reactance of the inductor,  $X_L$ , can be found from:

$$X_L = 2\pi fL = 2\pi \times 50 \times 0.5 = 157 \Omega.$$

$$\begin{aligned}\text{Power factor} = \cos \phi &= R/Z = \frac{R}{\sqrt{R^2 + X_L^2}} \\ &= \frac{85}{\sqrt{85^2 + 157^2}} = 0.48.\end{aligned}$$

The supply current,  $I$ , can be calculated from:

$$I = V/Z = V/\sqrt{R^2 + X_L^2} = 220/\sqrt{(85^2 + 157^2)}$$

$$I = 220/\sqrt{(7225 + 24\,649)}$$

$$I = 220/\sqrt{(31\,874)} = 220/178.5 = \mathbf{1.23\,A}.$$

### Power factor correction

When considering the utilisation of AC power, it is very important to understand the relationship between power factor and supply current. We have already shown that:

$$\text{Power factor} = \text{True power}/\text{Apparent power}.$$

Thus:

$$\text{Apparent power} = \text{True power}/\text{Power factor}.$$

or:

$$VI = \text{True power}/\text{Power factor}.$$

Since  $V$  is normally a constant (i.e. the supply voltage) we can infer that:

$$I \propto \text{True power}/\text{Power factor}.$$

This relationship shows that the supply current is inversely proportional to the power factor. This has some important implications. For example, it explains why a large value of supply current will flow when a load has a low value of power factor. Conversely, the smallest value of supply current will occur when the power factor is 1 (the largest possible value for the power factor).

### Example 3.3.67

A power of 400 W is supplied to an AC load when it is connected to a 110 V 60 Hz supply. If the supply current is 6 A, determine the power factor of the load and the phase angle between the supply voltage and current.

True power = 400 W.

Apparent power =  $110 \text{ V} \times 6 \text{ A} = 660 \text{ VA}$ .

Power factor = True power/Apparent power  
 $= 400 \text{ W}/660 \text{ VA} = 0.606$ .

Power factor =  $\cos \phi$  thus  $\phi = \cos^{-1}(0.606) = 52.7^\circ$ .

### Example 3.3.68

A motor produces an output of 750 W at an efficiency of 60% when operated from a 220 V AC mains supply. Determine the power factor of the motor if the supply current is 9.5 A. If the power factor is increased to 0.9 by means of a 'power factor correcting circuit', determine the new value of supply current.

True power =  $750 \text{ W}/0.6 = 1.25 \text{ kW}$ .

Power factor =  $1.25 \text{ kW}/(220 \text{ V} \times 9.5 \text{ A}) = 1.25 \text{ kW}/1.65 \text{ kVA}$   
 $= 0.76$ .

Supply current (for a power factor of 0.9) =  $1.25 \text{ kW}/(220 \text{ V} \times 0.9) = 6.3 \text{ A}$ .

It is frequently desirable to take steps to improve the power factor of an AC load. When the load is inductive, this can be achieved by connecting a capacitor in parallel with the load. Conversely, when the load is capacitive, the power factor can be improved by connecting an inductor in parallel with the load. In order to distinguish between the two types of load (inductive and capacitive) we often specify the power factor as either *lagging* (in the case of an inductor) or *leading* (in the case of a capacitor).

In either case, the object of introducing a component of *opposite* reactance is to reduce the overall phase angle (i.e. the angle between the supply current and the supply voltage). If, for example, a load has a (uncorrected) phase angle of  $15^\circ$  lagging, the phase angle could be reduced to  $0^\circ$  by introducing a component that would exhibit an equal, but opposite, phase angle of  $15^\circ$  leading. To put this into perspective, consider the following example:

### Example 3.3.69

A 1.5 kW AC load is operated from a 220 V 50 Hz AC supply. If the load has a power factor of 0.6 lagging, determine the value of capacitance that must be connected in parallel with it in order to produce a unity power factor.

$\cos \phi = 0.6$  thus  $\phi = \cos^{-1}(0.6) = 53.1^\circ$ .

Reactive power = True power  $\times \tan \phi = 1.5 \text{ kW} \times \tan(53.1^\circ)$ .

Thus reactive power =  $1.5 \times 1.33 \text{ kVA} = 2 \text{ kVA}$ .

In order to increase the power factor from 0.6 to 1, the capacitor connected in parallel with the load must consume an equal reactive power. We can use this to determine the current that must flow in the capacitor and hence its reactance.

Capacitor current,  $I_C = 2 \text{ kVA} / 220 \text{ V} = 9.1 \text{ A}$ .

Now  $X_C = V / I_C = 220 \text{ V} / 9.1 \text{ A} = 24.2 \Omega$ .

Since  $X_C = 1 / (2\pi f C)$ .

$C = 1 / (2\pi f X_C) = 1 / (2\pi \times 50 \times 24.2) = 1 / 7599 \text{ F}$ .

Thus  $C = 1.32 \times 10^{-4} \text{ F} = 132 \mu\text{F}$ .

### Problems 3.3.7

- (1) An AC load draws a current of 2.5 A from a 220 V supply. If the load has a power factor of 0.7, determine the active and reactive components of the current flowing in the load.
- (2) An AC load has an effective resistance of  $90 \Omega$  connected in series with a capacitive reactance of  $120 \Omega$ . Determine the power factor of the load and the apparent power that will be supplied when the load is connected to a 415 V 50 Hz supply.
- (3) The phase angle between the voltage and current supplied to an AC load is  $37^\circ$ . If the supply current is 2 A and the supply voltage is 415 V, determine the true power and reactive power in the load.
- (4) A coil has an inductance of 0.35 H and a resistance of  $65 \Omega$ . If the coil is connected to a 220 V 50 Hz supply determine its power factor and the value of capacitance that must be connected in parallel with it in order to raise the power factor to unity.
- (5) An AC load comprises an inductance of 1.5 H connected in series with a resistance of  $85 \Omega$ . If the load is to be used in conjunction with a 220 V 50 Hz AC supply, determine the power factor of the load and the supply current.
- (6) A power of 400 W is supplied to an AC load when it is connected to a 110 V 60 Hz supply. If the supply current is 6 A, determine the power factor of the load and the phase angle between the supply voltage and current.
- (7) A motor produces an output of 750 W at an efficiency of 60% when operated from a 220 V AC mains supply. Determine the power factor of the motor if the supply current is 9.5 A. If the power factor is increased to 0.9 by means of a 'power factor correcting circuit', determine the new value of supply current.
- (8) An AC load consumes a power of 800 W from a 110 V 60 Hz supply. If the load has a lagging power factor of 0.6, determine the value of parallel connected capacitor required to produce a unity power factor.

### Waveforms

#### *The sinusoidal waveform*

This is the most fundamental of all wave shapes and all other waveforms can be synthesised from sinusoidal components. To specify a sine wave,

we need to consider just three things: amplitude, frequency, and phase. Since no components are present at any other frequency (other than that of the fundamental sine wave) a sinusoidal wave is said to be a *pure tone*. All other waveforms can be reproduced by adding together sine waves of the correct amplitude, frequency, and phase. The study of these techniques is called Fourier Analysis (see Chapter 4).

### Fundamental and harmonic components

An integer multiple of a *fundamental* frequency is known as a *harmonic*. In addition, we often specify the order of the harmonic (second, third, and so on). Thus the *second harmonic* has twice the frequency of the fundamental, the *third harmonic* has three times the frequency of the fundamental, and so on. Consider, for example, a fundamental signal at 1 kHz. The second harmonic would have a frequency of 2 kHz, the third harmonic a frequency of 3 kHz, and the fourth harmonic a frequency of 4 kHz.

Note that, in musical terms, the relationship between notes that are one *octave* apart is simply that the two frequencies have a ratio of 2:1 (in other words, the higher frequency is double the lower frequency).

### Complex waveforms

Complex waveforms comprise a fundamental component together with a number of harmonic components, each having a specific amplitude and with a specific phase relative to the fundamental. The following example shows that this is not quite so complicated as it sounds!

#### Example 3.3.70

Consider a sinusoidal signal with an amplitude of 1 V at a frequency of 1 kHz. The waveform of this fundamental signal is shown in Figure 3.3.95.

Now consider the second harmonic of the first waveform. Let us suppose that this has an amplitude of 0.5 V and that it is in-phase with the fundamental. This 2 kHz signal is shown in Figure 3.3.96.

Finally, let us add the two waveforms together at each point in time. This produces the complex waveform shown in Figure 3.3.97.

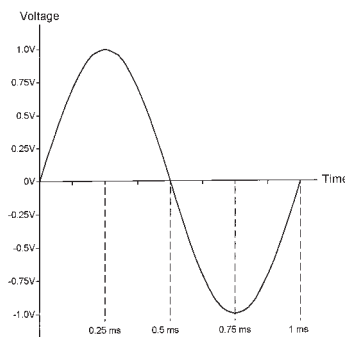


**Figure 3.3.95** Fundamental component

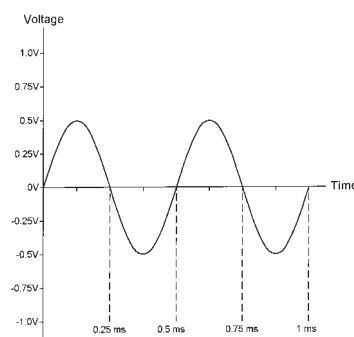


**Figure 3.3.96** Second harmonic component

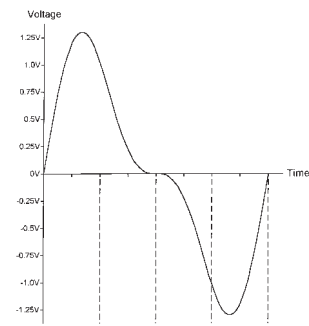


**Figure 3.3.97** Fundamental plus second harmonic components

We can describe a complex wave using an equation of the form:

$$v = V_1 \sin(\omega t) + V_2 \sin(2\omega t \pm \phi_2) + V_3 \sin(3\omega t \pm \phi_3) \\ + V_4 \sin(4\omega t \pm \phi_4) + \dots$$

Where  $v$  is the instantaneous voltage of the complex waveform at time,  $t$ .  $V_1$  is the amplitude of the fundamental,  $V_2$  is the amplitude of the second harmonic,  $V_3$  is the amplitude of the third harmonic, and so on. Similarly,  $\phi_2$  is the phase angle of the second harmonic (relative to the fundamental),  $\phi_3$  is the phase angle of the third harmonic (relative to the fundamental), and so on.

The important thing to note from the foregoing equation is that *all of the individual components that make up a complex waveform have a sine wave shape*.

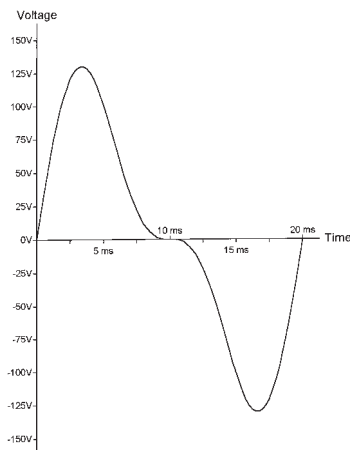


**Figure 3.3.98** See Example 3.3.71

### Example 3.3.71

The complex waveform shown in Figure 3.3.98 is given by:

$$v = 100 \sin(100\pi t) + 50 \sin(200\pi t) \text{ V.}$$

Determine:

- The amplitude of the fundamental.
- The frequency of the fundamental.
- The order of any harmonic components present.
- The amplitude of any harmonic components present.
- The phase angle of any harmonic components present (relative to the fundamental).

Comparing the foregoing equation with the general equation (see above) yields the following:

- The first term is the fundamental and this has an amplitude ( $V_1$ ) of 100 V.
- The frequency of the fundamental can be determined from:  
 $\sin(\omega t) = \sin(100\pi t)$ .  
 Thus  $\omega = 100\pi$  but  $\omega = 2\pi f$   
 thus  $2\pi f = 100\pi$  and  $f = 100\pi/2\pi = 50 \text{ Hz}$ .
- The frequency of the second term can similarly be determined from:  
 $\sin(\omega t) = \sin(200\pi t)$ .  
 Thus  $\omega = 200\pi$  but  $\omega = 2\pi f$   
 thus  $2\pi f = 200\pi$  and  $f = 200\pi/2\pi = 100 \text{ Hz}$ .  
 Thus the second term has twice the frequency of the fundamental, that is it is the second harmonic.
- The second harmonic has an amplitude ( $V_2$ ) of 50 V.
- Finally, since there is no phase angle included within the expression for the second harmonic (i.e. there is no  $\pm \phi_2$  term), the second harmonic component must be in phase with the fundamental.

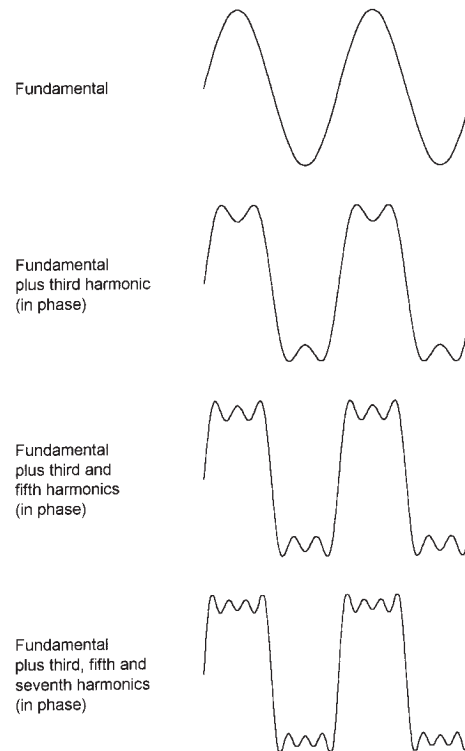


**Figure 3.3.99** *Fundamental and low-order harmonic components of a square wave*

### The square wave

The square wave can be created by adding a fundamental frequency to an infinite series of odd harmonics (i.e. the third, fifth, seventh, etc.). If the fundamental is at a frequency of  $f$ , then the third harmonic is at  $3f$ , the fifth is at  $5f$ , and so on.

The amplitude of the harmonics should decay in accordance with their harmonic order and they must all be in phase with the fundamental. Thus a square wave can be obtained from:

$$v = V \sin(\omega t) + V/3 \sin(3\omega t) + V/5 \sin(5\omega t) + V/7 \sin(7\omega t) + \dots$$

where  $V$  is the amplitude of the fundamental and  $\omega = 2\pi f$ .

Figure 3.3.99 shows this relationship graphically.

### The triangular wave

The triangular wave can similarly be created by adding a fundamental frequency to an infinite series of odd harmonics. However, in this case the harmonics should decay in accordance with the square of their harmonic order and they must alternate in phase so that the third, seventh, eleventh, etc. are in antiphase whilst the fifth, ninth, thirteenth, etc. are in phase with the fundamental. Thus a triangular wave can be obtained from:

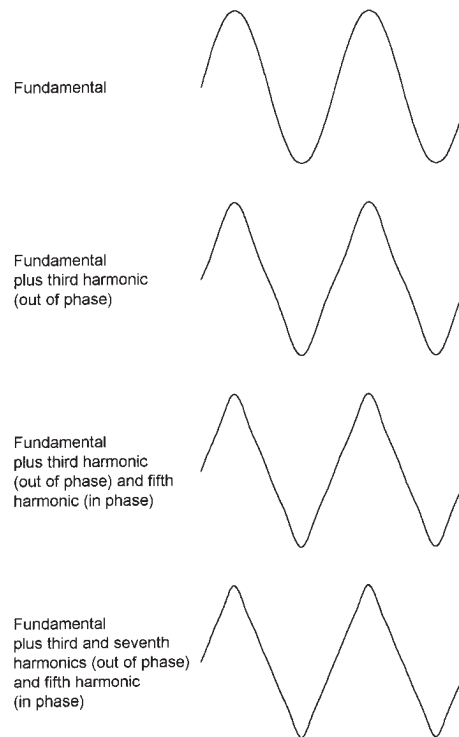
$$v = V \sin(\omega t) + V/9 \sin(3\omega t - \pi) + V/25 \sin(5\omega t) + V/49 \sin(7\omega t - \pi) + \dots$$

where  $V$  is the amplitude of the fundamental and  $\omega = 2\pi f$ .

Figure 3.3.100 shows this relationship graphically.



**Figure 3.3.100** *Fundamental and low-order harmonic components of a triangle wave*



### Question 3.3.4

A complex waveform is described by the equation:

$$v = 100 \sin(100\pi t) + 50 \sin(200\pi t - \pi/2).$$

Determine, graphically, the shape of the waveform and estimate the peak–peak voltage.

### Question 3.3.5

A complex waveform is described by the equation:

$$v = 100 \sin(100\pi t) + 250 \sin(200\pi t - \pi/2) + 450 \sin(300\pi t).$$

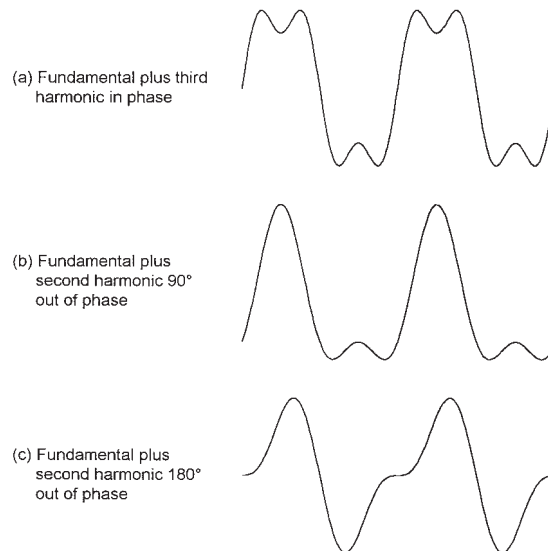
Determine, graphically, the shape of the waveform and estimate the peak–peak voltage.

### ***Effect of adding harmonics on overall waveshape***

The results of adding harmonics to a fundamental waveform with different phase relationships are listed below:

- (1) When odd harmonics are added to a fundamental waveform, regardless of their phase relative to the fundamental, *the positive and negative half-cycles will be similar* in shape and the resulting waveform will be symmetrical about the time axis – see Figure 3.3.101(a).

**Figure 3.3.101** Effects of harmonic components on waveshape



- (2) When even harmonics are added to a fundamental waveform with a phase shift other than 180°, the *positive and half-cycles will be dissimilar* in shape – see Figure 3.3.101(b).
- (3) When even harmonics are added to a fundamental waveform with a phase shift of 180°, the *positive half-cycles will be the mirror image of the negative half-cycles when reversed* – see Figure 3.3.101(c).

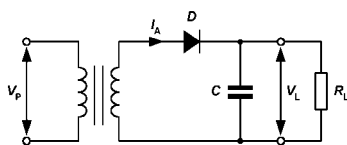


**Figure 3.3.102** A simple half-wave rectifier

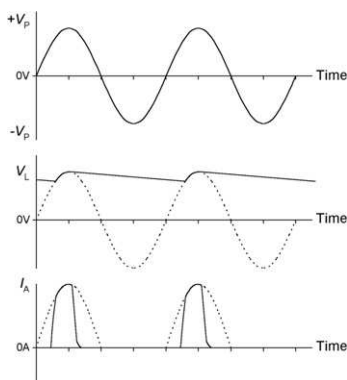


**Figure 3.3.103** Current and voltage waveforms for the simple half-wave rectifier

### Generation of harmonic components

Unwanted harmonic components may be produced in any system that incorporates non-linear circuit elements (such as diodes and transistors). Figure 3.3.102 shows the circuit diagram of a simple half-wave rectifier. The voltage and current waveforms for this circuit are shown in Figure 3.3.103. The important thing to note here is the shape of the current waveforms which comprises a series of relatively narrow pulses as the charge in the reservoir capacitor,  $C$ , is replenished at the crest of each positive-going half-cycle. These pulses of current are rich in unwanted harmonic content.

Harmonic components can also be generated in magnetic components (such as inductors and transformers) where the relationship between magnetic flux and applied current is non-linear (see Figure 3.3.104). In the case of a transformer, the voltage applied to the primary is sinusoidal and the flux (which lags the voltage by 90°) will also be sinusoidal. However, by virtue of the non-linearity of the magnetic flux/current characteristic, the current will not be purely sinusoidal but will contain a number of harmonic components, as shown in Figure 3.3.105.

Alternatively, there are a number of applications in which harmonic components are actually desirable. A typical application might be the frequency multiplier stage of a transmitter in which the non-linearity of an amplifier stage (operating in Class-C – see page 261) can be instrumental in producing an output which is rich in harmonic content. The desired harmonic can then be selected by means of a resonant circuit tuned to the desired multiple of the input frequency – see Figure 3.3.106.

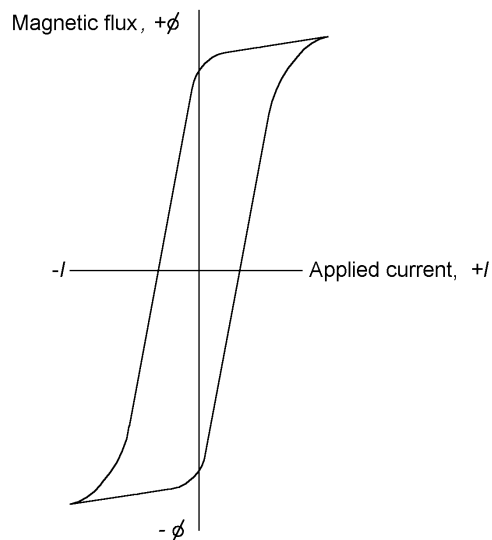


**Figure 3.3.104** Magnetic flux plotted against applied current for a transformer

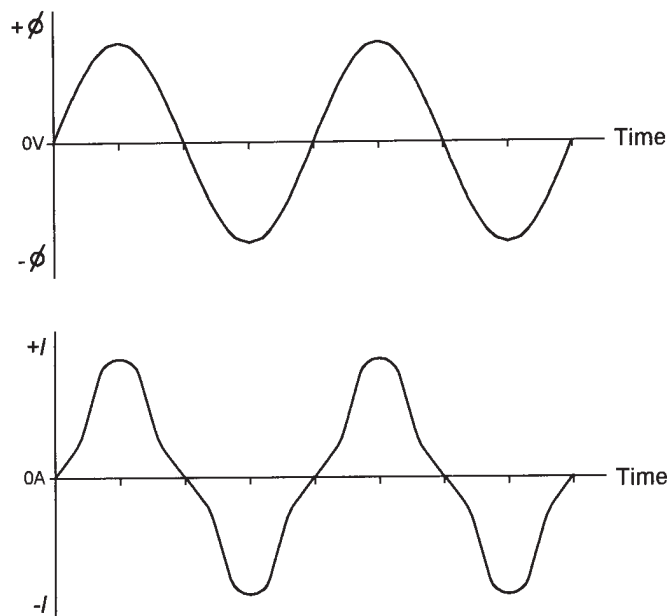


**Figure 3.3.105** Magnetic flux and applied current waveforms for a typical transformer

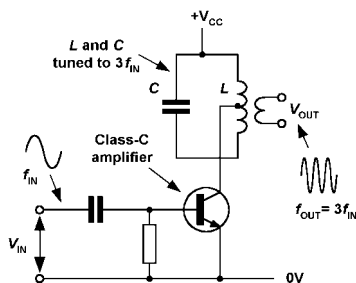


**Figure 3.3.106** A Class-C frequency multiplier stage

### Problems 3.3.8

- (1) A sinusoidal current is specified by the equation,  $i = 0.1 \sin(200\pi t + \pi/2)$  A. Determine:
  - (a) the peak value of the current;
  - (b) the frequency of the current;
  - (c) the phase angle of the current;
  - (d) the instantaneous value of the current at  $t = 5$  ms.
- (2) A complex waveform comprises a fundamental (sinusoidal) voltage which has a peak value of 36 V and a frequency of 400 Hz together with a third harmonic component having a peak value of 12 V leading the

fundamental by  $90^\circ$ . Write down an expression for the instantaneous value of the complex voltage, and use it to determine the value of voltage at 3 ms from the start of a cycle.

- (3) A complex waveform is given by the equation:

$$i = 10 \sin(100\pi t) + 3 \sin(300\pi t - \pi/2) + 1 \sin(500\pi t + \pi/2) \text{ mA.}$$

Determine:

- the amplitude of the fundamental;
- the frequency of the fundamental;
- the order of any harmonic components present;
- the amplitude of any harmonic components present;
- the phase angle of any harmonic components present (relative to the fundamental).

### 3.4 INFORMATION AND ENERGY CONTROL SYSTEMS

## Information systems

Much of electronics is concerned with the generation, transmission, reception, and storage of information using electronic circuits and devices. This information takes various forms including:

- speech or music (audio signals);
- visual images (video signals);
- alpha-numeric characters used in computer systems;
- numeric and other coded data in computer systems.

Electronic circuits provide different ways to process this information, including amplification of weak signals to usable levels, filtering to reduce noise and unwanted signal components, superimposition of a baseband signal onto a high-frequency carrier wave (*modulation*), recovery of a baseband signal from high-frequency carrier waves (*demodulation*), and logical operations such as those that take place in a computer.

The term *information system* applies to a system in which information is gathered, processed, and either displayed or stored for later analysis. *Information engineering* is used to describe the ‘engineering’ of information systems, that is the specification, design, construction, implementation, and maintenance of systems used for gathering, processing, and using information. So far, so good – but what actually is ‘information’?

To put it in its most simple terms, *information* is the way in which meaning is conveyed. Let us take an example to illustrate this. Imagine that you are in the driving seat of a car travelling on a motorway. You will be constantly receiving information. This information conveys meaning to you about the state of the car, of the motorway, and of other vehicles around you. You receive this information in various ways using the majority of your senses: sight, hearing, touch, and even smell. Signals are sent from sensory receptors via your nervous system to your brain which then uses the information presented to it in order to decide what to do next.

In the decision-making process, you might also make use of information that you previously stored in your memory. For example, you might recall that a long hill is around the next bend and this prompts you to move out to overtake a slow-moving vehicle well in advance.

In an engineering context, information usually takes the forms of analogue or digital signals (which may either be *baseband* signals or signals *modulated* onto a high-frequency carrier wave) transmitted by line, cable, optical fibre, point-to-point radio, microwave, or satellite.

### Information theory

*Information theory* was first developed in 1948 by the American electrical engineer Claude Shannon. Shannon was concerned with the measurement of information and, in particular, the information content of a message.

The term *information content* relates to the probability that a given message will contain completely new and unpredictable information and that this information is different from all other possible messages received. Essentially, when several possible messages are present in a communications channel:

- the highest value for the information content is assigned to the message that is the least probable;
- the least value for the information content is assigned to the message that is the most probable.

Shannon showed that the information content of a message is given by:

$$I = \log_2(1/p) \quad \text{or} \quad I = -\log_2(p)$$

where  $I$  is the information content and  $p$  is the probability of a message being transmitted.

#### Another view

A message that conveys very little new meaning has a very low information content. On the other hand, a message that conveys a large quantity of new information must have a very high information content.

## Mathematics in action

### Manipulating logarithms

We can prove the relationship,  $\log_2(1/p) = -\log_2(p)$  as follows:

Let  $y = \log_n(1/x)$ , thus  $y = \log_n(x^{-1})$ .

Now if  $y = \log_n(x^m)$ , then  $y = m \times \log_n(x)$ .

Hence  $y = \log_n(x^{-1}) = -1 \log_n(x)$ .

Thus  $y = \log_n(1/x) = -1 \log_n(x)$  and so  $\log_2(1/p) = -\log_2(p)$ .

#### Another view

Example 3.4.1 shows that the information content of this digital system is the same as the number of binary digits (or *bits*) used to convey the message. In other words, information content can be direct.

#### Example 3.4.1

A rotary position sensor has a resolution of  $22.5^\circ$  and provides its output in digital form using a four bit code. What is the information content of each of the following signals produced by this sensor?

- 0001
- x001
- xxx1

(where x = do not care, i.e. either 0 or 1).

- (a) There will be only one position that produces this code. Since there are 16 possible states (corresponding to codes 0000 to 1111) the probability of receiving this code (out of the 16 possible) is  $1/16$  or 0.0625. In this case:

$$\text{Now } I = \log_2(16) = 4.$$

- (b) Two positions can produce this code, 0001 and 1001. The probability of receiving either of these two codes (out of the 16 possible) is  $2/16$  or 0.125. In this case:

$$\text{Now } I = \log_2(1/0.125) = \log_2(8) = 3.$$

- (c) Eight positions can produce this code: 0001, 0011, 0101, 0111, 1001, 1011, 1101, and 1111. The probability of receiving any one of these eight codes (out of the 16 possible) is  $8/16$  or 0.5. In this case:

$$\text{Now } I = \log_2(1/0.5) = \log_2(2) = 1.$$

## ASCII code

The American Standard Code for Information Interchange (ASCII) is universally used in digital communications (see Table 3.4.1). The ASCII code is based on seven bits and these 128 different states are sufficient to convey upper and lower case letters, numbers, and basic punctuation as well as 32 *control characters* needed to provide information about how the data is to be transferred and how the characters are to be displayed on the printed page or on a visual display unit (VDU) screen. The control characters use the lowest five bits of the binary code and they are present when the two most-significant bits (MSBs) of the seven bit code are zero. Thus 0011011 is a control character (ESC) whereas 1001011 is a *printable character* (K).

Note that the decimal (Dec.) and hexadecimal (Hex.) values are often used to refer to ASCII characters rather than the binary codes that the digital circuits recognise directly.

ASCII characters are often transmitted along an eight bit data path where the MSB is *redundant* (this simply means that, as far as the standard ASCII code is concerned, the MSB conveys no meaning). Many computer and IT manufacturers (e.g. IBM, Olivetti, etc.) have made use of the redundant MSB to extend the basic character set with foreign accents and block graphic characters. These codes are referred to as an *extended character set*.

### Question 3.4.1

How many extra codes are available with an extended character set (i.e. a character set based on a full eight bits rather than a standard ASCII seven bit code). What is the information content of these additional codes?

**Table 3.4.1** ASCII code

| Dec. | Binary  | Hex. | ASCII | Dec. | Binary  | Hex. | ASCII | Dec. | Binary  | Hex. | ASCII |
|------|---------|------|-------|------|---------|------|-------|------|---------|------|-------|
| 0    | 0000000 | 00   | NUL   | 43   | 0101011 | 2B   | +     | 86   | 1010110 | 56   | V     |
| 1    | 0000001 | 01   | SOH   | 44   | 0101100 | 2C   | ,     | 87   | 1010111 | 57   | W     |
| 2    | 0000010 | 02   | STX   | 45   | 0101101 | 2D   | -     | 88   | 1011000 | 58   | X     |
| 3    | 0000011 | 03   | ETX   | 46   | 0101110 | 2E   | .     | 89   | 1011001 | 59   | Y     |
| 4    | 0000100 | 04   | EOT   | 47   | 0101111 | 2F   | /     | 90   | 1011010 | 5A   | Z     |
| 5    | 0000101 | 05   | ENQ   | 48   | 0110000 | 30   | 0     | 91   | 1011011 | 5B   | [     |
| 6    | 0000110 | 06   | ACK   | 49   | 0110001 | 31   | 1     | 92   | 1011100 | 5C   | \     |
| 7    | 0000111 | 07   | BEL   | 50   | 0110010 | 32   | 2     | 93   | 1011101 | 5D   | ]     |
| 8    | 0001000 | 08   | BS    | 51   | 0110011 | 33   | 3     | 94   | 1011110 | 5E   | ^     |
| 9    | 0001001 | 09   | HT    | 52   | 0110100 | 34   | 4     | 95   | 1011111 | 5F   | ␣     |
| 10   | 0001010 | 0A   | LF    | 53   | 0110101 | 35   | 5     | 96   | 1100000 | 60   | ␣     |
| 11   | 0001011 | 0B   | VT    | 54   | 0110110 | 36   | 6     | 97   | 1100001 | 61   | a     |
| 12   | 0001100 | 0C   | FF    | 55   | 0110111 | 37   | 7     | 98   | 1100010 | 62   | b     |
| 13   | 0001101 | 0D   | CR    | 56   | 0111000 | 38   | 8     | 99   | 1100011 | 63   | c     |
| 14   | 0001110 | 0E   | SO    | 57   | 0111001 | 39   | 9     | 100  | 1100100 | 64   | d     |
| 15   | 0001111 | 0F   | SI    | 58   | 0111010 | 3A   | :     | 101  | 1100101 | 65   | e     |
| 16   | 0010000 | 10   | DLE   | 59   | 0111011 | 3B   | ;     | 102  | 1100110 | 66   | f     |
| 17   | 0010001 | 11   | DC1   | 60   | 0111100 | 3C   | <     | 103  | 1100111 | 67   | g     |
| 18   | 0010010 | 12   | DC2   | 61   | 0111101 | 3D   | =     | 104  | 1101000 | 68   | h     |
| 19   | 0010011 | 13   | DC3   | 62   | 0111110 | 3E   | >     | 105  | 1101001 | 69   | i     |
| 20   | 0010100 | 14   | DC4   | 63   | 0111111 | 3F   | ?     | 106  | 1101010 | 6A   | j     |
| 21   | 0010101 | 15   | NAK   | 64   | 1000000 | 40   | @     | 107  | 1101011 | 6B   | k     |
| 22   | 0010110 | 16   | SYN   | 65   | 1000001 | 41   | A     | 108  | 1101100 | 6C   | l     |
| 23   | 0010111 | 17   | ETB   | 66   | 1000010 | 42   | B     | 109  | 1101101 | 6D   | m     |
| 24   | 0011000 | 18   | CAN   | 67   | 1000011 | 43   | C     | 110  | 1101110 | 6E   | n     |
| 25   | 0011001 | 19   | EM    | 68   | 1000100 | 44   | D     | 111  | 1101111 | 6F   | o     |
| 26   | 0011010 | 1A   | SUB   | 69   | 1000101 | 45   | E     | 112  | 1110000 | 70   | p     |
| 27   | 0011011 | 1B   | ESC   | 70   | 1000110 | 46   | F     | 113  | 1110001 | 71   | q     |
| 28   | 0011100 | 1C   | FS    | 71   | 1000111 | 47   | G     | 114  | 1110010 | 72   | r     |
| 29   | 0011101 | 1D   | GS    | 72   | 1001000 | 48   | H     | 115  | 1110011 | 73   | s     |
| 30   | 0011110 | 1E   | RS    | 73   | 1001001 | 49   | I     | 116  | 1110100 | 74   | t     |
| 31   | 0011111 | 1F   | US    | 74   | 1001010 | 4A   | J     | 117  | 1110101 | 75   | u     |
| 32   | 0100000 | 20   | SP    | 75   | 1001011 | 4B   | K     | 118  | 1110110 | 76   | v     |
| 33   | 0100001 | 21   | !     | 76   | 1001100 | 4C   | L     | 119  | 1110111 | 77   | w     |
| 34   | 0100010 | 22   | "     | 77   | 1001101 | 4D   | M     | 120  | 1111000 | 78   | x     |
| 35   | 0100011 | 23   | #     | 78   | 1001110 | 4E   | N     | 121  | 1111001 | 79   | y     |
| 36   | 0100100 | 24   | \$    | 79   | 1001111 | 4F   | O     | 122  | 1111010 | 7A   | z     |
| 37   | 0100101 | 25   | %     | 80   | 1010000 | 50   | P     | 123  | 1111011 | 7B   | {     |
| 38   | 0100110 | 26   | &     | 81   | 1010001 | 51   | Q     | 124  | 1111100 | 7C   |       |
| 39   | 0100111 | 27   | '     | 82   | 1010010 | 52   | R     | 125  | 1111101 | 7D   | }     |
| 40   | 0101000 | 28   | (     | 83   | 1010011 | 53   | S     | 126  | 1111110 | 7E   | ~     |
| 41   | 0101001 | 29   | )     | 84   | 1010100 | 54   | T     | 127  | 1111111 | 7F   | DEL   |
| 42   | 0101010 | 2A   | *     | 85   | 1010101 | 55   | U     |      |         |      |       |

**Figure 3.4.1** Basic elements of an information system

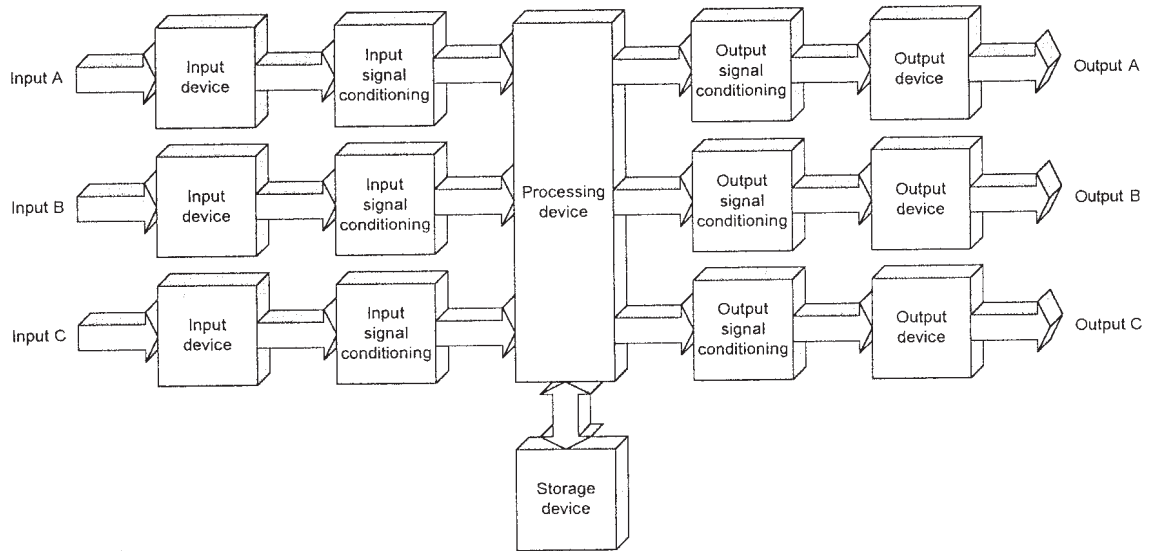
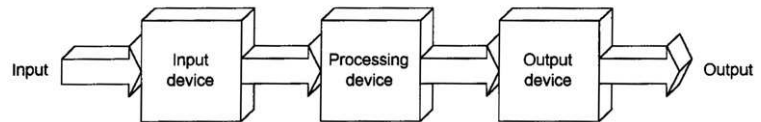


**Figure 3.4.2** A more complex information system

### Elements of an information system

The basic elements of an information system are shown in Figure 3.4.1. In this simple model, the *input device* is responsible for gathering or extracting information from the outside world. Input elements can be a wide variety of sensing devices and transducers, such as temperature sensors, strain gauges, proximity sensors, keyboards, microphones, etc.

The *processing device* can typically perform a variety of tasks, including making decisions and storing the information for later analysis. Processing elements can be microprocessors, microcomputers, programmed logic arrays (PLA), programmable logic controllers (PLC), etc.

The *output device* is responsible for displaying or communicating the information (or the results of the analysis performed by the processing element). Output elements can be cathode ray tube (CRT) displays, alpha-numeric displays, indicators, warning devices, etc.

Figure 3.4.2 shows an information system in which a number of input devices (*sensors* or *transducers*) produce electrical signals that are an analogy of the particular physical quantity present at the input. In order to standardise the signal levels that are presented to the processing device, the electrical signals from each input device are subjected to *signal conditioning*. For example, some input transducers produce continuously variable analogue outputs which will require conversion to equivalent digital signals that can be applied to a *digital processing device* (e.g. a microprocessor or microcomputer). This conversion is carried out using a *digital-to-analogue converter (DAC)*. Conversely, some output transducers produce discrete state digital outputs which will require conversion to equivalent analogue signals that can be applied to an analogue output



**Table 3.4.2** Some examples of information systems

|                            | <i>Plant monitoring system</i>                                         | <i>Satellite global positioning system</i>      | <i>Traffic information system</i>                                              |
|----------------------------|------------------------------------------------------------------------|-------------------------------------------------|--------------------------------------------------------------------------------|
| Input devices              | Temperature sensors, flow-rate sensors, etc.                           | UHF antenna, keypad                             | Magnetic sensors, TV cameras, keyboard                                         |
| Input signal conditioning  | ADCs, amplifiers, filters                                              | Frequency changing, amplification, demodulation | Optical fibre transmitters and receivers, microwave radio links                |
| Processing                 | Industrial microcomputer                                               | Dedicated microprocessor                        | Networked computer system                                                      |
| Storage                    | Hard disk drive with backup tape drive                                 | Semiconductor memory                            | Magneto-optical disk                                                           |
| Output signal conditioning | CRT interface card, alpha-numeric display driver, parallel output port | Liquid crystal display (LCD) driver             | Solid-state-switching devices, serial interface to remote information displays |
| Output devices             | CRT display, alpha-numeric display, piezo-electric warning device      | LCD screen                                      | Hazard warning indicators, traffic information displays                        |

device (e.g. a direct current (DC) motor). This conversion is carried out using an *analogue-to-digital converter (ADC)*. Other common examples of signal conditioning devices are level shifters, amplifiers, filters, attenuators, modulators, and demodulators.

### Examples of information systems

Typical examples of information systems are plant monitoring and control systems, networked computer systems, and many other systems based on microcomputers and microprocessors in which information (or *data*) is exchanged between ‘intelligent’ devices. Note also that, whether they are based on analogue or digital technology, all types of communication system can be classified as ‘information systems’.

Table 3.4.2 lists some examples of information systems.

### Question 3.4.2

A local radio station provides entertainment and information for its listening audience.

Treat this as an information system and identify the following elements:

Input devices, input signal conditioning, processing, storage, output signal conditioning, and output devices.

### Signals

Within an information system, information is conveyed by means of *signals*. These may take a number of different forms:

- switching between two states (i.e. on/off, open/closed or high/low);
- a varying DC voltage or current (the instantaneous voltage or current represents the value of the signal at the time of sampling);
- a waveform in which one or more properties are varied (e.g. frequency or amplitude).

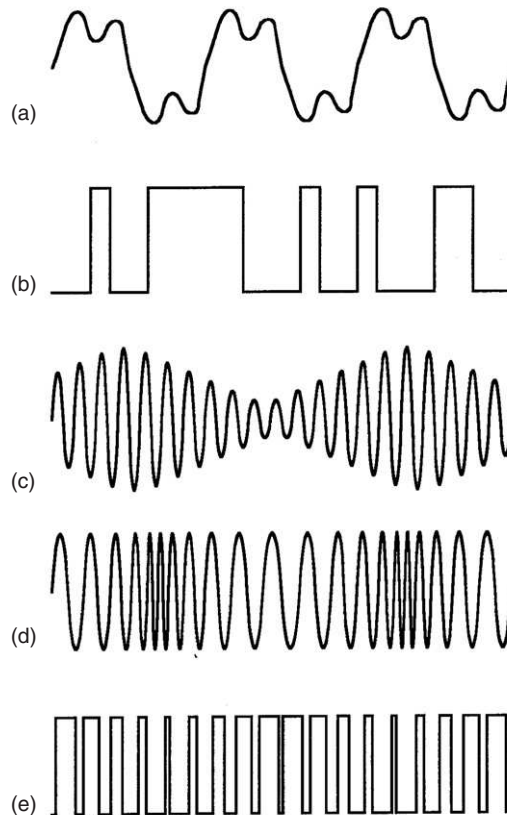


**Figure 3.4.3** Various forms of signal: (a) analogue signal, (b) digital signal, (c) an AM signal, (d) FM signal, and (e) PWM signal

Signals can be sent using wires, cables, optical, and radio links and they can be processed in various ways using amplifiers, modulators, filters, etc. Signals can also be classified as either *analogue* (continuously variable) or *digital* (based on discrete states).

Figure 3.4.3 shows various forms of signal, including baseband analogue and digital signals, amplitude modulated (AM) and frequency modulated (FM) signals, and a pulse width modulated (PWM) signal.

## Noise

Signals in information systems are susceptible to corruption by stray unwanted signals (both *natural* and *man-made*) which we refer to as *noise*. Noise may occur regularly (in which case it is referred to as *impulse noise*) or may occur randomly (in which case it is called *random noise*).

### Impulse noise

Noise generated by an electrical plant (such as transformers, rectifiers, and electrical machines) is, by its very nature, impulsive since it is likely to repeat at the frequency of the supply. Impulse noise can also be generated by natural causes such as electrical discharges in the atmosphere.

### Random noise

Random noise results from thermal agitation of electrons within electrical conductors and components such as resistors, transistors, and integrated circuits. Random noise in electrical conductors and components is very much dependent on temperature.

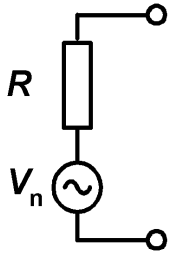


**Figure 3.4.4** Equivalent circuit of a noise source (Thévenin equivalent circuit)

### Effective noise voltage

The effective noise voltage generated within an electrical conductor is given by:

$$V_n = \sqrt{(4kTRf_w)}$$

where  $V_n$  is the effective noise voltage (in V),  $k$  is Boltzmann's constant ( $1.38 \times 10^{-23}$  J/K),  $T$  is the absolute temperature (in K),  $R$  is the resistance of the conductor ( $\Omega$ ), and  $f_w$  is the bandwidth in which the noise is measured.

The Thévenin equivalent circuit of a noise source is shown in Figure 3.4.4. Note that the voltage source ( $V_n$ ) appears in series with its associated resistance ( $R$ ).

#### Example 3.4.2

Determine the effective noise voltage produced in a bandwidth of 10 kHz by a conductor having a resistance of 10 k $\Omega$  at a temperature of 17°C:

$$V_n = \sqrt{(4kTRf_w)} = \sqrt{(4 \times 1.372 \times 10^{-23} \times (273 + 17) \times 10 \times 10^3 \times 10 \times 10^3)}$$

$$V_n = \sqrt{(4 \times 1.372 \times 10^{-15} \times 290)} = 2 \times 10^{-7} \times \sqrt{(1.372 \times 290)} \\ = 1.26 \mu\text{V}.$$

### Noise and its effect on small signals

In a communication system, thermally generated noise ultimately limits the ability of a system to respond to very small signals. What is of paramount importance here is the ratio of signal power to noise power. Let us illustrate this with an example: suppose the level of the signal applied to an information system is progressively reduced whilst the level of noise generated internally remains constant. A point will eventually be reached where the signal becomes lost in the noise. At this point, the signal will have become masked by the noise to the extent that it is no longer discernible. You might now be tempted to suggest that all we need to do is to increase the amount of amplification (gain) within the system to improve the level of the signal. Unfortunately, this will not work simply because the extra gain will increase the level of the noise as well as the signal. In fact, *no amount of amplification will result in an improvement* in the ratio of signal power to noise power.

### Signal-to-noise ratio

The signal-to-noise ratio in a system is normally expressed in decibels (dB) and is defined as:

$$S/N = 10 \log_{10}(P_{\text{signal}}/P_{\text{noise}}) \text{ dB}.$$

In practice, it is difficult to separate the signal present in a system from the noise. If, for example, you measure the output power produced by an

**Table 3.4.3** Effect of signal-to-noise ratio on signal quality

| (S + N)/N ratio (dB) | Effect                                                                                                          |
|----------------------|-----------------------------------------------------------------------------------------------------------------|
| 70                   | Sound produced by a good quality compact disk audio system (signal appears to be completely free from noise)    |
| 50                   | Sound produced by a good quality FM radio system when receiving a strong local signal (no noise detectable)     |
| 30                   | Noise not noticeable – sound quality produced by an AM radio when receiving a strong local signal               |
| 20                   | Noise becomes noticeable – sound quality not sufficient for the reproduction of music but acceptable for speech |
| 10                   | Signal noticeably degraded by noise – quality only sufficient for voice grade communication                     |
| 6                    | Very noisy voice communication channel – signal badly degraded by noise                                         |
| 0                    | Noise and signal powers equal – voice signal severely degraded by noise and of unacceptable quality             |
| –10                  | Voice signal completely lost in noise and unusable                                                              |

amplifier you will actually be measuring the signal power *together with* any noise that may be present. Hence a more practical quantity to measure is the (signal-plus-noise)-to-noise ratio. Furthermore, provided that the noise power is very much smaller than the signal power, there will not be very much difference between the signal-to-noise ratio and the (signal-plus-noise)-to-noise ratio.

Thus:

$$(S + N)/N = 10 \log_{10}(P_{\text{signal} + \text{noise}}/P_{\text{noise}}) \text{ dB.}$$

### Example 3.4.3

In the absence of a signal, the noise power present at the output of an amplifier is 10 mW. When a signal is present, the power increases to 2 W. Determine the (signal-plus-noise)-to-noise ratio expressed in decibels.

$$\text{Now } (S + N)/N = 10 \log_{10}(P_{\text{signal} + \text{noise}}/P_{\text{noise}}) \text{ dB}$$

$$\begin{aligned} \text{thus } (S + N)/N &= 10 \log_{10}(2 \text{ W}/10 \text{ mW}) = 10 \log_{10}(200) \\ &= 10 \times 2.3 = \mathbf{23 \text{ dB.}} \end{aligned}$$

Note that (signal-plus-noise)-to-noise ratio may also be determined from the *voltages* present at a point in a circuit, in which case:

$$(S + N)/N = 20 \log_{10}(V_{\text{signal} + \text{noise}}/V_{\text{noise}}) \text{ dB.}$$

As a rough guide, Table 3.4.3 will give you some idea of the effect of different values of signal-to-noise ratio on the quality of audio signals when noise is present. Alternatively, Figure 3.4.5 shows the effects of different degrees of noise on a visual image.

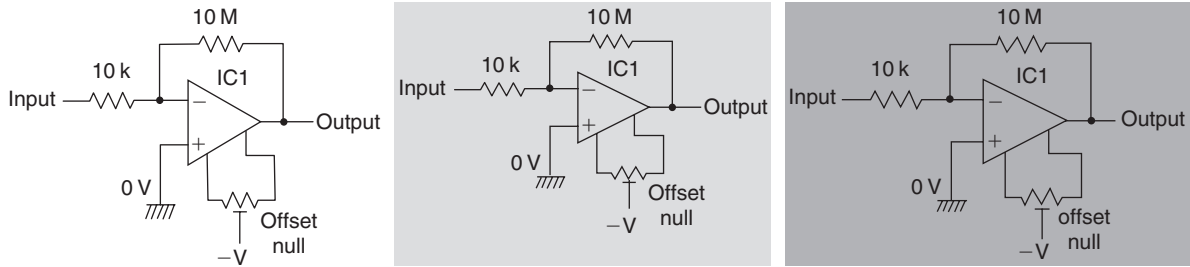


**Figure 3.4.5** Effect of various degrees of noise present with a visual image: (a) no noise present, (b) 20 dB signal-to-noise ratio, and (c) 10 dB signal-to-noise ratio

### Problems 3.4.1

- (1) A satellite receiver has an equivalent noise input resistance of  $50\ \Omega$ . If it operates with an effective bandwidth of 50 kHz at a temperature of  $20^\circ\text{C}$ , determine the effective noise voltage at the input of the receiver.
- (2) An instrumentation amplifier employs an input transducer which has an effective noise resistance of  $100\ \Omega$ . If the transducer operates at a temperature of  $-10^\circ\text{C}$  and it is connected to an amplifier with a gain of 2000 and a bandwidth of 2 kHz, determine the noise voltage produced at the output of the amplifier (you may assume that the amplifier itself is 'noise free').
- (3) In the absence of a signal, the Y-amplifier of an oscilloscope produces an output of 50 mV at a temperature of  $26^\circ\text{C}$ . If the amplifier has a bandwidth of 10 MHz and a voltage gain of 1000, determine the effective noise resistance at the input of the Y-amplifier.
- (4) An amplifier produces an output of 4 W when a signal is applied. When the input signal is disconnected, the output power falls to 10 mW. Determine the (signal-plus-noise)-to-noise ratio expressed in decibels.
- (5) The output signal from an audio pre-amplifier has a (signal-plus-noise)-to-noise ratio of 40 dB. If the output power is 10 W when a signal is present, determine the output power when the signal is absent.

## Amplifiers

### Types of amplifier

Many different types of amplifier are found in electronic circuits. Before we explain the operation of transistor amplifiers in detail, it is worth describing some of the types of amplifier used in electronic circuits.

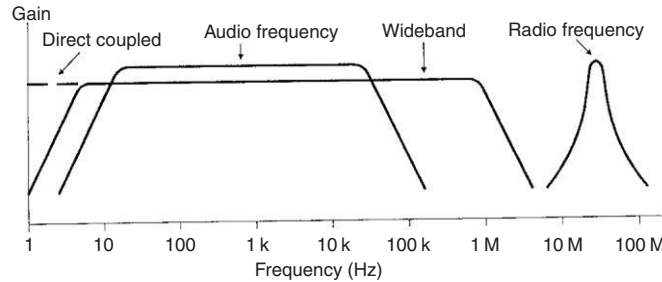
#### AC coupled amplifiers

In AC coupled amplifiers, stages are coupled together in such a way that DC levels are isolated and only the AC components of a signal are transferred from stage to stage.

#### DC coupled amplifiers

In DC (or direct) coupled amplifiers, stages are coupled together in such a way that stages are not isolated to DC potentials. Both AC and DC signal components are transferred from stage to stage.

**Figure 3.4.6** Frequency response of various types of amplifier



### Large-signal amplifiers

Large-signal amplifiers are designed to cater for appreciable voltage and/or current levels (typically from 1 to 100 V, or more).

### Small-signal amplifiers

Small-signal amplifiers are designed to cater for low level signals (normally less than 1 V and often much smaller).

### Audio frequency amplifiers

Audio frequency amplifiers operate in the band of frequencies that is normally associated with audio signals (e.g. 20 Hz to 20 kHz).

### Wideband amplifiers

Wideband amplifiers are capable of amplifying a very wide range of frequencies, typically from a few tens of Hz to several MHz (see Figure 3.4.6).

### Radio frequency amplifiers

Radio frequency amplifiers operate in the band of frequencies that is normally associated with radio signals (e.g. from 100 kHz to over 1 GHz). Note that it is desirable for amplifiers of this type to be *frequency selective* and thus their frequency response may be restricted to a relatively narrow band of frequencies (see Figure 3.4.6).

### Low-noise amplifiers

Low-noise amplifiers are designed so that they contribute negligible noise (signal disturbance) to the signal being amplified. These amplifiers are usually designed for use with very small-signal levels (usually less than 10 mV, or so).

## Gain

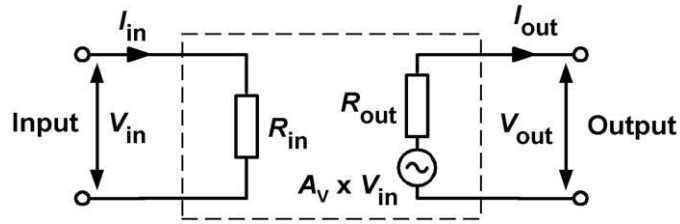
One of the most important parameters of an amplifier is the amount of amplification or *gain* that it provides. Gain is simply the ratio of output voltage to input voltage, output current to input current, or output power to input power (see Figure 3.4.7). These three ratios give, respectively, the voltage gain, current gain, and power gain. Thus:

$$\text{voltage gain, } A_v = V_{\text{out}}/V_{\text{in}}$$

$$\text{current gain, } A_i = I_{\text{out}}/I_{\text{in}}$$

$$\text{and power gain, } A_p = P_{\text{out}}/P_{\text{in}}.$$

**Figure 3.4.7** Definition of amplifier gain



Note that, since power is the product of current and voltage ( $P = IV$ ), we can infer that:

$$A_p = P_{out}/P_{in} = (I_{out}V_{out})/(I_{in}V_{in}) = (I_{out}/I_{in}) \times (V_{out}/V_{in}).$$

$$\text{Thus } A_p = A_i \times A_v.$$

### Example 3.4.4

An amplifier produces an output voltage of 2 V for an input of 50 mV. If the input and output currents in this condition are, respectively, 4 and 200 mA, determine:

- the voltage gain,
- the current gain,
- the power gain.

- The voltage gain is calculated from:

$$A_v = V_{out}/V_{in} = 2 \text{ V}/50 \text{ mV} = 40.$$

- The current gain is calculated from:

$$A_i = I_{out}/I_{in} = 200 \text{ mA}/4 \text{ mA} = 50.$$

- The power gain is calculated from:

$$\begin{aligned} A_p &= P_{out}/P_{in} = (2 \text{ V} \times 200 \text{ mA})/(50 \text{ mV} \times 4 \text{ mA}) \\ &= 0.4 \text{ W}/200 \mu\text{W} = \mathbf{2000}. \end{aligned}$$

(Note that  $A_p = A_i \times A_v = 50 \times 40 = 2000$ .)

Gain can also be expressed in decibels rather than as a straightforward ratio. The decibel is one tenth of a Bel. This, in turn, is defined as the logarithm (to the base 10) of the ratio of the two powers levels:

$$\text{Bels} = \log_{10}(P_{out}/P_{in}) \quad \text{and} \quad \text{dB} = 10 \log_{10}(P_{out}/P_{in}).$$

$$\text{Thus, power gain (in dB): } A_{p(\text{dB})} = 10 \log_{10}(P_{out}/P_{in}).$$

We use decibels (rather than Bels) simply because they are a more convenient unit to work with. Voltage and current gain can also be expressed in decibels:

$$\text{Voltage gain (in dB): } A_{v(\text{dB})} = 20 \log_{10}(V_{out}/V_{in})$$

$$\text{and, current gain (in dB): } A_{i(\text{dB})} = 20 \log_{10}(I_{out}/I_{in}).$$

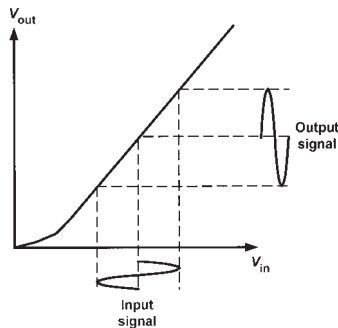


**Figure 3.4.8** Class A (linear operation)

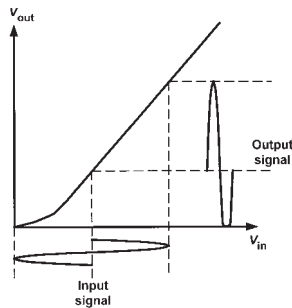


**Figure 3.4.9** Effect of reducing bias and increasing signal amplitude (the output waveform is no longer a faithful replica of the input)

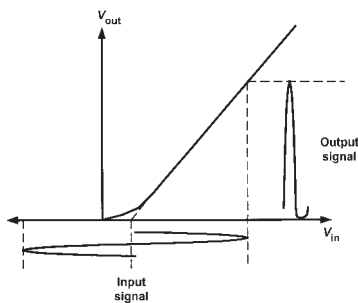


**Figure 3.4.10** Class AB operation (bias set at projected cut-off)

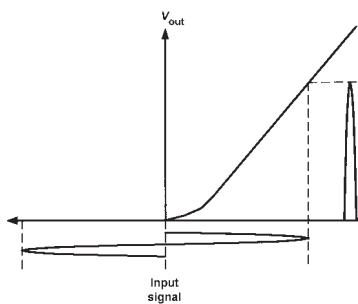


**Figure 3.4.11** Class B operation (no bias applied)

### Example 3.4.5

A telephone line amplifier provides a voltage gain of 18 dB. Determine the input signal voltage required to produce an output of 2 V.

$$\text{Now, } A_{V(\text{dB})} = 20 \log_{10}(V_{\text{out}}/V_{\text{in}}).$$

$$\text{Thus } A_{V(\text{dB})}/20 = \log_{10}(V_{\text{out}}/V_{\text{in}}).$$

$$V_{\text{out}}/V_{\text{in}} = \text{antilog}_{10}(A_{V(\text{dB})}/20)$$

$$\text{or } V_{\text{in}} = V_{\text{out}}/\text{antilog}_{10}(A_{V(\text{dB})}/20) = 2/\text{antilog}_{10}(18/20)$$

$$V_{\text{in}} = 2/\text{antilog}_{10}(18/20) = 2/\text{antilog}_{10}(0.9) = 2/7.94 = 0.252 \text{ V} = 252 \text{ mV}.$$

### Class of operation

A requirement of most amplifiers is that the output signal should be a faithful copy of the input signal, albeit somewhat larger in amplitude. Other types of amplifier are 'non-linear', in which case their input and output waveforms will not necessarily be similar. In practice, the degree of linearity provided by an amplifier can be affected by a number of factors including the amount of bias applied and the amplitude of the input signal. It is also worth noting that a linear amplifier will become non-linear when the applied input signal exceeds a threshold value.

Beyond this value the amplifier is said to be 'over-driven' and the output will become increasingly distorted if the input signal is further increased.

Amplifiers are usually designed to be operated with a particular value of *bias* supplied to the active devices (i.e. transistors). For linear operation, the active device(s) must be operated in the linear part of their transfer characteristic ( $V_{\text{out}}$  plotted against  $V_{\text{in}}$ ).

Figure 3.4.8 shows the input and output signals for an amplifier operating in linear mode. This form of operation is known as *Class A* and the bias point is adjusted to the mid-point of the linear part of the transfer characteristic. Furthermore, current will flow in the active devices used in a Class A amplifier during a complete cycle of the signal waveform. At no time does the current fall to zero.

Figure 3.4.9 shows the effect of moving the bias point down the transfer characteristic and, at the same time, increasing the amplitude of the input signal. From this, you should notice that the extreme negative portion of the output signal has become distorted. This effect arises from the non-linearity of the transfer characteristic that occurs near the origin (i.e. the zero point). Despite the obvious non-linearity in the output waveform, the active device(s) will conduct current during a complete cycle of the signal waveform.

Now consider what will happen if the bias is even further reduced and the amplitude of the input signal is further increased. Figure 3.4.10 shows the bias point set at the *projected cut-off point*. The negative-going portion of the output signal becomes cut-off (or *clipped*) and the active device(s) will cease to conduct for this part of the cycle. This mode of operation is known as *Class AB*.

Now let us consider what will happen if no bias at all is applied to the amplifier (see Figure 3.4.11). The output signal will only comprise a



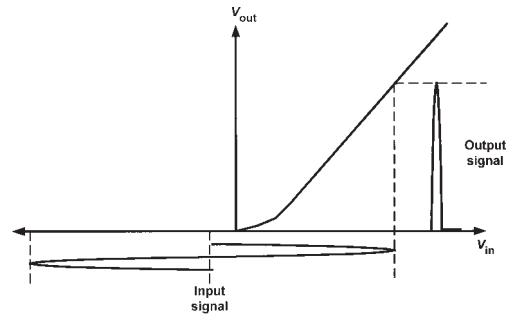


**Figure 3.4.12** Class B operation (bias is set beyond cut-off)

**Table 3.4.4** Classes of operation

| Class of operation | Bias point     | Conduction angle (°) | Efficiency (typical) (%) | Application (typical)      |
|--------------------|----------------|----------------------|--------------------------|----------------------------|
| A                  | Mid-point      | 360                  | 5–25                     | Linear audio amplifiers    |
| AB                 | Projected      | 210                  | 25–45                    | Push–pull audio amplifiers |
| B                  | At cut-off     | 180                  | 45–65                    | Push–pull audio amplifiers |
| C                  | Beyond cut-off | 120                  | 65–85                    | Radio frequency amplifiers |

series of positive-going half-cycles and the active device(s) will only be conducting during half-cycles of the waveform (i.e. they will only be operating 50% of the time). This mode of operation is known as *Class B* and is commonly used in push–pull power amplifiers where the two active devices in the output stage conduct on alternate half-cycles of the waveform.

Finally, there is one more class of operation to consider. The input and output waveforms for *Class C* operation are shown in Figure 3.4.12. Here the bias point is set at beyond the cut-off (zero) point and a very large input signal is applied. The output waveform will then comprise a series of quite sharp positive-going pulses. These pulses of current or voltage can be applied to a tuned circuit load in order to recreate a sinusoidal signal. In effect, the pulses will excite the tuned circuit and its inherent *flywheel action* will produce a sinusoidal output waveform. This mode of operation is only used in radio frequency power amplifiers which must operate at high levels of efficiency. Table 3.4.4 summarises the classes of operation used in amplifiers.

### Input and output resistance

Input resistance is the ratio of input voltage to input current and it is expressed in ohms. The input of an amplifier is normally purely resistive (i.e. any reactive component is negligible) in the middle of its working frequency range (i.e. the *mid-band*). In some cases, the reactance of the input may become appreciable (e.g. if a large value of stray capacitance

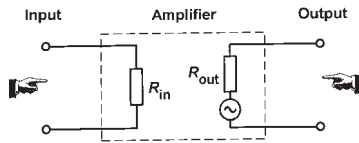


**Figure 3.4.13** Input and output resistance 'seen' looking into the input and output terminals, respectively

appears in parallel with the input resistance). In such cases we would refer to *input impedance* rather than input resistance.

Output resistance is the ratio of open-circuit output voltage to short-circuit output current and is measured in ohms. Note that this resistance is internal to the amplifier and should not be confused with the resistance of a load connected externally. As with input resistance, the output of an amplifier is normally purely resistive and we can safely ignore any reactive component. If this is not the case, we would once again refer to *output impedance* rather than output resistance.

The *equivalent circuit* in Figure 3.4.13 shows how the input and output resistances are 'seen', respectively, looking into the input and output terminals.

### Frequency response

The frequency response of an amplifier is usually specified in terms of the upper and lower *cut-off frequencies* of the amplifier. These frequencies are those at which the output power has dropped to 50% (otherwise known as the  $-3\text{ dB}$  points) or where the voltage gain has dropped to 70.7% of its mid-band value. Figures 3.4.14 and 3.4.15, respectively, show how the bandwidth can be expressed in terms of power or voltage. In either case, the cut-off frequencies ( $f_1$  and  $f_2$ ) and bandwidth are identical.

Note that, for response curves of this type, frequency is almost invariably plotted on a logarithmic scale.

#### Example 3.4.6

Determine the mid-band voltage gain and upper and lower cut-off frequencies for the amplifier whose frequency response is shown in Figure 3.4.16.

The mid-band voltage gain corresponds to the flat part of the frequency response characteristic. At this point the voltage gain reaches a maximum of 35 (see Figure 3.4.16).

The voltage gain at the cut-off frequencies can be calculated from:

$$A_{V \text{ cut-off}} = 0.707 \times A_{V \text{ max.}} = 0.707 \times 35 = \mathbf{24.7}$$

This value of gain intercepts the frequency response at  $f_1 = 57\text{ Hz}$  and  $f_2 = 590\text{ kHz}$  (see Figure 3.4.16).

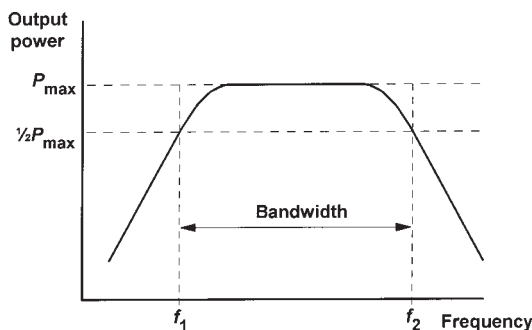


**Figure 3.4.14** Frequency response and bandwidth (output power plotted against frequency)

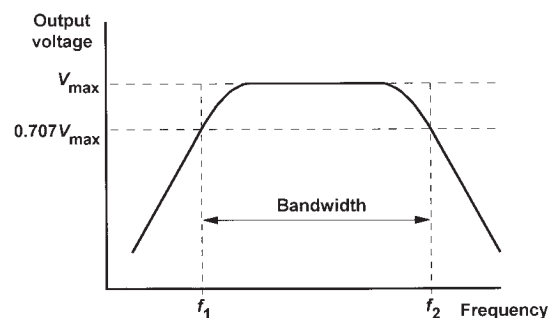


**Figure 3.4.15** Frequency response and bandwidth (output voltage plotted against frequency)

### Bandwidth

The bandwidth of an amplifier is usually taken as the difference between the upper and lower cut-off frequencies (i.e.  $f_2 - f_1$  in Figures 3.4.14 and 3.4.15). The bandwidth of an amplifier must be sufficient to accommodate the range of frequencies present within the signals that is to be presented with. Many signals contain *harmonic* components (i.e. signals at  $2f$ ,  $3f$ ,  $4f$ , etc. where  $f$  is the frequency of the *fundamental* signal). To perfectly reproduce a square wave, for example, it requires an amplifier with an infinite bandwidth (note that a square wave comprises an infinite series of harmonics). Clearly it is not possible to perfectly reproduce such a wave, but it does explain why it can be desirable for an amplifier's bandwidth to greatly exceed the highest signal frequency that is required to handle!

### Phase shift

Phase shift is the phase angle between the input and output voltages measured in degrees. The measurement is usually carried out in the mid-band where, for most amplifiers, the phase shift remains relatively constant. Note also that conventional single-stage transistor amplifiers usually provide phase shifts of either  $180^\circ$  or  $360^\circ$  (i.e.  $0^\circ$ ).

### Negative feedback

Many practical amplifiers use negative feedback in order to precisely control the gain, reduce distortion, and improve bandwidth by feeding back a small proportion of the output. The amount of feedback determines the overall (or *closed-loop*) gain. Because this form of feedback has the effect of reducing the overall gain of the circuit, this form of feedback is known as *negative feedback*. An alternative form of feedback, where the output is fed back in such a way as to reinforce the input (rather than to subtract from it) is known as *positive feedback*. This form of feedback is used in oscillator circuits (see page 267).

Figure 3.4.17 shows the block diagram of an amplifier stage with negative feedback applied. In this circuit, the proportion of the output voltage fed back to the input is given by  $\beta$  and the overall voltage gain will be given by:

$$\text{Overall gain} = V_{\text{out}}/V_{\text{in}}.$$

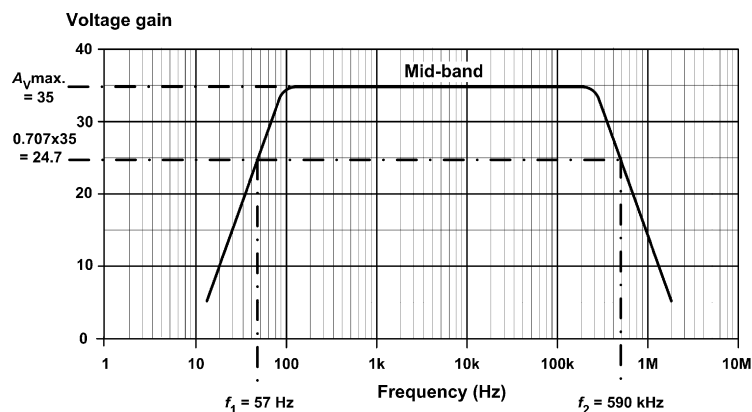


**Figure 3.4.16** See Example 3.4.6

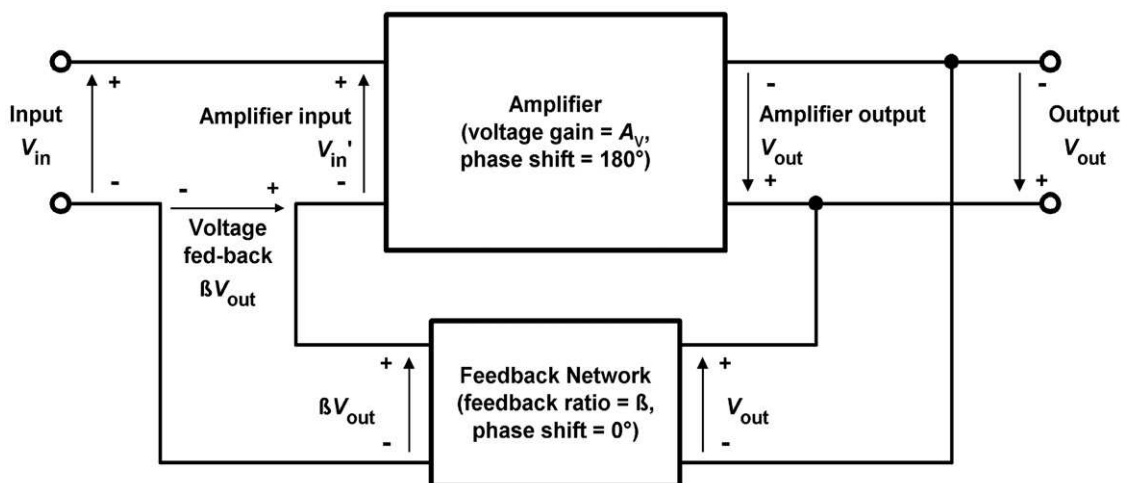


**Figure 3.4.17** Amplifier with negative feedback applied

Now  $V'_{in} = V_{in} - \beta V_{out}$  (by applying Kirchhoff's Voltage Law)

(note that the amplifier's input voltage has been *reduced* by applying negative feedback)

thus  $V_{in} = V'_{in} + \beta V_{out}$

and  $V_{out} = A_v \times V'_{in}$  ( $A_v$  is the *internal gain* of the amplifier).

Hence, overall gain =  $V_{out}/V_{in} = (A_v \times V'_{in})/(V'_{in} + \beta V_{out}) = A_v/(1 + \beta A_v)$ .

The overall gain with negative feedback applied will thus be less than the gain without feedback. Furthermore, if  $A_v$  is very large (as is the case with an operational amplifier) the overall gain with negative feedback applied will be given by:

overall gain (when  $A_v$  is very large) =  $1/\beta$ .

Note, also, that the *loop gain* of a feedback amplifier is defined as the product of  $\beta$  and  $A_v$ .

### Example 3.4.7

An amplifier with negative feedback applied has an open-loop voltage gain of 50 and one tenth of its output is fed back to the input (i.e.  $\beta = 0.1$ ). Determine the overall voltage gain with negative feedback applied.

With negative feedback applied the overall voltage gain will be given by:

$$A_v/(1 + \beta A_v) = 50/(1 + (0.1 \times 50)) = 50/(1 + 5) \\ = 50/6 = 8.33.$$

### Example 3.4.8

If, in Example 3.4.7, the amplifier's open-loop voltage gain increases by 20%, determine the percentage increase in overall voltage gain.

The new value of open-loop gain will be given by:

$$A_V = A_V + 0.2A_V = 50 + (0.2 \times 50) = 60.$$

The overall voltage gain with negative feedback will then be:

$$A_V/(1 + \beta A_V) = 60/(1 + (0.1 \times 60)) = 60/(1 + 6) \\ = 60/7 = 8.57.$$

The increase in overall voltage gain, expressed as a percentage will thus be:

$$(8.57 - 8.33)/8.33 \times 100\% = \mathbf{2.88\%}.$$

(Note that this example illustrates one of the important benefits of negative feedback in stabilising the overall gain of an amplifier stage.)

### Example 3.4.9

An integrated circuit that produces an open-loop gain of 100 is to be used as the basis of an amplifier stage having a precise voltage gain of 20. Determine the amount of feedback required.

Rearranging the formula,  $A_V/(1 + \beta A_V)$ , to make  $\beta$  the subject gives:

$$\beta = 1/A'_V - 1/A_V$$

where  $A'_V$  is the overall voltage gain with feedback applied, and  $A_V$  is the open-loop voltage gain.

$$\text{Thus } \beta = 1/20 - 1/100 = 0.05 - 0.01 = \mathbf{0.04}.$$

### Problems 3.4.2

- (1) An amplifier produces an output voltage of 6 V for an input of 150 mV. If the input and output currents in this condition are, respectively, 600  $\mu$ A and 120 mA, determine:
  - (a) the voltage gain,
  - (b) the current gain,
  - (c) the power gain.
- (2) An amplifier with negative feedback applied has an open-loop voltage gain of 40 and one fifth of its output is fed back to the input. Determine the overall voltage gain with negative feedback applied.
- (3) A manufacturer produces integrated circuit amplifier modules that have voltage gains which vary from 12 000 to 24 000 due to production tolerances. If the modules are to be used with 5% negative feedback applied, determine the maximum and minimum values of voltage gain.
- (4) An amplifier has an open-loop gain of 200. If this device is to be used in an amplifier stage having a precise voltage gain of 50, determine the amount of negative feedback that will be required.

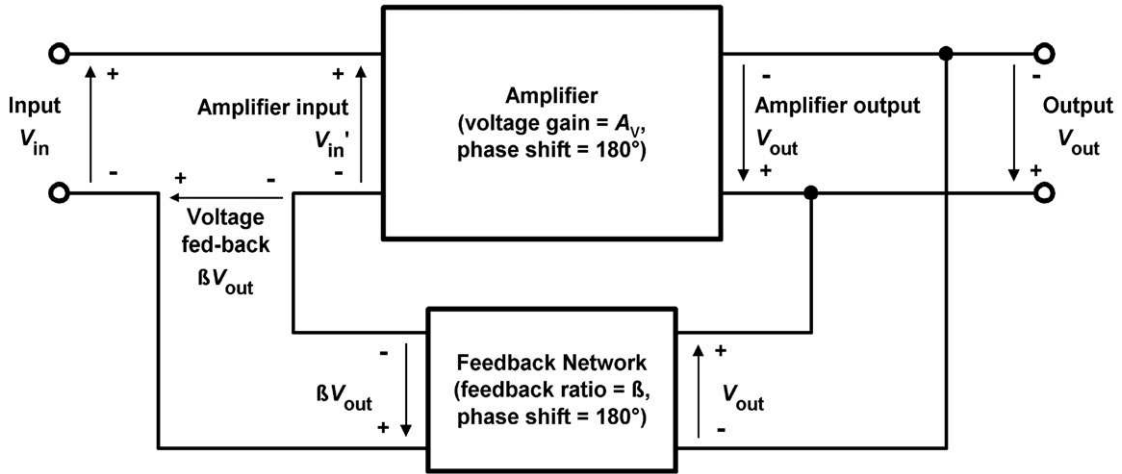


Figure 3.4.18 Amplifier with positive feedback applied

## Oscillators

This section describes some basic oscillator circuits that generate sine wave, square wave, and triangular waveforms. These circuits form the basis of clocks and timing arrangements as well as signal and function generators.

On page 165, we showed how negative feedback can be applied to an amplifier to form the basis of a stage which has a precisely controlled gain. An alternative form of feedback, where the output is fed back in such a way as to reinforce the input (rather than to subtract from it) is known as *positive feedback*. Figure 3.4.18 shows the block diagram of an amplifier stage with positive feedback applied. Note that the amplifier provides a phase shift of  $180^\circ$  and the feedback network provides a further  $180^\circ$ . Thus the overall phase shift is  $0^\circ$ . The overall voltage gain is given by:

$$\text{Overall gain} = V_{\text{out}}/V_{\text{in}}.$$

Now  $V'_{\text{in}} = V_{\text{in}} + \beta V_{\text{out}}$  (by applying Kirchhoff's Voltage Law)  
thus  $V_{\text{in}} = V'_{\text{in}} - \beta V_{\text{out}}$

and  $V_{\text{out}} = A_V \times V'_{\text{in}}$  ( $A_V$  is the *internal gain* of the amplifier).

Hence, overall gain  $= A_V \times V'_{\text{in}}/V_{\text{in}} - \beta V_{\text{out}} = A_V \times V'_{\text{in}}/V_{\text{in}} - \beta(A_V \times V'_{\text{in}})$ .

$$\text{Thus overall gain} = A_V/(1 - \beta A_V).$$

Now consider what will happen when the *loop gain* ( $\beta A_V$ ) approaches unity. The denominator ( $1 - \beta A_V$ ) will become close to zero. This will have the effect of increasing the overall gain, that is the overall gain with positive feedback applied will be *greater* than the gain without feedback.

It is worth illustrating this difficult concept using some practical figures. Assume that an amplifier has a gain of 9 and one tenth of its output is fed back to the input (i.e.  $\beta = 0.1$ ). In this case the loop gain ( $\beta \times A_V$ ) is 0.9.

With negative feedback applied the overall voltage gain will fall to:

$$A_V/(1 + \beta A_V) = 9/(1 + 0.1(9)) = 9/1.9 = 4.7.$$

With positive feedback applied the overall voltage gain will be:

$$A_V/(1 - \beta A_V) = 9/(1 - 0.1 \times 9) = 9/(1 - 0.9) = 9/0.1 = 90.$$

Now assume that the amplifier has a gain of 10 and, once again, one tenth of its output is fed back to the input (i.e.  $\beta = 0.1$ ). In this example the loop gain ( $\beta \times A_V$ ) is exactly 1.

With negative feedback applied, the overall voltage gain will fall to:

$$A_V/(1 + \beta A_V) = 10/(1 + 0.1 \times 10) = 10/2 = 5.$$

With positive feedback applied the overall voltage gain will be:

$$A_V/(1 - \beta A_V) = 10/(1 - 0.1 \times 10) = 10/(1 - 1) = 10/0 = \infty.$$

This simple example shows that a loop gain of unity (or larger) will result in infinite gain and an amplifier which is unstable. In fact, the amplifier will *oscillate* since any small disturbances will be amplified and result in an output. Clearly, as far as an amplifier is concerned, positive feedback may have an undesirable effect – instead of reducing the overall gain the effect is that of reinforcing any signal present and the output can build up into continuous oscillation if the loop gain is 1, or greater. To put this another way, an oscillator can simply be thought of as an amplifier that generates an output signal without the need for an input signal!

### Conditions for oscillation

From the foregoing we can deduce that the conditions for oscillation are:

- (a) the feedback must be positive (i.e. the signal fed back must arrive back in phase with the signal at the input);
- (b) the overall loop voltage gain must be greater than 1 (i.e. the amplifier's gain must be sufficient to overcome the losses associated with any frequency selective feedback network).

Hence, to create an oscillator we simply need an amplifier with sufficient gain to overcome the losses of the network that provides positive feedback. Assuming that the amplifier provides  $180^\circ$  phase shift, the frequency of oscillation will be that at which there is  $180^\circ$  phase shift in the feedback network. A number of circuits can be used to provide  $180^\circ$  phase shift, one of the simplest being a three-stage *C-R* ladder network. Alternatively, if the amplifier produces  $0^\circ$  phase shift, the circuit will oscillate at the frequency at which the feedback network produces  $0^\circ$  phase shift. In both cases, the essential point is that the feedback should be positive so that the output signal arrives back at the input in such a sense as to reinforce the original signal.

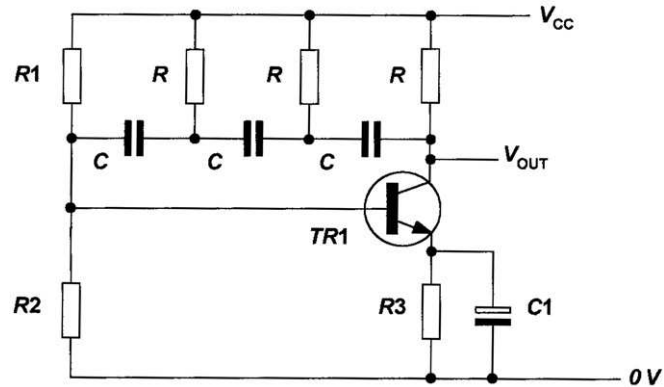
### Types of oscillator

Many different types of oscillator are used in electronic circuits. Oscillators tend to fall into two main categories:

- Oscillators that produce sine wave outputs (i.e. *sinusoidal oscillators*).
- Oscillators that produce square wave or rectangular pulse wave outputs (i.e. square wave oscillators). These types of oscillator are sometimes also known as *astable oscillators* (or *multivibrators*).

We shall briefly explore the operation of some of the most common types of sinusoidal and astable oscillator.

**Figure 3.4.19** Sine wave oscillator based on a three-stage C–R ladder network



### C–R ladder network oscillator

A simple phase-shift oscillator based on a three-stage C–R ladder network is shown in Figure 3.4.19. The total phase shift provided by the C–R ladder network (connected between collector and base) is  $180^\circ$  at the frequency of oscillation. The amplifier provides the other  $180^\circ$  phase shift in order to realise an overall phase shift of  $360^\circ$  or  $0^\circ$ .

The approximate frequency of oscillation of the circuit shown in Figure 3.4.19 will be given by:

$$f = 1/(2\pi\sqrt{6CR})$$

However, since the transistor's input and output resistance appears in parallel with the ladder network (i.e. the transistor *loads* the network), the actual frequency of operation will be slightly different. At the operating frequency, the loss associated with the ladder network is 29, thus the transistor amplifier must provide a gain of at least 29 in order for the circuit to oscillate.

#### Example 3.4.10

Determine the approximate frequency of oscillation of a three-stage ladder network oscillator where  $C = 10 \text{ nF}$  and  $R = 10 \text{ k}\Omega$ .

Using  $f = 1/(2\pi\sqrt{6CR})$

gives  $f = 1/(6.28 \times 2.45 \times 10 \times 10^{-9} \times 10 \times 10^3) \text{ Hz}$

thus  $f = 1/(6.28 \times 2.45 \times 10^{-4}) = 1 \times 10^4/15.386 = 650 \text{ Hz}$ .

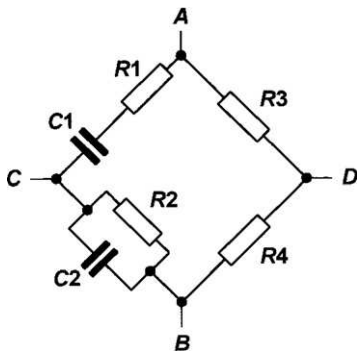


**Figure 3.4.20** A Wien bridge network

### Wien bridge oscillator

An alternative approach to providing the phase shift required is the use of a Wien bridge network (Figure 3.4.20). Like the C–R ladder, this network provides a phase shift which varies with frequency. The input signal is applied to A and B whilst the output is taken from C and D. At one particular frequency, the phase shift produced by the network will be exactly zero (i.e. the input and output signals will be in phase). If we connect the network to an amplifier producing  $0^\circ$  phase shift which has sufficient gain to overcome the losses of the Wien bridge, oscillation will result.



The minimum amplifier gain required to sustain oscillation is given by:

$$A_V = 1 + C1/C2 + R2/R1$$

and the frequency at which the phase shift will be zero is given by:

$$f = 1/2\pi\sqrt{(C1 \times C2 \times R1 \times R2)}.$$

In practice, we normally make  $R1 = R2$  and  $C1 = C2$  hence:

$$A_V = 1 + C/C + R/R = 1 + 1 + 1 = 3$$

and the frequency at which the phase shift will be zero is given by:

$$f = 1/(2\pi CR)$$

where  $R = R1 = R2$  and  $C = C1 = C2$ .

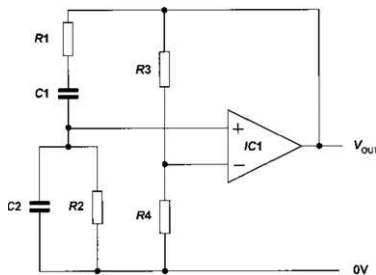


**Figure 3.4.21** Sine wave oscillator based on a Wien bridge (see Example 3.4.11)

### Example 3.4.11

Figure 3.4.21 shows the circuit of a Wien bridge oscillator based on an operational amplifier. If  $C1 = C2 = 100 \text{ nF}$ , determine the output frequencies produced by this arrangement (a) when  $R1 = R2 = 1 \text{ k}\Omega$  and (b) when  $R1 = R2 = 6 \text{ k}\Omega$ .

(a) When  $R1 = R2 = 1 \text{ k}\Omega$

$$f = 1/(2\pi CR)$$

where  $R = R1 = R2$  and  $C = C1 = C2$ .

$$\text{Thus } f = 1/(2\pi CR) = 1/(6.28 \times 100 \times 10^{-9} \times 1 \times 10^3) = 1/(6.28 \times 10^{-4})$$

$$\text{or } f = 10\,000/6.28 = 1592 \text{ Hz} = \mathbf{1.592 \text{ kHz}}.$$

(b) When  $R1 = R2 = 6 \text{ k}\Omega$

$$f = 1/(2\pi CR)$$

where  $R = R1 = R2$  and  $C = C1 = C2$ .

$$\text{Thus } f = 1/(2\pi CR) = 1/(6.28 \times 100 \times 10^{-9} \times 6 \times 10^3) = 1/37.68 \times 10^{-4} \text{ or } f = 10\,000/37.68 = \mathbf{265.4 \text{ Hz}}.$$

## Multivibrators

Multivibrators are a family of oscillator circuits that produce output waveforms consisting of one or more rectangular pulses. The term *multivibrator* simply originates from the fact that this type of waveform is rich in harmonics (i.e. ‘multiple vibrations’).

Multivibrators use regenerative (positive) feedback; the active devices present within the oscillator circuit being operated as switches, being alternately cut-off and driven into saturation.

The principal types of multivibrator are:

- *astable multivibrators* that provide a continuous train of pulses (these are sometimes also referred to as *free-running multivibrators*);
- *monostable multivibrators* that produce a single output pulse (they have one stable state and are thus sometimes also referred to as *one-shot* circuits);
- *bistable multivibrators* that have two stable states and require a trigger pulse or control signal to change from one state to another.

### The astable multivibrator

Figure 3.4.22 shows a classic form of astable multivibrator based on two transistors. Figure 3.4.23 shows how this circuit can be re-drawn in an arrangement that more closely resembles a two-stage common-emitter amplifier with its output connected back to its input.

In Figure 3.4.22 the values of the base resistors,  $R_3$  and  $R_4$ , are such that the sufficient base current will be available to completely saturate the respective transistor. The values of the collector load resistors,  $R_1$  and  $R_2$ , are very much smaller than  $R_3$  and  $R_4$ . When power is first applied to the circuit, assume that  $TR_2$  saturates before  $TR_1$  (in practice one transistor would always saturate before the other due to variations in component tolerances and transistor parameters).

As  $TR_2$  saturates, its collector voltage will fall rapidly from  $+V_{CC}$  to 0 V. This drop in voltage will be transferred to the base of  $TR_1$  via  $C_2$ . This negative-going voltage will ensure that  $TR_1$  is initially placed in the non-conducting state. As long as  $TR_1$  remains cut-off,  $TR_2$  will continue to be saturated. During this time,  $C_2$  will charge via  $R_4$  and  $TR_1$ 's base voltage will rise exponentially from  $-V_{CC}$  towards  $+V_{CC}$ . However,  $TR_1$ 's base voltage will not rise much above 0 V because as soon as it reaches +0.7 V (sufficient to cause base current to flow)  $TR_1$  will begin to conduct. As  $TR_1$  begins to turn on, its collector voltage will fall rapidly from  $+V_{CC}$  to 0 V. This fall in voltage is transferred to the base of  $TR_2$  via  $C_1$  and, as a consequence,  $TR_2$  will turn-off.  $C_1$  will then charge via  $R_3$  and  $TR_2$ 's base voltage will rise exponentially from  $-V_{CC}$  towards  $+V_{CC}$ . As before,  $TR_2$ 's base voltage will not rise much above 0 V because as soon as it reaches +0.7 V (sufficient to cause base current to flow)  $TR_2$  will start to conduct. The cycle is then repeated indefinitely.

The time for which the collector voltage of  $TR_2$  is low and  $TR_1$  is high ( $T_1$ ) will be determined by the time constant,  $R_4 \times C_2$ . Similarly, the time for which the collector voltage of  $TR_1$  is low and  $TR_2$  is high ( $T_2$ ) will be determined by the time constant,  $R_3 \times C_1$ .

The following approximate relationships apply:

$$T_1 = 0.7 C_2 R_4 \quad \text{and} \quad T_2 = 0.7 C_1 R_3.$$

Since one complete cycle of the output occurs in a time,  $T = T_1 + T_2$ , the periodic time of the output is given by:

$$T = 0.7(C_2 R_4 + C_1 R_3).$$

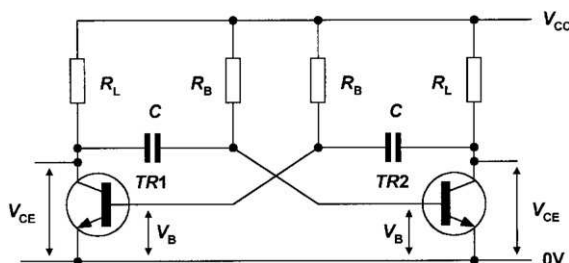


**Figure 3.4.22** Classic form of astable multivibrator using bipolar transistors

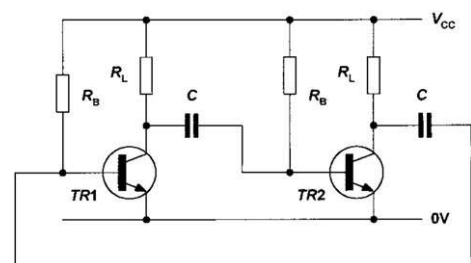


**Figure 3.4.23** Circuit of Figure 3.4.22 re-drawn to show two common-emitter stages

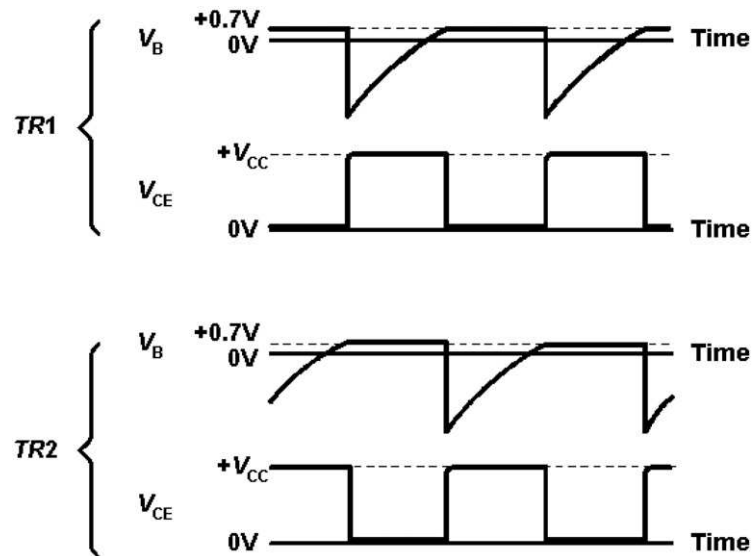


**Figure 3.4.24** Waveforms for the bipolar transistor astable multivibrator

Finally, we often require a symmetrical 'square wave' output where  $T_1 = T_2$ . To obtain such an output, we should make  $R_3 = R_4$  and  $C_1 = C_2$ , in which case the periodic time of the output will be given by:

$$T = 1.4 CR$$

where  $C = C_1 = C_2$  and  $R = R_3 = R_4$ .

Waveforms for the astable oscillator are shown in Figure 3.4.24.

### Example 3.4.12

The astable multivibrator in Figure 3.4.22 is required to produce a square wave output at 1 kHz. Determine suitable values for  $R_3$  and  $R_4$  if  $C_1$  and  $C_2$  are both 10 nF.

Since a square wave is required and  $C_1$  and  $C_2$  have identical values,  $R_3$  must be made equal to  $R_4$ . Now:

$$T = 1/f = 1/1 \times 10^3 = 1 \times 10^{-3} \text{ s.}$$

Rearranging  $T = 1.4 CR$  to make  $R$  the subject gives:

$$R = T/1.4C = 1 \times 10^{-3}/1.4 \times 10 \times 10^{-9} = 0.071 \times 10^6 \Omega = 71 \text{ k}\Omega.$$

### Other forms of astable oscillator

Figure 3.4.25 shows the circuit diagram of an alternative form of astable oscillator which produces a triangular output waveform. Operational amplifier,  $IC_1$  forms an integrating stage whilst  $IC_2$  is connected with positive feedback to ensure that oscillation takes place.

Assume that the output from  $IC_2$  is initially at or near the supply voltage,  $+V_{CC}$  and capacitor,  $C$ , is uncharged. The voltage at the output of  $IC_2$  will be passed, via  $R$ , to  $IC_1$ . Capacitor,  $C$ , will start to charge and the output voltage of  $IC_1$  will begin to fall. Eventually, the output voltage

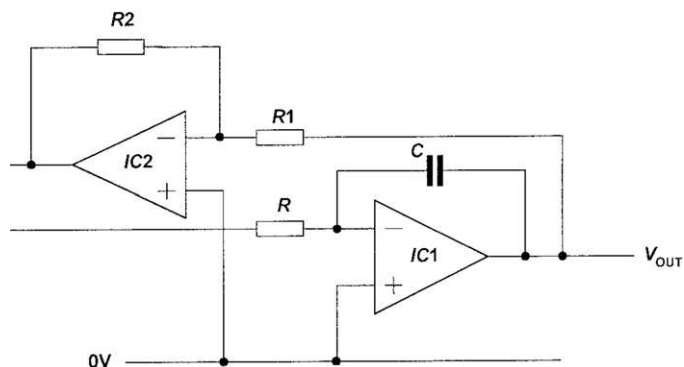


**Figure 3.4.25** Astable oscillator using operational amplifiers

will have fallen to a value that causes the polarity of the voltage at the non-inverting input of IC2 to change from positive to negative. At this point, the output of IC2 will rapidly fall to  $-V_{CC}$ . Again, this voltage will be passed, via  $R$ , to IC1. Capacitor,  $C$ , will then start to charge in the other direction and the output voltage of IC1 will begin to rise. Eventually, the output voltage will have risen to a value that causes the polarity of the non-inverting input of IC2 to revert to its original (positive) state and the cycle will continue indefinitely.

The *upper threshold voltage* (i.e. the maximum positive value for  $V_{out}$ ) will be given by:

$$V_{UT} = V_{CC} \times (R1/R2).$$

The *lower threshold voltage* (i.e. the maximum negative value for  $V_{out}$ ) will be given by:

$$V_{LT} = -V_{CC} \times (R1/R2).$$

### Single-stage astable oscillator

A simple form of astable oscillator producing a square wave output can be built using just one operational amplifier, as shown in Figure 3.4.26. This circuit employs positive feedback with the output fed back to the non-inverting input via the potential divider formed by  $R1$  and  $R2$ .

Assume that  $C$  is initially uncharged and the voltage at the inverting input is slightly less than the voltage at the non-inverting input. The output voltage will rise rapidly to  $+V_{CC}$  and the voltage at the inverting input will begin to rise exponentially as capacitor  $C$  charges through  $R$ . Eventually, the voltage at the inverting input will have reached a value that causes the voltage at the inverting input to exceed that present at the non-inverting input. At this point, the output voltage will rapidly fall to  $-V_{CC}$ . Capacitor,  $C$ , will then start to charge in the other direction and the voltage at the inverting input will begin to fall exponentially. Eventually, the voltage at the inverting input will have reached a value that causes the voltage at the inverting input to be less than that present at the non-inverting input. At this point, the output voltage will rise rapidly to  $+V_{CC}$  once again and the cycle will continue indefinitely.

The *upper threshold voltage* (i.e. the maximum positive value for the voltage at the inverting input) will be given by:

$$V_{UT} = V_{CC} \times R2/(R1 + R2).$$

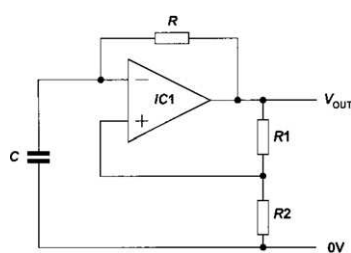


**Figure 3.4.26** Single-stage astable oscillator using an operational amplifier

The *lower threshold voltage* (i.e. the maximum negative value for the voltage at the inverting input) will be given by:

$$V_{LT} = -V_{CC} \times R1/(R1 + R2).$$

### Problems 3.4.3

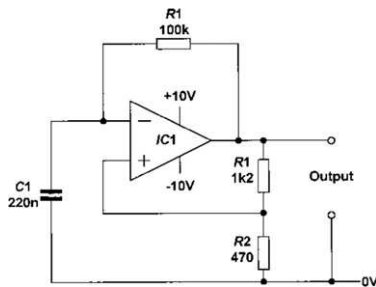


**Figure 3.4.27** See Problem 3.4.3

- (1) An amplifier with a gain of 8 has 10% of its output fed back to the input. Determine the gain of the stage (a) with negative feedback and (b) with positive feedback.
- (2) A phase-shift oscillator is to operate with an output at 1 kHz. If the oscillator is based on a three-stage ladder network, determine the required values of resistance if three capacitors of 10 nF are to be used.
- (3) A Wien bridge oscillator is based on the circuit shown in Figure 3.4.21 but  $R1$  and  $R2$  are replaced by a dual-gang potentiometer. If  $C1 = C2 = 22$  nF determine the values of  $R1$  and  $R2$  required to produce an output at exactly 400 Hz.
- (4) Determine the peak-peak output voltage produced by the circuit shown in Figure 3.4.27.
- (5) An astable multivibrator circuit is required to produce an asymmetrical rectangular output which has a period of 4 ms and is to be 'high' for 1 ms and 'low' for 3 ms. If the timing capacitors are both to be 10 nF, determine the values of the two timing resistors required.

### Transducers

Transducers are devices that convert energy in the form of sound, light, heat, etc. into an equivalent electrical signal, or vice versa. Before we go further, let us consider a couple of examples that you will already be familiar with. A loudspeaker is a device that converts low-frequency electrical current into sound. A thermocouple, on the other hand, is a transducer that converts temperature into voltage.

Transducers may be used both as system inputs and system outputs. From the two previous examples, it should be obvious that a loudspeaker is an *output transducer* designed for use in conjunction with an audio system. Whereas, a thermocouple is an *input transducer* designed for use in conjunction with a temperature control system.

Tables 3.4.5 and 3.4.6 provide examples of transducers that can be used to input and output three important physical quantities: sound, temperature, and angular position.

### Sensors

A *sensor* is a special kind of transducer that is used to generate an input signal to a measurement, instrumentation, or control system. The signal produced by a sensor is an *electrical analogy* of a physical quantity, such as distance, velocity, acceleration, temperature, pressure, light level, etc. The signals returned from a sensor, together with control inputs from the operator (where appropriate), will subsequently be used to determine the output from the system. The choice of sensor is governed by a number of factors including accuracy, resolution, cost, and physical size.

Sensors can be categorised as either *active* or *passive*. An active sensor *generates* a current or voltage output. A passive transducer requires

**Table 3.4.5** Some examples of input transducers

| Physical quantity       | Input transducer     | Notes                                                                                                                                                                                                                                                                                                      |
|-------------------------|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sound (pressure change) | Dynamic microphone   | Diaphragm attached to a coil is suspended in a magnetic field. Movement of the diaphragm causes current to be induced in the coil.                                                                                                                                                                         |
| Temperature             | Thermocouple         | Small e.m.f. generated at the junction between two dissimilar metals (e.g. copper and constantan). Requires reference junction and compensated cables for accurate measurement.                                                                                                                            |
| Angular position        | Rotary potentiometer | Fine wire resistive element is wound around a circular former. Slider attached to the control shaft makes contact with the resistive element. A stable DC voltage source is connected across the ends of the potentiometer. Voltage appearing at the slider will then be proportional to angular position. |

**Table 3.4.6** Some examples of output transducers

| Physical quantity       | Output transducer         | Notes                                                                                                                                                                                      |
|-------------------------|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sound (pressure change) | Loudspeaker               | Diaphragm attached to a coil is suspended in a magnetic field. Current in the coil causes movement of the diaphragm which alternately compresses and rarefies the air mass in front of it. |
| Temperature             | Resistive heating element | Metallic conductor is wound onto a ceramic or mica former. Current flowing in the conductor produces heat.                                                                                 |
| Angular position        | Stepper motor             | Multi-phase motor provides precise rotation in discrete steps of 15° (24 steps per revolution), 7.5° (48 steps per revolution), and 1.8° (200 steps per revolution).                       |

a source of current or voltage and it modifies this in some way (e.g. by virtue of a change in the sensor's resistance). The result may still be a voltage or current *but it is not generated by the sensor on its own*.

Sensors can also be classed as either *digital* or *analogue*. The output of a digital sensor can exist in only two discrete states, either 'on' or 'off', 'low' or 'high', 'logic 1' or 'logic 0', etc. The output of an analogue sensor can take any one of an infinite number of voltage or current levels. It is thus said to be *continuously variable*.

Table 3.4.7 shows some common types of sensor.

## Energy control systems

Energy control systems are systems that control or regulate the supply of energy to a plant or process. Unlike information systems, energy control systems usually operate continuously and are invariably automatic, requiring little or no human intervention under normal operating conditions.

The basic elements of an energy control system are shown in Figure 3.4.28. In this simple model, the *input setting device* is responsible for setting the required level of energy output. Input setting devices can be switches, rotary or linear potentiometers, keypads, or keyboards.

The *controlling element* or *controller* accepts the signal from the input device and, like the processing element of an information system, it can typically perform a variety of tasks, including making decisions and operating according to a stored algorithm (or *program*). Controllers can take the form of power amplifiers or power-switching devices together with logic circuits, microprocessors, microcontrollers, or PLC.

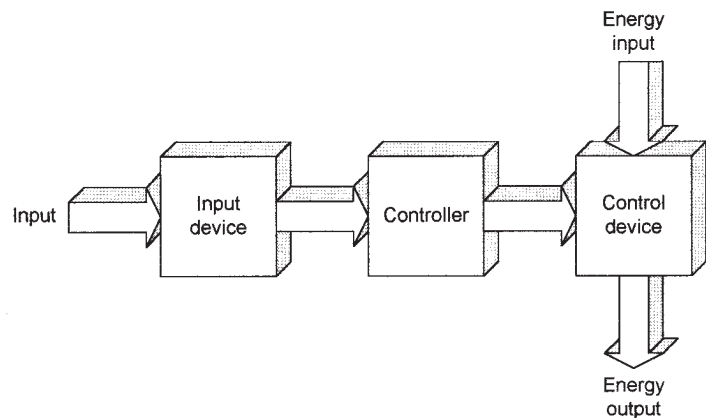
**Table 3.4.7** *Some common types of sensor*

| <i>Physical parameter</i> | <i>Type of sensor</i>                                                                                                  | <i>Notes</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------|------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Angular position          | Resistive rotary position sensor<br>Optical shaft encoder                                                              | Rotary track potentiometer with linear law produces analogue voltage proportional to angular position.<br>Encoded disk interposed between optical transmitter and receiver (infra-red light emitting diode (LED) and photodiode or phototransistor).                                                                                                                                                                                                                                                                                                                                                                   |
| Angular velocity          | Tachogenerator<br>Toothed rotor tachometer                                                                             | Small DC generator with linear output characteristic.<br>Analogue output voltage proportional to shaft speed.<br>Magnetic pick-up responds to the movement of a toothed ferrous disk. The pulse repetition frequency of the output is proportional to the angular velocity.                                                                                                                                                                                                                                                                                                                                            |
| Flow                      | Rotating vane flow sensor                                                                                              | Turbine rotor driven by fluid. Turbine interrupts infra-red beam. Pulse repetition frequency of output is proportional to flow rate.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Linear position           | Resistive linear position sensor<br>Linear variable differential transformer (LVDT)<br>Magnetic linear position sensor | Linear track potentiometer with linear law produces analogue voltage proportional to linear position. Limited linear range.<br>Miniature transformer with split secondary windings and moving core attached to a plunger. Requires AC excitation and phase-sensitive detector.<br>Magnetic pick-up responds to movement of a toothed ferrous track. Pulses are counted as the sensor that moves along the track.                                                                                                                                                                                                       |
| Light level               | Photocell<br>Light dependent resistor (LDR)<br>Photodiode<br>Phototransistor                                           | Voltage-generating device. The analogue output voltage produced is proportional to light level.<br>An analogue output voltage results from a change of resistance within a cadmium sulphide (CdS) sensing element. Usually connected as part of a potential divider or bridge.<br>Two-terminal device connected as a current source.<br>An analogue output voltage is developed across a series resistor of appropriate value.<br>Three-terminal device connected as a current source. An analogue output voltage is developed across a series resistor of appropriate value.                                          |
| Liquid level              | Float switch<br>Capacitive proximity switch<br>Diffuse scan proximity switch                                           | Simple switch element which operates when a particular level is detected.<br>Switching device which operates when a particular level is detected. Ineffective with some liquids.<br>Switching device which operates when a particular level is detected. Ineffective with some liquids.                                                                                                                                                                                                                                                                                                                                |
| Pressure                  | Microswitch pressure sensor<br>Differential pressure vacuum switch<br>Piezo-resistive pressure sensor                  | Microswitch fitted with actuator mechanism and range setting springs. Suitable for high-pressure applications.<br>Microswitch with actuator driven by a diaphragm. May be used to sense differential pressure. Alternatively, one chamber may be evacuated and the sensed pressure applied to a second input.<br>Pressure exerted on diaphragm causes changes of resistance in attached piezo-resistive transducers.<br>Transducers are usually arranged in the form of a four active element bridge which produces an analogue output voltage.                                                                        |
| Proximity                 | Reed switch<br>Inductive proximity switch<br>Capacitive proximity switch<br>Optical proximity switch                   | Reed switch and permanent magnet actuator. Only effective over short distances.<br>Target object modifies magnetic field generated by the sensor. Only suitable for metals (non-ferrous metals with reduced sensitivity).<br>Target object modifies electric field generated by the sensor. Suitable for metals, plastics, wood, and some liquids and powders.<br>Available in diffuse and through scan types. Diffuse scan types require reflective targets. Both types employ optical transmitters and receivers (usually infra-red emitting LEDs and photodiodes or phototransistors). Digital input port required. |
| Strain                    | Resistive strain gauge                                                                                                 | Foil type resistive element with polyester backing for attachment to body under stress. Normally connected in full bridge configuration with temperature-compensating gauges to provide an analogue output voltage.                                                                                                                                                                                                                                                                                                                                                                                                    |

*(continued)*

**Table 3.4.7** (Continued)

| Physical parameter | Type of sensor                   | Notes                                                                                                                                                                                                               |
|--------------------|----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Temperature        | Semiconductor strain gauge       | Piezo-resistive elements provide greater outputs than comparable resistive foil types. More prone to temperature changes and also inherently non-linear.                                                            |
|                    | Thermocouple                     | Small e.m.f. generated by a junction between two dissimilar metals. For accurate measurement, requires compensated connecting cables and specialised interface.                                                     |
|                    | Thermistor                       | Usually connected as part of a potential divider or bridge. An analogue output voltage results from resistance changes within the sensing element.                                                                  |
| Weight             | Semiconductor temperature sensor | Two-terminal device connected as a current source. An analogue output voltage is developed across a series resistor of appropriate value.                                                                           |
|                    | Load cell                        | Usually comprises four strain gauges attached to a metal frame. This assembly is then loaded and the analogue output voltage produced is proportional to the weight of the load.                                    |
| Vibration          | Electromagnetic vibration sensor | Permanent magnet seismic mass suspended by springs within a cylindrical coil. The frequency and amplitude of the analogue output voltage are respectively proportional to the frequency and amplitude of vibration. |

**Figure 3.4.28** Basic elements of an energy control system

The *controlled device* is a specialised output device that is capable of regulating the energy input in order to provide the required energy output. Typical examples of control devices are motorised valves, pumps, heaters, etc. However, when an electrical supply is being controlled, the control device can be a relay, or a solid-state-switching device such as a thyristor, triac, or a power transistor (see pages 282–284).

### Automatic energy control systems

Figure 3.4.29 shows an energy control system which employs negative feedback in order to automatically regulate its output. This shows how the energy level at the output is determined by means of a *sensor* and, after signal conditioning, is compared with the desired signal level provided by the input device. The comparator produces an output which is the difference between the input signal and the signal that has been fed back from the output. This output signal is often referred to as the *error signal* and, when it is present, it indicates that there is a difference



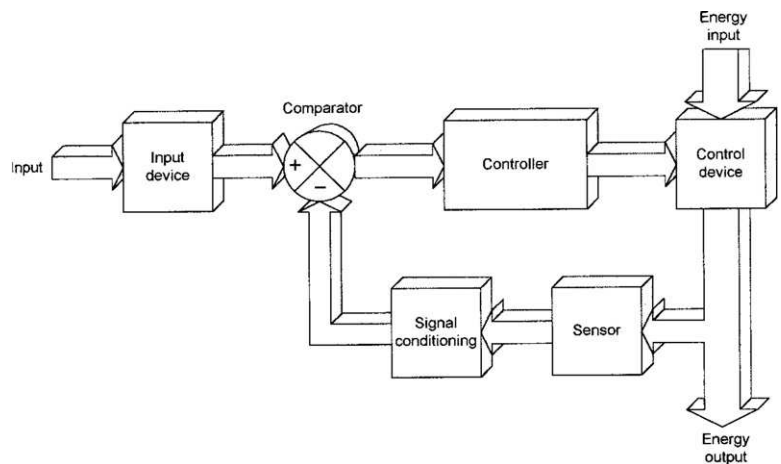


**Figure 3.4.29** An automatic energy control system

**Table 3.4.8** Some examples of energy control systems

|                       | Industrial heating system | Hydraulic ram                     | Belt conveyor                                    |
|-----------------------|---------------------------|-----------------------------------|--------------------------------------------------|
| Input setting device  | Variable resistor         | Keypad                            | Switches                                         |
| Comparator            | Operational amplifier     |                                   | Microcontroller<br>Programmable logic controller |
| Controller            | Power amplifier           | Servo driver                      | Programmable logic controller                    |
| Controlled device     | Motorised gas valve       | Servo valve                       | Optically coupled triacs (three-phase)           |
| Output sensing device | Thermistor                | Resistive displacement transducer | Infra-red proximity sensors                      |
| Signal conditioning   | None                      | ADC and DAC                       | Amplifier/level shifter                          |

between the desired output and the actual output from the system. The error signal is fed to the controller, which in turn acts upon the control device. As with earlier systems, there may be a need for signal conditioning in order to regularise signal levels and to provide digital-to-analogue conversion or analogue-to-digital conversion, as appropriate.

**Examples of energy control systems**

Typical examples of energy control systems are shown in Table 3.4.8.

**Question 3.4.3**

A gravel washing plant uses a constant head water supply which is pumped from a lake into a concrete reservoir. Treat this as an automatic energy control system and identify the following elements: input setting device, comparator, controller, controlled device, output sensing device, signal conditioning.

## Power control

Many engineering processes involve the control of appreciable levels of electrical energy. The electronic circuits that provide this control are often referred to as *power controllers*. Power controllers are designed to switch high values of current and/or voltage. Since we are usually concerned with average power (rather than instantaneous values), the supply of energy is often controlled by rapidly interrupting an electrical supply (i.e. switching) rather than using a resistive control device (which would dissipate unwanted energy as heat). This process is best illustrated by taking an example.

### Example 3.4.13

Two alternative forms of energy controller are shown in Figure 3.4.30(a) and (b). The circuit in Figure 3.4.30(a) makes use of a resistive control device whilst that in Figure 3.4.30(b) uses a switching device. For the purpose of this example we will assume that the switching device is perfect (i.e. that its 'on' resistance is zero, its 'off' resistance is infinite and that it changes from 'on' to 'off' in zero time).

*For the resistive control device (Figure 3.4.30(a)):*

Now  $R_L = 1 \Omega$ , thus the load voltage ( $V_L$ ) and load current ( $I_L$ ) for a load power ( $P_L$ ) of 100 W will be given by:

$$V_L = \sqrt{P \times R_L} = \sqrt{100 \times 1} = \sqrt{100} = 10 \text{ V.}$$

$$I_L = P_L / V_L = 100 / 10 = 10 \text{ A.}$$

The power supplied to the circuit ( $P_S$ ) will be given by:

$$P_S = V_S \times I_L = 50 \times 10 = 500 \text{ W.}$$

The power dissipated in the resistive control device ( $P_R$ ) will be:

$$P_R = P_S - P_L = 500 - 100 = 400 \text{ W.}$$

The efficiency of the energy control system under these conditions will be:

$$\eta = (P_L / P_S) \times 100\% = (100 / 500) \times 100\% = 20\%.$$

*For the switching control device (Figure 3.4.30(b)):*

When the switch is closed (see Figure 3.4.30(c)) the full supply will appear across the load. In this condition the peak power supplied to the load will be:

$$P_{PK} = V_S^2 / R_L = 50^2 / 1 = 2500 \text{ W.}$$

The waveform showing the instantaneous power delivered to the load is shown in Figure 3.4.30(c). From this you should note that, from the area under the graph, the peak power ( $P_{PK}$ ) is related to the average power in the load ( $P_{L(AV)}$ ) as follows:

$$P_{PK} \times t_{on} = P_{L(AV)} \times (t_{on} + t_{off})$$

$$\text{thus } P_{L(AV)} = P_{PK} \times t_{on} / (t_{on} + t_{off}) = \tau \times P_{PK}$$

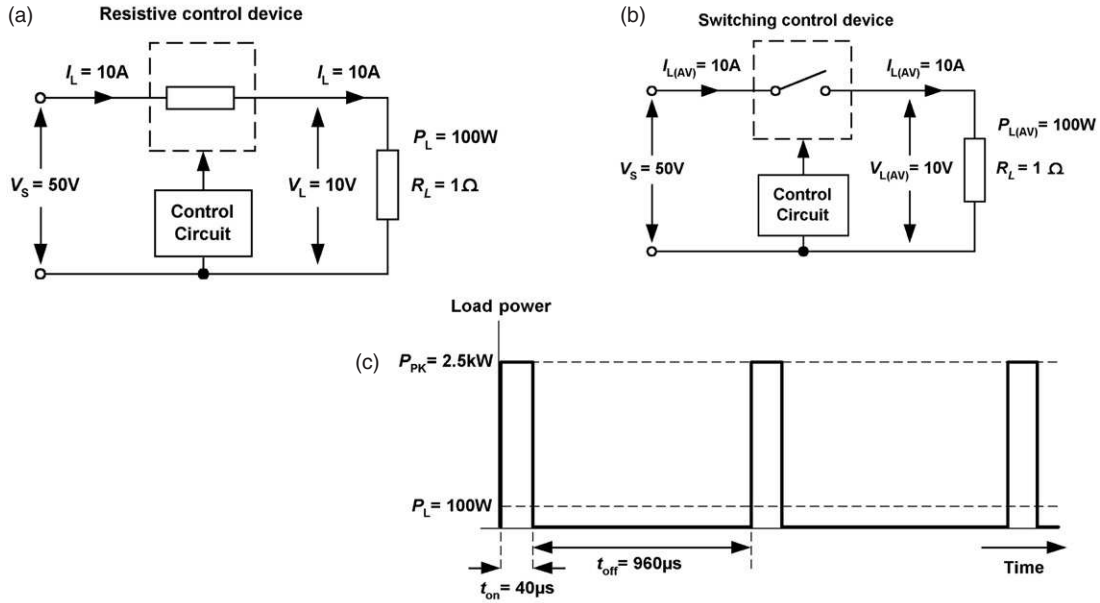


**Figure 3.4.30** Two alternative forms of electrical energy controller: (a) using a resistive control device, (b) using a switching control device, and (c) variation of load power with time

where  $\tau$  is the duty cycle of the control signal and  $\tau = t_{on}/(t_{on} + t_{off})$ .

Now  $P_{L(AV)} = 100W$  and  $P_{PK} = 2500W$

thus  $\tau = P_{L(AV)}/P_{PK} = 100/2500 = 0.04$ .

Assuming that the switching frequency is 1 kHz (i.e. a complete on/off cycle takes 1 ms), the duty cycle for this control signal is thus 1 ms (i.e.  $t_{on} + t_{off} = 1ms$ ) hence:

$$\tau = t_{on}/(t_{on} + t_{off}) \quad \text{or}$$

$$t_{on} = \tau \times (t_{on} + t_{off}) = 0.04 \times 1ms = 40\mu s$$

$$\text{and } t_{off} = 1ms - 40\mu s = 960\mu s.$$

Thus, to produce an average load power of 100 W using a 1 kHz switching frequency, the 'on' time ( $t_{on}$ ) must be 40  $\mu s$  and the 'off' time ( $t_{off}$ ) must be 960  $\mu s$ .

Note that, since the switching control device has been assumed to be perfect, no power will be dissipated in it and thus the average power supplied to the circuit,  $P_s$ , will be the same as that supplied to the load (i.e. 100 W). The circuit thus has an efficiency which approaches 100%.

### Power-switching devices

The switching device employed in a power control circuit is crucial to the successful operation of the circuit. The basic requirements for a power-switching device are as follows:

- (a) The switching device should only require a very small input current in order to control a very much larger current (flowing through the load).

- (b) The switching device should operate very rapidly (i.e. the time between 'on' and 'off' states should be negligible).
- (c) During conduction, the switching device should be capable of continuously carrying the rated load current. It should also be capable of handling momentary surge currents (which may be an order of magnitude greater than the continuously rated load current).
- (d) There should be minimal power dissipation within the switching device (it should exhibit a very low resistance in the on/conducting state and a very high resistance in the off/non-conducting state).
- (e) In the non-conducting state, the switching device should be capable of continuously sustaining the peak value of rated supply voltage. It should also be capable of coping with momentary surge voltages (which may exceed the normal peak supply voltage by a factor of two, or more).

## Relays

The traditional method of switching current through a load which requires isolation from the controlling circuit involves the use of an electromechanical relay. Relays offer many of the desirable characteristics of an 'ideal'-switching device (notably a very low 'on' resistance and virtually infinite 'off' resistance coupled with a coil to contact breakdown voltage which is usually in excess of several kV). Unfortunately, relays also have several shortcomings which prevent their use in a number of applications. These shortcomings include:

- the *contact bounce* which occurs during the transitory state which exists when the contacts make contact;
- *arcing* (ionisation of the air between the contacts) which may occur when the contacts break and which can result in the generation of heat which can literally burn out the contact surfaces and which can produce a large amount of radio frequency interference;
- the need for regular inspection and routine maintenance with periodic replacement of relays when contacts wear out.

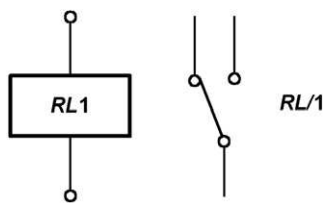


**Figure 3.4.31** Relay symbol

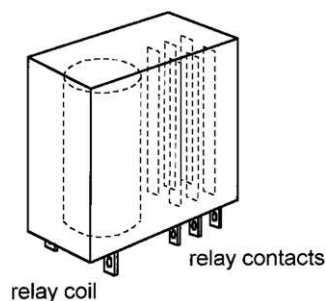


**Figure 3.4.32** A typical relay package

A typical electromechanical relay may be rated for around 1 000 000 operations, or more. To put this into context, if operated once every minute, the contact set on such a relay can be expected to give satisfactory operation for a period of about 2 years. It is important to note, however, that electromechanical relays are prone to *both* mechanical and electrical failure (the latter being more prevalent if the device is operated at, or near, its maximum rating).

Despite this, in simple low-speed 'on/off'-switching applications, the conventional electromechanical relay can still provide a cost-effective solution to controlling currents of up to about 10 A or more, at voltages of up to 250 V AC and 100 V DC. Furthermore (unlike solid-state-switching devices), relays are available with a variety of different contact sets, including single-pole on/off-switching, single-pole changeover (SPCO), double-pole changeover (DPCO), and four-pole changeover (4PCO). The coils which provide the necessary magnetic flux to operate a relay are available for operation on a variety of voltages between 5 and 115 V DC.

The symbol for a relay is shown in Figure 3.4.31 while Figure 3.4.32 shows the typical packaging used for such a device.

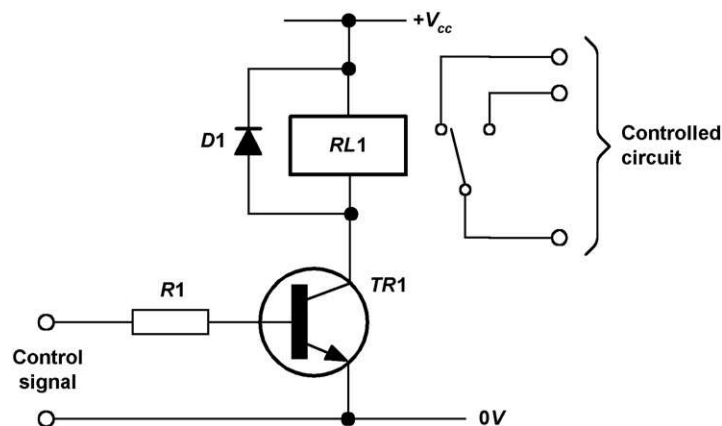


**Figure 3.4.33** On/off power-switching arrangement based on a relay

A typical on/off power-switching arrangement based on a relay is shown in Figure 3.4.33. The diode connected across the relay coil is designed to absorb the *back* electromotive force (*e.m.f.*) produced when the magnetic flux rapidly collapses at the point at which the current is removed from the coil.

### Thyristors

Semiconductor power control devices provide fast, efficient, and reliable methods of switching high currents and voltages. Thyristors (or ‘silicon controlled rectifiers’ as they are sometimes called) are three-terminal semiconductor devices which can switch very rapidly from a conducting to a non-conducting state. In the ‘off’ state, the thyristor exhibits negligible leakage current, whilst in the ‘on’ state the device exhibits very low resistance. This results in very little power loss within the thyristor even when appreciable power levels are being controlled.

Once switched into the conducting state, the thyristor will remain conducting (i.e. it is *latched* into the ‘on’ state) until the forward current is removed from the device. It is important to note that, in DC applications, this necessitates the interruption (or disconnection) of the supply before the device can be reset into its non-conducting state. However, where a thyristor is used with an alternating supply, the device will automatically become reset whenever the mains supply reverses. The device can then be triggered on the next half-cycle having correct polarity to permit conduction.

The symbol for a thyristor is shown in Figure 3.4.34 while Figure 3.4.35 shows the typical encapsulation used for such a device.

A typical power control arrangement based on a thyristor is shown in Figure 3.4.35. This circuit provides *half-wave control*. Assuming that the thyristor is perfect, the following formulae apply:

| Parameter              | Minimum | Maximum         |
|------------------------|---------|-----------------|
| Load voltage ( $V_L$ ) | 0 V     | $0.707 V_L$     |
| Load power ( $P_L$ )   | 0 W     | $0.5 V_S^2/R_L$ |
| Power curve frequency  | 50 Hz   |                 |

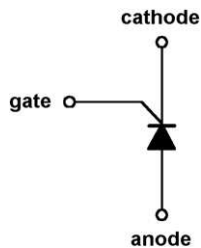


**Figure 3.4.34** Thyristor symbol

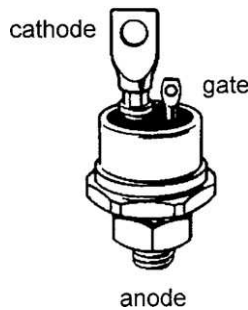


**Figure 3.4.35** A typical thyristor package

An improved power control arrangement based on two thyristors is shown in Figures 3.4.36 and 3.4.37. This circuit provides *full-wave*

**Figure 3.4.36** Half-wave power controller using a thyristor

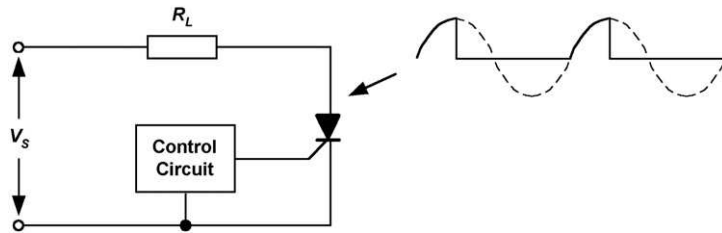
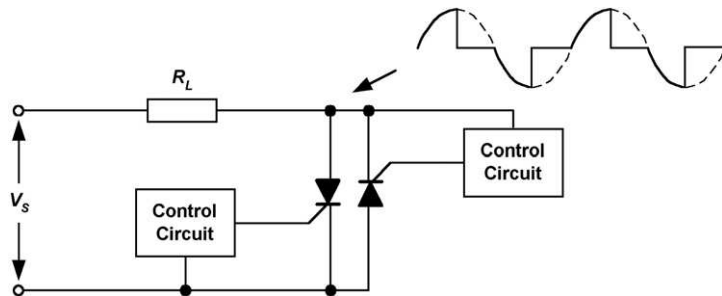


**Figure 3.4.37** Full-wave power controller using a thyristor



control. Assuming that each thyristor is perfect, the following formulae apply:

| Parameter              | Minimum | Maximum     |
|------------------------|---------|-------------|
| Load voltage ( $V_L$ ) | 0 V     | $V_s$       |
| Load power ( $P_L$ )   | 0 W     | $V_s^2/R_L$ |
| Power curve frequency  | 100 Hz  |             |

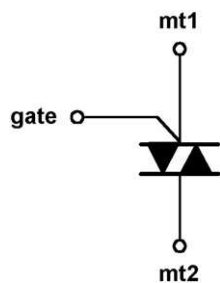


**Figure 3.4.38** Triac symbol

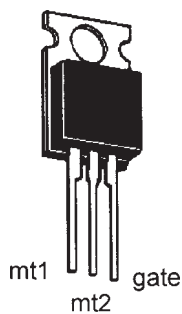


**Figure 3.4.39** A typical triac package

## Triacs

Triacs are a refinement of the thyristor which, when triggered, conduct on both positive and negative half-cycles of the applied voltage. Triacs have three terminals, known as main-terminal one (MT1), main-terminal two (MT2), and gate (G), as shown in Figure 3.4.38. Triacs can be triggered by both positive and negative voltages applied between G and MT1 with positive and negative voltages present at MT2, respectively. Triacs thus provide *full-wave control*.

In order to simplify the design of triggering circuits, triacs are often used in conjunction with diacs (equivalent to a bi-directional zener diode). A typical *diac* conducts heavily whenever the applied voltage exceeds approximately 30 V in either direction. Once triggered into the conducting state, the resistance of the diac falls to a very low value and thus a relatively large current will flow. The typical encapsulation used for a triac is shown in Figure 3.4.39.

A power control arrangement based on a triac is shown in Figure 3.4.40. This circuit provides *full-wave control*. Assuming that each triac is perfect, the following formulae apply:

| Parameter              | Minimum | Maximum     |
|------------------------|---------|-------------|
| Load voltage ( $V_L$ ) | 0 V     | $V_s$       |
| Load power ( $P_L$ )   | 0 W     | $V_s^2/R_L$ |
| Power curve frequency  | 100 Hz  |             |

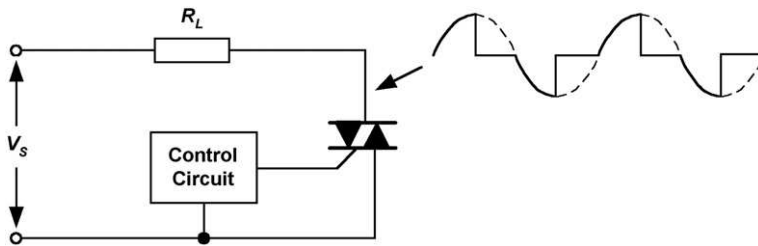


Figure 3.4.40 Power controller using a triac

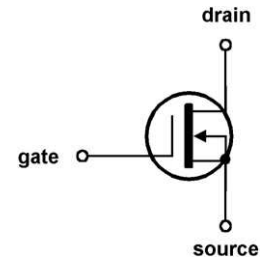


Figure 3.4.41 A typical power MOSFET

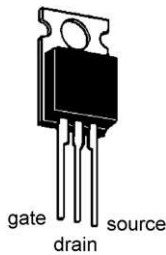


Figure 3.4.42 Power MOSFET symbol

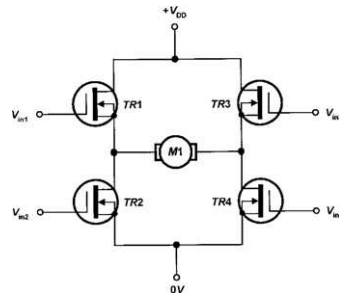


Figure 3.4.43 DC motor controller based on power MOSFETs

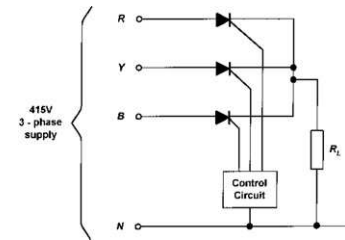


Figure 3.4.44 See Question 3.4.4

### Transistors

When DC, rather than AC is to be controlled, transistors (both bipolar and field-effect types) can be used. Both types offer the ability to control high currents and high voltages when supplied with only a low-level signal. However, unlike thyristors and triacs, transistors do not remain *latched* in the conducting state. The symbol and packaging for a typical power MOSFET transistor is shown in Figures 3.4.41 and 3.4.42, respectively. Figure 3.4.43 shows a typical DC motor controller based on a bridge arrangement of four power MOSFET devices.

#### Question 3.4.4

A three-phase power controller based on thyristors is shown in Figure 3.4.44. Sketch time related waveforms showing the gate control signals and load voltage assuming that each thyristor is triggered for one third of a cycle of the supply current.

# 4

# Electrical and electronic principles

## Summary

This unit covers the electrical principles that students in many branches of electrical and electronic engineering need to understand. It builds on some of the basic theory introduced in Chapter 3 and provides the basis for further study in the more specialised electrical units.

## 4.1 CIRCUIT THEORY

This section develops some of the basic electrical and electronic principles that we met in the previous chapter. It begins by explaining a number of useful circuit theorems including those developed by Thévenin and Norton. If you have previously studied Electrical and Electronic Principles at level N (or an equivalent GNVQ Advanced level unit) you should be quite comfortable with this material. If not, you should make sure that you *fully* understand Section 3.3 before going further!

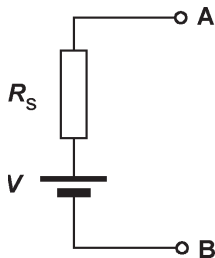


**Figure 4.1.1** Equivalent circuit of a voltage source

## Voltage and current sources

Let us start by considering the properties of a ‘perfect’ battery. No matter how much current is drawn from it, this device – if it were to exist – would produce a constant voltage between its positive and negative terminals. In practice, you will be well aware that the terminal voltage of a battery falls progressively as we draw more current from it. We account for this fact by referring to its *internal resistance*,  $R_s$  (see Figure 4.1.1).

### Internal resistance

The notion of internal resistance is fundamental to understanding the behaviour of electrical and electronic circuits and it is worth exploring this idea a little further before we take a look at circuit theorems in some detail. In Figure 4.1.1, the internal resistance is shown as a single discrete resistance connected in series with a voltage source. The electromotive force (e.m.f.) source has *exactly* the same voltage as that which would be measured between the battery’s terminals when no current is being drawn from it. As more current is drawn from a battery, its terminal voltage will fall



whereas the internal e.m.f. will remain constant. It might help to illustrate this with an example:

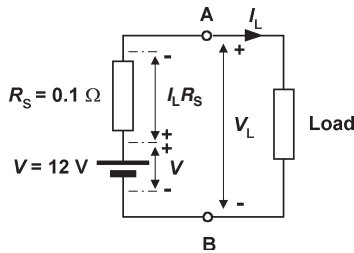


**Figure 4.1.2** See Example 4.1.1

### Example 4.1.1

A battery has an open-circuit (no-load) terminal voltage of 12 V and an internal resistance of 0.1  $\Omega$ . Determine the terminal voltage of the battery when it supplies currents of (a) 1 A, (b) 10 A, and (c) 100 A.

The terminal voltage of the battery when 'on-load' will always be less than the terminal voltage when there is no load connected. The reduction in voltage (i.e. the voltage that is effectively 'lost' inside the battery) will be equal to the load current multiplied by the internal resistance. Applying Kirchhoff's voltage law to the circuit of Figure 4.1.2 gives:

$$V_L = V - I_L R_S$$

where  $V_L$  is the terminal voltage on-load,  $V$  is the terminal voltage with no load connected,  $I_L$  is the load current, and  $R_S$  is the internal resistance of the battery.

In case (a):  $I_L = 1\text{ A}$ ,  $V = 12\text{ V}$ , and  $R_S = 0.1\ \Omega$  and thus:

$$V_L = V - I_L R_S = 12\text{ V} - (1\text{ A} \times 0.1\ \Omega) = 12 - 0.1 = 11.9\text{ V}.$$

In case (b):  $I_L = 10\text{ A}$ ,  $V = 12\text{ V}$ , and  $R_S = 0.1\ \Omega$  and thus:

$$V_L = V - I_L R_S = 12\text{ V} - (10\text{ A} \times 0.1\ \Omega) = 12 - 1 = 11\text{ V}.$$

In case (a):  $I_L = 100\text{ A}$ ,  $V = 12\text{ V}$ , and  $R_S = 0.1\ \Omega$  and thus:

$$V_L = V - I_L R_S = 12\text{ V} - (100\text{ A} \times 0.1\ \Omega) = 12 - 10 = 2\text{ V}.$$

The previous example shows that, provided that the internal resistance of a battery is very much smaller than the resistance of any circuit connected to the battery's terminals, a battery will provide a reasonably constant source of voltage. In fact, when a battery becomes exhausted, its internal resistance begins to rise sharply. This, in turn, reduces the terminal voltage when we try to draw current from the battery as the following example shows:

### Example 4.1.2

A battery has an open-circuit (no-load) terminal voltage of 9 V. If the battery is required to supply a load which has a resistance of 90  $\Omega$ , determine the terminal voltage of the battery when its internal resistance is (a) 1  $\Omega$  and (b) 10  $\Omega$ .

As we showed before, the terminal voltage of the battery when 'on-load' will always be less than the terminal voltage when there is no load connected. Applying Kirchhoff's voltage law gives:

$$V_L = V - I_L R_S$$

where  $V_L$  is the terminal voltage on-load,  $V$  is the terminal voltage with no load connected,  $I_L$  is the load current, and  $R_S$  is the internal resistance of the battery.

Now  $I_L = V_L/R_L$ .

Thus  $V_L = V - (V_L/R_L) \times R_S$

or  $V_L + (V_L/R_L) \times R_S = V$

or  $V_L(1 + (1/R_L) \times R_S) = V$

or  $V_L(1 + R_S/R_L) = V$ .

Hence  $V_L = V/(1 + R_S/R_L)$ .

In case (a):  $V = 12\text{ V}$ ,  $R_S = 1\ \Omega$ , and  $R_L = 90\ \Omega$ .

Thus  $V_L = V/(1 + R_S/R_L) = 12/(1 + 1/90) = 12/(1 + 0.011)$

or  $V_L = 12/1.011 = 11.87\text{ V}$ .

In case (b):  $V = 12\text{ V}$ ,  $R_S = 10\ \Omega$ , and  $R_L = 90\ \Omega$ .

Thus  $V_L = V/(1 + R_S/R_L) = 12/(1 + 10/90) = 12/(1 + 0.111)$

or  $V_L = 12/1.111 = 10.8\text{ V}$ .

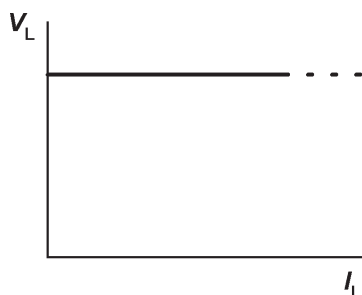


**Figure 4.1.3**  $V/I$  characteristic for a constant voltage source

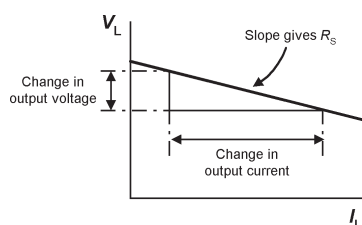


**Figure 4.1.4**  $V/I$  characteristic for a voltage source showing the effect of internal resistance



**Figure 4.1.5** Symbols used for DC and AC constant voltage sources

### The constant voltage source

A battery is the most obvious example of a voltage source. However, as we shall see later, any linear circuit with two terminals, no matter how complex, can be represented by an equivalent circuit based on a voltage source which uses just two components: a source of e.m.f.,  $V$ , connected in series with an internal resistance,  $R_S$ . Furthermore, the voltage source can either be a source of direct current (DC) (i.e. a battery), or a source of alternating current (AC) (i.e. an AC generator, a signal source, or the output of an oscillator).

An ideal voltage source – a *constant voltage source* – would have negligible internal resistance ( $R_S = 0$ ) and its voltage/current characteristic would look like that shown in Figure 4.1.3.

In practice,  $R_S \neq 0$  and the terminal voltage will fall as the load current,  $I_L$ , increases, as shown in Figure 4.1.4.

The internal resistance,  $R_S$ , can be found from the slope of the graph in Figure 4.1.4 as follows:

$$R_S = \frac{\text{Change in output voltage, } V_L}{\text{Corresponding change in output current, } I_L}.$$

The symbols commonly used for constant voltage DC and AC sources are shown in Figure 4.1.5.

### Example 4.1.3

A battery has a no-load terminal voltage of 12 V. If the terminal voltage falls to 10.65 V when a load current of 1.5 A is supplied, determine the internal resistance of the battery.

Now:

$$R_S = \frac{\text{Change in output voltage, } V_L}{\text{Corresponding change in output current, } I_L}$$

$$= \frac{12 - 10.65}{1.5 - 0} = \frac{1.35}{1.5} = 0.9 \, \Omega.$$

Note that we could have approached this problem in a different way by considering the voltage 'lost' inside the battery (i.e. the voltage dropped across  $R_S$ ). The lost voltage will be  $(12 - 10.65) = 1.35 \text{ V}$ . In this condition, the current flowing is  $1.5 \text{ A}$ , hence applying Ohm's law gives:

$$R_S = V/I = 1.35/1.5 = 0.9 \, \Omega.$$

### Thévenin's theorem

Thévenin's theorem states that:

*Any two-terminal network can be replaced by an equivalent circuit consisting of a voltage source and a series resistance equal to the internal resistance seen looking into the two terminals.*

In order to determine the Thévenin equivalent of a two-terminal network you need to:

- Determine the voltage that will appear between the two terminals with no load connected to the terminals. This is the *open-circuit voltage*.
- Replace any voltage sources with a short-circuit connection (or replace any current sources with an open-circuit connection) and then determine the *internal resistance* of the circuit (i.e. the resistance that appears between the two terminals).

#### Example 4.1.4

Determine the Thévenin equivalent of the two-terminal network shown in Figure 4.1.6.

First we need to find the voltage that will appear between the two terminals, A and B, with no load connected. In this condition, no current will be drawn from the network and thus there will be no voltage dropped across  $R_3$ . The voltage that

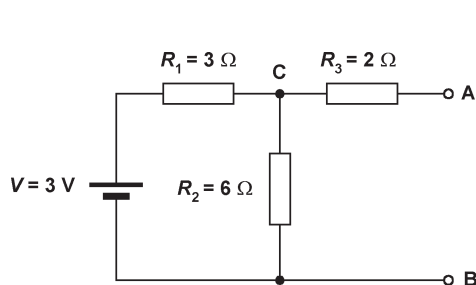


Figure 4.1.6 See Example 4.1.4

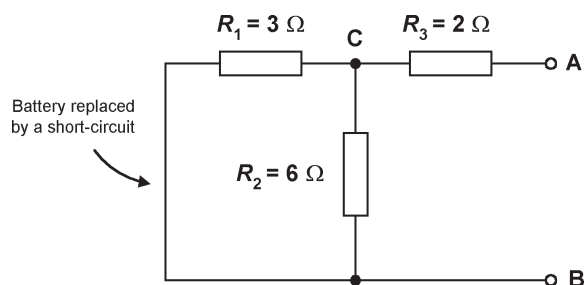


Figure 4.1.7 See Example 4.1.4

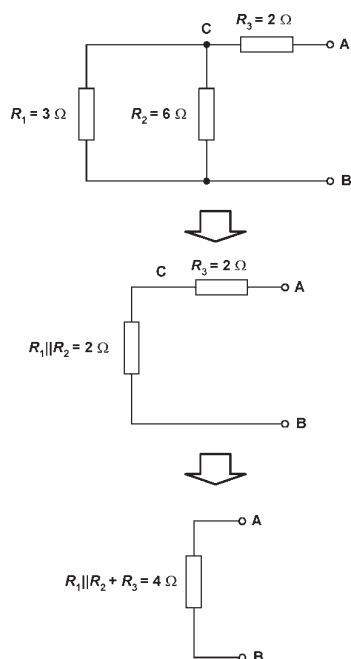


Figure 4.1.8 See Example 4.1.4

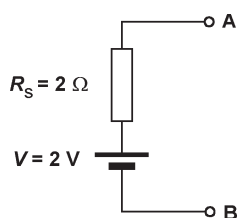


Figure 4.1.9 See Example 4.1.4

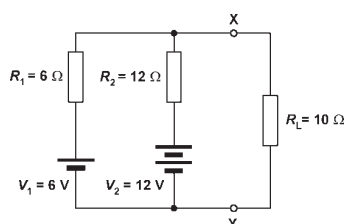


Figure 4.1.10 See Example 4.1.5

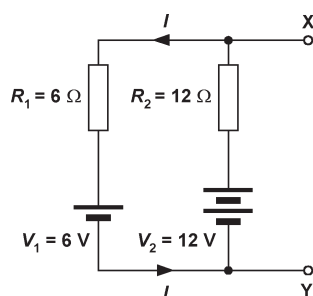


Figure 4.1.11 See Example 4.1.5

appears between A and B,  $V_{AB}$ , will then be identical to that which is dropped across  $R_2$  ( $V_{CB}$ ).

The voltage dropped across  $R_2$  (with no load connected) can be found by applying the potential divider theorem, thus:

$$V_{CB} = V \times R_2 / (R_1 + R_2) = 3 \times 6 / (3 + 6) = 18/9 = 2 \text{ V.}$$

Thus the open-circuit output voltage will be 2 V.

Next we need to find the internal resistance of the circuit with the battery replaced by a short-circuit (see Figure 4.1.7).

The circuit can be progressively reduced to a single resistor as shown in Figure 4.1.8.

Thus the internal resistance is 4 Ω.

The Thévenin equivalent of the two-terminal network is shown in Figure 4.1.9.

### Example 4.1.5

Determine the current flowing in, and voltage dropped across, the 10 Ω resistor shown in Figure 4.1.10.

First we will determine the Thévenin equivalent circuit of the network.

With no load connected to the network (i.e. with  $R_L$  disconnected) the current supplied by the 12 V battery will flow in an anti-clockwise direction around the circuit as shown in Figure 4.1.11.

The total e.m.f. present in the circuit will be:

$V = 12 - 6 = 6 \text{ V}$  (note that the two batteries oppose one another).

The total resistance in the circuit is:

$R = R_1 + R_2 = 6 + 12 = 18 \text{ Ω}$  (the two resistors are in series with one another).

Hence, the current flowing in the circuit when  $R_L$  is disconnected will be given by:

$$I = V/R = 6/18 = 0.333 \text{ A.}$$

The voltage developed between the output terminals (X and Y) will be equal to 12 V *minus* the voltage dropped across the 12 Ω resistor (or 6 V *plus* the voltage dropped across the 6 Ω resistor).

$$\text{Hence } V_{XY} = 12 - (0.333 \times 12) = 12 - 4 = 8 \text{ V}$$

$$\text{or } V_{XY} = 6 + (0.333 \times 6) = 6 + 2 = 8 \text{ V.}$$

Thus the open-circuit output voltage will be 8 V.

The internal resistance with  $V_1$  and  $V_2$  replaced by short-circuit will be given by:

$$R = 6 \times 12 / (6 + 12) = 72/18 = 4 \text{ Ω.}$$

Thus the internal resistance will be 4 Ω.

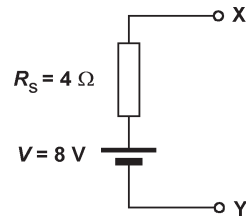


**Figure 4.1.12** See Example 4.1.5

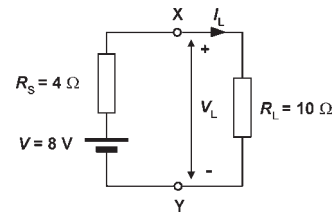


**Figure 4.1.13** See Example 4.1.5

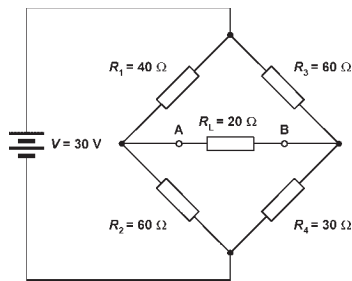


**Figure 4.1.14** See Example 4.1.6

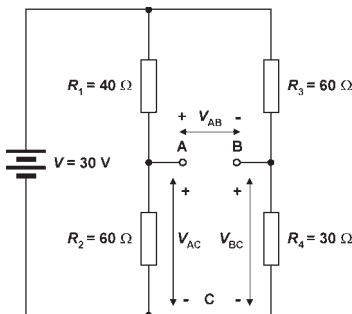


**Figure 4.1.15** See Example 4.1.6

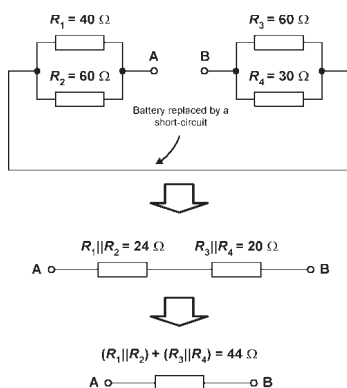


**Figure 4.1.16** See Example 4.1.6

Figure 4.1.12 shows the Thévenin equivalent of the circuit.

Now, with  $R_L = 10\ \Omega$  connected between X and Y the circuit looks like that shown in Figure 4.1.13.

The total resistance present will be:

$$R = R_S + R_L = 4 + 10 = 14\ \Omega.$$

The current drawn from the circuit in this condition will be given by:

$$I_L = V/R = 8/14 = 0.57\text{ A}.$$

Finally, the voltage dropped across  $R_L$  will be given by:

$$V_L = I \times R_L = 0.57 \times 10 = 5.7\text{ V}.$$

### Example 4.1.6

Determine the current flowing in a  $20\ \Omega$  resistor connected to the Wheatstone bridge shown in Figure 4.1.14.

Once again, we will begin by determining the Thévenin equivalent circuit of the network. To make things a little easier, we will redraw the circuit so that it looks a little more familiar (see Figure 4.1.15).

With A and B open-circuit, the circuit simply takes the form of two potential dividers ( $R_1$  and  $R_2$  forming one potential divider, whilst  $R_3$  and  $R_4$  form the other potential divider). It is thus a relatively simple matter to calculate the open-circuit voltage drop between A and B, since:

$$V_{AB} = V_{AC} - V_{BC} \text{ (from Kirchhoff's voltage law)}$$

$$\text{where } V_{AC} = V \times R_2 / (R_1 + R_2) = 30 \times 60 / (40 + 60) = 18\text{ V}$$

$$\text{and } V_{BC} = V \times R_4 / (R_3 + R_4) = 30 \times 30 / (30 + 60) = 10\text{ V}.$$

$$\text{Thus } V_{AB} = 18 - 10 = 8\text{ V}.$$

Next, to determine the internal resistance of the network with the battery replaced by a short-circuit, we can again redraw the circuit and progressively reduce it to one resistance. Figure 4.1.16 shows how this is done.

Thus the internal resistance is  $44\ \Omega$ .

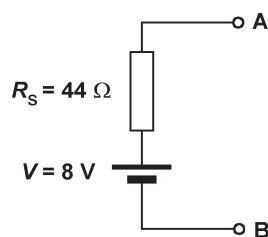


**Figure 4.1.17** See Example 4.1.6

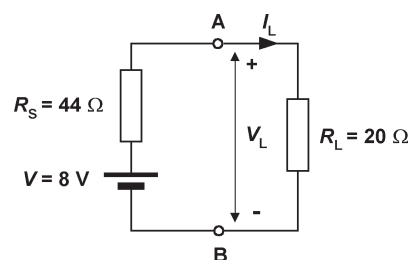


**Figure 4.1.18** See Example 4.1.6

The Thévenin equivalent of the bridge network is shown in Figure 4.1.17.

Now, with  $R_L = 20\ \Omega$  connected between A and B the circuit looks like that shown in Figure 4.1.18.

The total resistance present will be:

$$R = R_S + R_L = 44 + 20 = 64\ \Omega.$$

The current drawn from the circuit in this condition will be given by:

$$I_L = V/R = 8/64 = 0.125\ \text{A}.$$

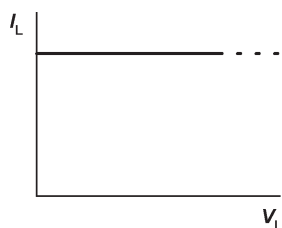


**Figure 4.1.19**  $I/V$  characteristic for a constant current source

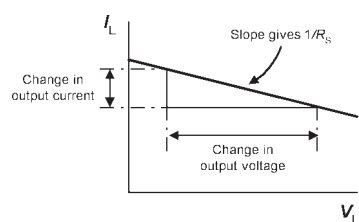


**Figure 4.1.20**  $I/V$  characteristic for a constant current source showing the effect of internal resistance

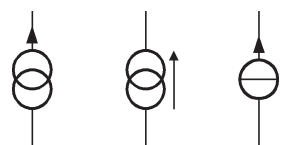


**Figure 4.1.21** Symbols used for DC and AC constant current sources

### The constant current source

By now, you should be familiar with the techniques for analysing simple electric circuits by considering the voltage sources that may be present and their resulting effect on the circuit in terms of the currents and voltage drops that they produce. However, when analysing some types of circuit it can often be more convenient to consider sources of current rather than voltage. In fact, any linear circuit with two terminals, no matter how complex, can be thought of as a current source; that is, a supply of current,  $I$ , connected in parallel with an internal resistance,  $R_p$ . As with voltage sources, current sources can either be a source of DC or AC.

An ideal current source – a *constant current source* – would have infinite internal resistance ( $R_p = \infty$ ) and its voltage/current characteristic would look like that shown in Figure 4.1.19.

In practice,  $R_p \neq \infty$  and the current will fall as the load voltage,  $V_L$ , increases, as shown in Figure 4.1.20.

The internal resistance,  $R_p$ , can be found from the reciprocal of the slope of the graph in Figure 4.1.20 as follows:

$$\begin{aligned} R_p &= \frac{1}{\left( \frac{\text{Change in output current, } I_L}{\text{Corresponding change in output voltage, } V_L} \right)} \\ &= \frac{\text{Change in output voltage, } V_L}{\text{Corresponding change in output current, } I_L} \end{aligned}$$

The symbols commonly used for constant current DC and AC sources are shown in Figure 4.1.21.

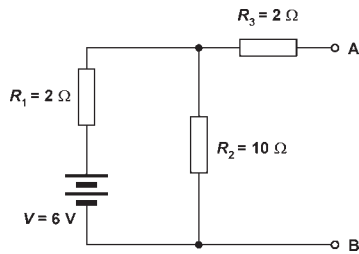


Figure 4.1.22 See Example 4.1.7

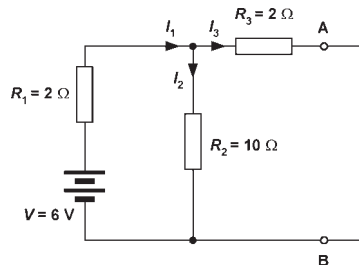


Figure 4.1.23 See Example 4.1.7

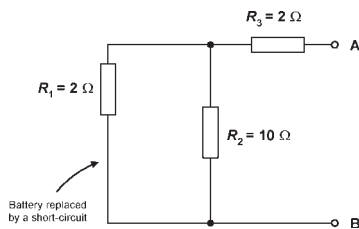


Figure 4.1.24 See Example 4.1.7

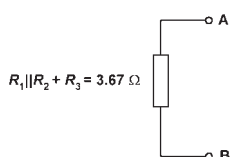
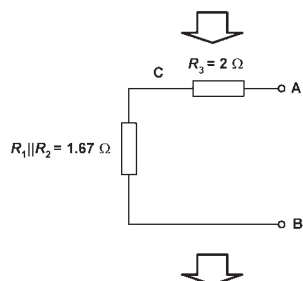
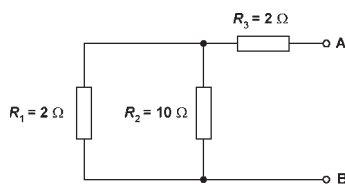


Figure 4.1.25 See Example 4.1.7

## Norton's theorem

Norton's theorem states that:

*Any two-terminal network can be replaced by an equivalent circuit consisting of a current source and a parallel resistance equal to the internal resistance seen looking into the two terminals.*

In order to determine the Norton equivalent of a two-terminal network you need to:

- Determine the current that will appear between the two terminals when they are short-circuited together. This is the *short-circuit current*.
- Replace any current sources with an open-circuit connection (or replace any voltage sources with a short-circuit connection) and then determine the *internal resistance* of the circuit (i.e. the resistance that appears between the two terminals).

### Example 4.1.7

Determine the Norton equivalent of the two-terminal network shown in Figure 4.1.22.

First, we need to find the current that would flow between terminals A and B with the two terminals linked together by a short-circuit (see Figure 4.1.23).

The total resistance that appears across the 6V battery with the output terminals linked together is given by:

$$R_1 + (R_2 \times R_3)/(R_2 + R_3) = 2 + (2 \times 10)/(2 + 10) \\ = 2 + 20/12 = 3.67 \Omega.$$

The current supplied by the battery,  $I_1$ , will then be given by:

$$I_1 = V/R = 6/3.67 = 1.63 \text{ A.}$$

Using the current divider theorem, the current in  $R_3$  (and that flowing between terminals A and B) will be given by:

$$I_3 = I_1 \times R_2/(R_2 + R_3) = 1.63 \times 10/(10 + 2) \\ = 16.3/12 = 1.35 \text{ A.}$$

Thus the short-circuit current will be 1.35 A.

Second, we need to find the internal resistance of the circuit with the battery replaced by a short-circuit (see Figure 4.1.24).

The circuit can be progressively reduced to a single resistor as shown in Figure 4.1.25.

Thus the internal resistance is 3.67 Ω.

The Norton equivalent of the two-terminal network is shown in Figure 4.1.26.

### Example 4.1.8

Determine the voltage dropped across a 6Ω resistor connected between X and Y in the circuit shown in Figure 4.1.27.

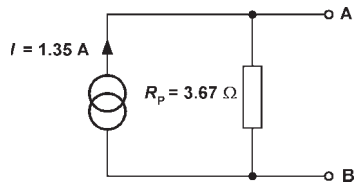


**Figure 4.1.26** See Example 4.1.7

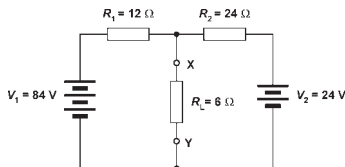


**Figure 4.1.27** See Example 4.1.8

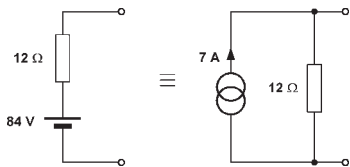


**Figure 4.1.28** See Example 4.1.8

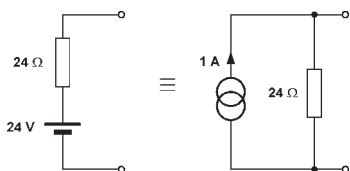


**Figure 4.1.29** See Example 4.1.8

First, we will determine the Norton equivalent circuit of the network. Since there are two voltage sources ( $V_1$  and  $V_2$ ) in this circuit, we can derive the Norton equivalent of each branch and then combine them together into one current source and one parallel resistance. This makes life a lot easier!

The left-hand branch of the circuit is equivalent to a current source of 7 A with an internal resistance of 12 Ω (see Figure 4.1.28).

The right-hand branch of the circuit is equivalent to a current source of 1 A with an internal resistance of 24 Ω (see Figure 4.1.29).

The combined effect of the two current sources is a single constant current generator of  $(7 + 1) = 8$  A with a parallel resistance of  $12 \times 24 / (12 + 24) = 8$  Ω (see Figure 4.1.30).

Now, with  $R_L = 6$  Ω connected between X and Y the circuit looks like that shown in Figure 4.1.31.

The total resistance present will be:

$$R = R_p + R_L / (R_p + R_L) = 8 \times 6 / (8 + 6) = 3.43 \Omega.$$

When 8 A flows in this resistance, the voltage dropped across the parallel combination of resistors will be:

$$V_L = 8 / 3.43 = 2.33 \text{ V}.$$

### Thévenin to Norton conversion

The previous example shows that it is relatively easy to convert from one equivalent circuit to the other (do not overlook the fact that, whichever equivalent circuit is used to solve a problem, the result will be identical!).

To convert from the Thévenin equivalent circuit to the Norton equivalent circuit:

$$R_p = R_s \quad \text{and} \quad V_{TH} = I_N \times R_p.$$

To convert from Norton equivalent circuit to the Thévenin equivalent circuit:

$$R_s = R_p \quad \text{and} \quad I_N = V_{TH} / R_s.$$

This equivalence is shown in Figure 4.1.32.

### Superposition theorem

The superposition theorem is a simple yet powerful method that can be used to analyse networks containing a number of voltage sources and linear resistances.

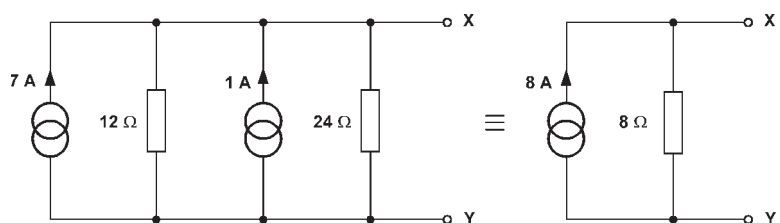


**Figure 4.1.30** See Example 4.1.8



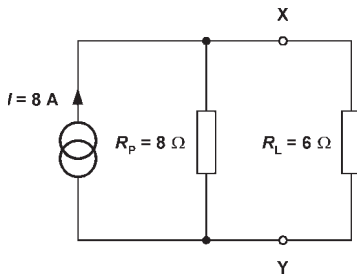


**Figure 4.1.31** See Example 4.1.8

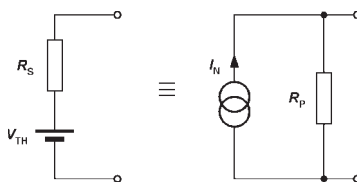


**Figure 4.1.32** Norton and Thévenin equivalent circuits

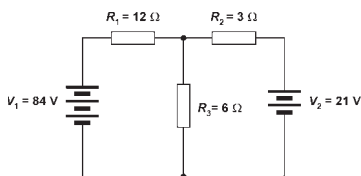


**Figure 4.1.33** See Example 4.1.9

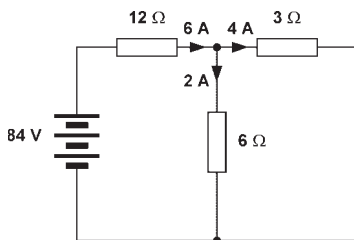


**Figure 4.1.34** See Example 4.1.9

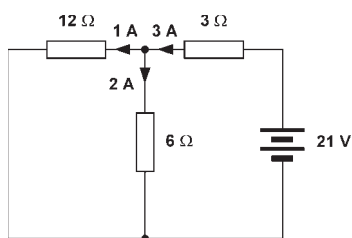


**Figure 4.1.35** See Example 4.1.9

The superposition theorem states that:

*In any network containing more than one voltage source, the current in, or potential difference developed across, any branch can be found by considering the effects of each source separately and adding their effects. During this process, any temporarily omitted source must be replaced by its internal resistance (or a short-circuit if it is a perfect voltage source).*

This might sound a little more complicated than it really is so we shall explain the use of the theorem with a simple example:

### Example 4.1.9

Use the superposition theorem to determine the voltage dropped across the  $6\ \Omega$  resistor in Figure 4.1.33.

First we shall consider the effects of the left-hand voltage source ( $V_1$ ) when taken on its own. Figure 4.1.34 shows the voltages and currents (you can easily check these by calculation, if you wish!).

Next we shall consider the effects of the right-hand voltage source ( $V_2$ ) when taken on its own. Figure 4.1.35 shows the voltages and currents (again, you can easily check these by calculation!).

Finally, we can combine these two sets of results, taking into account the direction of current flow to arrive at the values of current in the complete circuit. Figure 4.1.36 shows how this is done. The current in the  $6\ \Omega$  resistor is thus  $4\ \text{A}$  and the voltage dropped across it will be:

$$V = I \times R = 4 \times 6 = 24\ \text{V}.$$

## T-, $\pi$ -, star-, and delta-networks

Figure 4.1.37 shows four basic electrical network configurations that you will encounter in various forms. In order to solve network problems it is often very helpful to convert from one type of circuit to another that has equivalent properties. We shall consider the T and  $\pi$  circuits first.

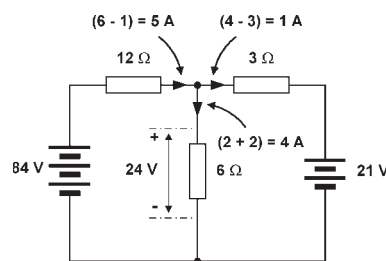


**Figure 4.1.36** See Example 4.1.9

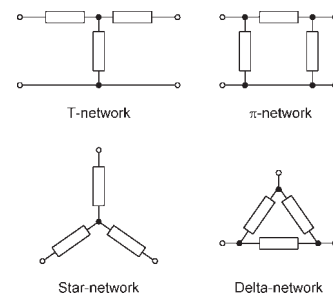


**Figure 4.1.37** Four basic electrical network configurations

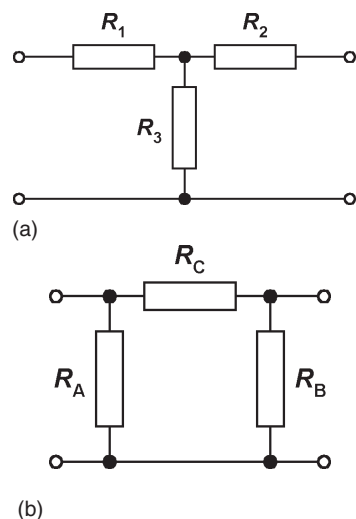


**Figure 4.1.38** T- and \$\pi\$-networks

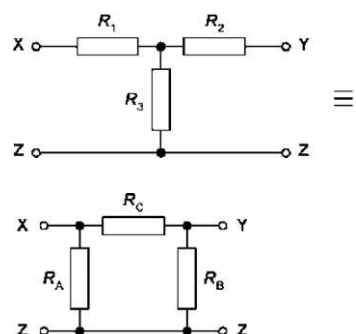


**Figure 4.1.39** T-network equivalent of a \$\pi\$-network

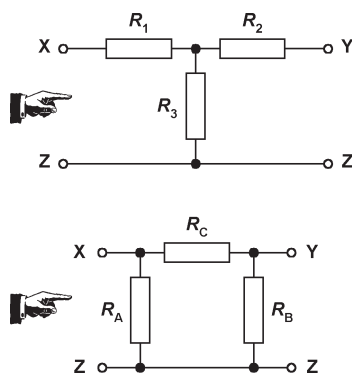


**Figure 4.1.40** Comparing the resistance seen looking into the networks at X-Z

### T-\$\pi\$- and \$\pi\$-T-transformation

T- and \$\pi\$-networks are commonly encountered in attenuators and filters. Figure 4.1.38(a) shows a T-network (comprising \$R\_1\$, \$R\_2\$, and \$R\_3\$) whilst Figure 4.1.38(b) shows a \$\pi\$-network (comprising \$R\_A\$, \$R\_B\$, and \$R\_C\$).

The T-equivalent of the \$\pi\$-network can be derived as follows. The resistance seen looking into each network at X-Z will be (Figures 4.1.39–4.1.42):

$$R_1 + R_3 = \frac{R_A(R_B + R_C)}{R_A + R_B + R_C} \quad (i)$$

The resistance seen looking into each network at Y-Z will be:

$$R_2 + R_3 = \frac{R_B(R_A + R_C)}{R_A + R_B + R_C} \quad (ii)$$

The resistance seen looking into each network at X-Y will be:

$$R_1 + R_2 = \frac{R_C(R_A + R_B)}{R_A + R_B + R_C} \quad (iii)$$

Subtracting (ii) from (i) gives:

$$R_1 - R_2 = \frac{R_A(R_B + R_C) - R_B(R_A + R_C)}{R_A + R_B + R_C} \quad (iv)$$

Adding (iii) and (iv) gives:

$$2R_1 = \frac{R_A(R_B + R_C) - R_B(R_A + R_C) + R_C(R_A + R_B)}{R_A + R_B + R_C}$$

$$\begin{aligned} \text{thus } 2R_1 &= \frac{R_A R_B + R_A R_C - R_A R_B - R_B R_C + R_A R_C + R_B R_C}{R_A + R_B + R_C} \\ &= \frac{2R_A R_C}{R_A + R_B + R_C} \end{aligned}$$

$$\text{thus } R_1 = \frac{R_A R_C}{R_A + R_B + R_C} \quad (v)$$

Similarly:

$$R_2 = \frac{R_B R_C}{R_A + R_B + R_C} \quad (vi)$$

and

$$R_3 = \frac{R_A R_B}{R_A + R_B + R_C} \quad (vii)$$

The \$\pi\$-equivalent of the T-network is derived as follows: Dividing (vi) by (v) gives:

$$\frac{R_2}{R_1} = \frac{R_B R_C}{R_A R_C} = \frac{R_B}{R_A}$$

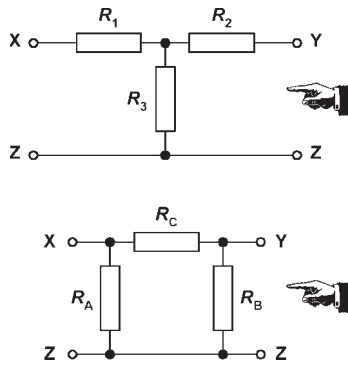


**Figure 4.1.41** Comparing the resistance seen looking into the networks at Y-Z

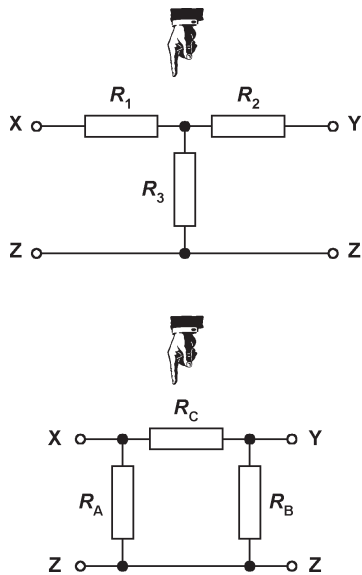


**Figure 4.1.42** Comparing the resistance seen looking into the networks at X-Y

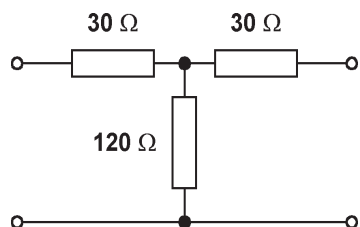


**Figure 4.1.43** See Example 4.1.10

$$\text{therefore, } R_B = \frac{R_2}{R_1} \times R_A. \quad (\text{viii})$$

Similarly, dividing (vi) by (vii) gives:

$$\frac{R_2}{R_3} = \frac{R_B R_C}{R_A R_B} = \frac{R_C}{R_A}.$$

$$\text{therefore, } R_C = \frac{R_2}{R_3} \times R_A. \quad (\text{ix})$$

Substituting (viii) and (ix) for  $R_B$  and  $R_C$ , respectively, in (v) gives:

$$\begin{aligned} R_1 &= \frac{R_A \left( \frac{R_2}{R_3} \times R_A \right)}{R_A + \left( \frac{R_2}{R_1} \times R_A \right) + \left( \frac{R_2}{R_3} \times R_A \right)} \\ &= \frac{R_A^2 \times \frac{R_2}{R_3}}{R_A \left( 1 + \frac{R_2}{R_1} + \frac{R_2}{R_3} \right)} = \frac{R_A \times \frac{R_2}{R_3}}{1 + \frac{R_2}{R_1} + \frac{R_2}{R_3}} \end{aligned}$$

$$\begin{aligned} \text{thus } R_A &= \frac{R_1 R_3}{R_2} \left( 1 + \frac{R_2}{R_1} + \frac{R_2}{R_3} \right) = \frac{R_1 R_3}{R_2} + \frac{R_1 R_2 R_3}{R_1 R_2} + \frac{R_1 R_2 R_3}{R_2 R_3} \\ &= \frac{R_1 R_3}{R_2} + \frac{R_2 R_3}{R_2} + \frac{R_1 R_2}{R_2} \end{aligned}$$

$$\text{hence } R_A = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_2}. \quad (\text{x})$$

Similarly:

$$R_B = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_1} \quad (\text{xi})$$

and

$$R_C = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_3}. \quad (\text{xii})$$

The six transformation equations can be summarised as follows:

*The equivalent  $\pi$ -network resistance between two terminals of a T-network is equal to the sum of the products of each pair of adjacent resistances divided by the opposite resistance – see Equations (x)–(xii).*

*The equivalent T-network resistance between two adjacent terminals of a  $\pi$ -network is equal to the product of the two adjacent resistances divided by the sum of the three resistances – see Equations (v)–(vii).*

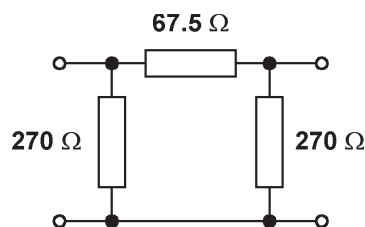


**Figure 4.1.44**  $\pi$ -network equivalent of the T-network shown in Figure 4.1.43

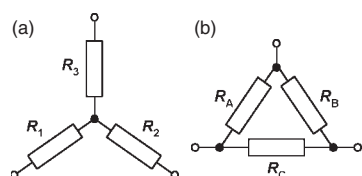


**Figure 4.1.45** Star- and delta-networks

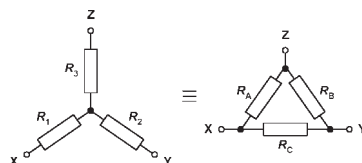


**Figure 4.1.46** Star-equivalent of a delta-network

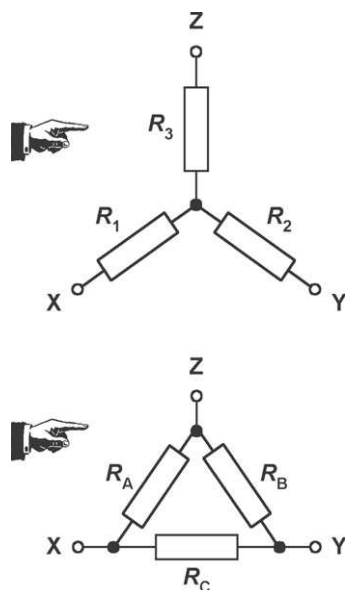


**Figure 4.1.47** Comparing the resistance seen looking into the networks at X-Z

### Example 4.1.10

Determine the  $\pi$ -network equivalent of the T-network shown in Figure 4.1.43.

From Figure 4.1.43,  $R_1 = R_2 = 30\ \Omega$ , and  $R_3 = 120\ \Omega$ .

$$\text{Now } R_A = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_2}.$$

$$\text{Now } R_1 R_2 + R_1 R_3 + R_2 R_3 = (30 \times 30) + (30 \times 120) + (30 \times 120) = 8100$$

$$\text{and } R_A = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_2} = \frac{8100}{30} = 270\ \Omega$$

$$\text{similarly } R_B = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_1} = \frac{8100}{30} = 270\ \Omega$$

$$\text{and } R_C = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_3} = \frac{8100}{120} = 67.5\ \Omega.$$

Figure 4.1.44 shows the  $\pi$ -equivalent of the T-network shown in Figure 4.1.43.

### Star-delta- and delta-star-transformation

Star- and delta-networks are commonly encountered in three-phase systems. Figure 4.1.45(a) shows a star-network (comprising  $R_1$ ,  $R_2$ , and  $R_3$ ) whilst Figure 4.1.45(b) shows a delta-network (comprising  $R_A$ ,  $R_B$ , and  $R_C$ ).

The star-equivalent of the delta-network is derived as follows. The resistance seen looking into each network at X-Z will be (Figures 4.1.46–4.1.49):

$$R_1 + R_3 = \frac{R_A(R_B + R_C)}{R_A + R_B + R_C}. \quad (\text{i})$$

The resistance seen looking into each network at Y-Z will be:

$$R_2 + R_3 = \frac{R_B(R_A + R_C)}{R_A + R_B + R_C}. \quad (\text{ii})$$

The resistance seen looking into each network at X-Y will be:

$$R_1 + R_2 = \frac{R_C(R_A + R_B)}{R_A + R_B + R_C}. \quad (\text{iii})$$

Subtracting (ii) from (i) gives:

$$R_1 - R_2 = \frac{R_A(R_B + R_C) - R_B(R_A + R_C)}{R_A + R_B + R_C}. \quad (\text{iv})$$

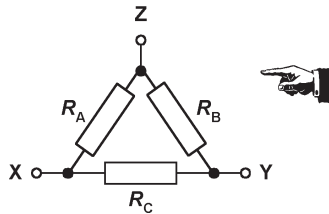
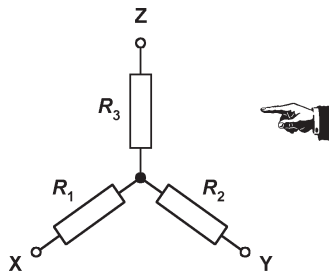


**Figure 4.1.48** Comparing the resistance seen looking into the networks at Y–Z

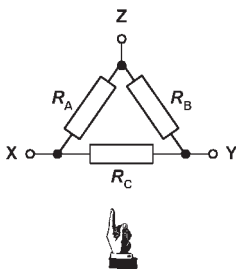
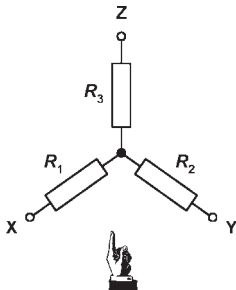


**Figure 4.1.49** Comparing the resistance seen looking into the networks at X–Y

Adding (iii) and (iv) gives:

$$2R_1 = \frac{R_A(R_B + R_C) - R_B(R_A + R_C) + R_C(R_A + R_B)}{R_A + R_B + R_C}.$$

$$\begin{aligned} \text{thus } 2R_1 &= \frac{R_A R_B + R_A R_C - R_A R_B - R_B R_C + R_A R_C + R_B R_C}{R_A + R_B + R_C} \\ &= \frac{2R_A R_C}{R_A + R_B + R_C} \end{aligned}$$

$$\text{Thus } R_1 = \frac{R_A R_C}{R_A + R_B + R_C}. \quad (\text{v})$$

Similarly:

$$R_2 = \frac{R_B R_C}{R_A + R_B + R_C} \quad (\text{vi})$$

and

$$R_3 = \frac{R_A R_B}{R_A + R_B + R_C}. \quad (\text{vii})$$

The delta-equivalent of the star-network is derived as follows: Dividing (vi) by (v) gives:

$$\frac{R_2}{R_1} = \frac{R_B R_C}{R_A R_C} = \frac{R_B}{R_A}$$

$$\text{therefore, } R_B = \frac{R_2}{R_1} \times R_A. \quad (\text{viii})$$

Similarly, dividing (vi) by (vii) gives:

$$\frac{R_2}{R_3} = \frac{R_B R_C}{R_A R_B} = \frac{R_C}{R_A}$$

$$\text{therefore, } R_C = \frac{R_2}{R_3} \times R_A. \quad (\text{ix})$$

Substituting (viii) and (ix) for  $R_B$  and  $R_C$ , respectively, in (v) gives:

$$\begin{aligned} R_1 &= \frac{R_A \left( \frac{R_2}{R_3} \times R_A \right)}{R_A + \left( \frac{R_2}{R_1} \times R_A \right) + \left( \frac{R_2}{R_3} \times R_A \right)} \\ &= \frac{R_A^2 \times \frac{R_2}{R_3}}{R_A \left( 1 + \frac{R_2}{R_1} + \frac{R_2}{R_3} \right)} = \frac{R_A \times \frac{R_2}{R_3}}{1 + \frac{R_2}{R_1} + \frac{R_2}{R_3}}. \end{aligned}$$

$$\begin{aligned}\text{Thus } R_A &= \frac{R_1 R_3}{R_2} \left( 1 + \frac{R_2}{R_1} + \frac{R_2}{R_3} \right) = \frac{R_1 R_3}{R_2} + \frac{R_1 R_2 R_3}{R_1 R_2} + \frac{R_1 R_2 R_3}{R_2 R_3} \\ &= \frac{R_1 R_3}{R_2} + \frac{R_2 R_3}{R_2} + \frac{R_1 R_2}{R_2}\end{aligned}$$

$$\text{hence } R_A = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_2} \quad (\text{x})$$

Similarly:

$$R_B = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_1} \quad (\text{xi})$$

and

$$R_C = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_3} \quad (\text{xii})$$

The six transformation equations can be summarised as follows:

*The equivalent delta resistance between two terminals of a star-network is equal to the sum of the products of each pair of adjacent resistances divided by the opposite resistance – see Equations (x)–(xii).*

*The equivalent star resistance between two adjacent terminals of a delta-network is equal to the product of the two adjacent resistances divided by the sum of the three resistances – see Equations (v)–(vii).*

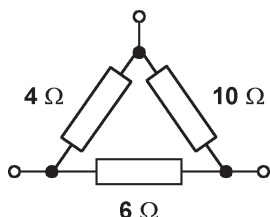


**Figure 4.1.50** See Example 4.1.11

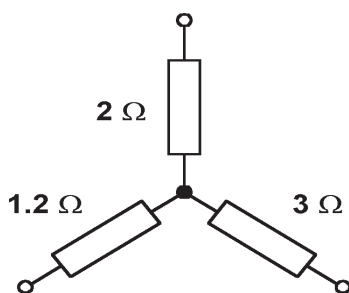


**Figure 4.1.51** Star-network equivalent of the delta-network shown in Figure 4.1.50

### Example 4.1.11

Determine the star-network equivalent of the delta-network shown in Figure 4.1.50.

From Figure 4.1.50,  $R_A = 4 \Omega$ ,  $R_B = 10 \Omega$ , and  $R_C = 6 \Omega$ .

$$\text{Now } R_1 = \frac{R_A R_C}{R_A + R_B + R_C}$$

$$\text{Now } R_A + R_B + R_C = 4 + 10 + 6 = 20 \Omega$$

$$\text{and } R_1 = \frac{R_A R_C}{R_A + R_B + R_C} = \frac{4 \times 6}{20} = \frac{24}{20} = 1.2 \Omega$$

$$\text{similarly } R_2 = \frac{R_B R_C}{R_A + R_B + R_C} = \frac{10 \times 6}{20} = \frac{60}{20} = 3 \Omega$$

$$\text{and } R_3 = \frac{R_A R_B}{R_A + R_B + R_C} = \frac{4 \times 10}{20} = \frac{40}{20} = 2 \Omega.$$

Figure 4.1.51 shows the star-equivalent of the delta-network shown in Figure 4.1.50.

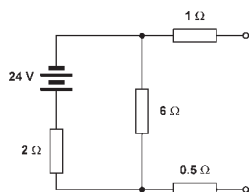


**Figure 4.1.52** See Questions 4.1.1

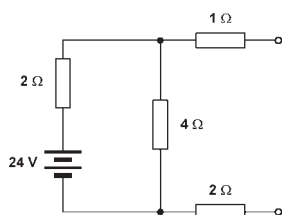


**Figure 4.1.53** See Questions 4.1.1

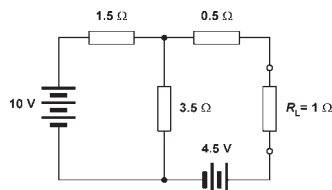


**Figure 4.1.54** See Questions 4.1.1

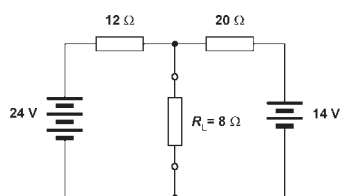


**Figure 4.1.55** See Questions 4.1.1

## Questions 4.1.1

- (1) The terminal voltage of a battery falls from 24 V with no load connected to 21.6 V when supplying a current of 6 A. Determine the internal resistance of the battery. Also determine the terminal voltage of the battery when it supplies a current of 10 A.
- (2) Determine the Thévenin equivalent of the circuit shown in Figure 4.1.52.
- (3) Determine the Norton equivalent of the circuit shown in Figure 4.1.53.
- (4) Use Thévenin's theorem to determine the current flowing in  $R_L$  in the circuit shown in Figure 4.1.54.
- (5) Use Norton's theorem to determine the voltage developed across  $R_L$  in the circuit shown in Figure 4.1.55.
- (6) Use the Superposition theorem to determine the voltage dropped across the 15 Ω resistor in Figure 4.1.56.
- (7) Determine the  $\pi$ -network equivalent of the T-network shown in Figure 4.1.57.
- (8) Determine the T-network equivalent of the  $\pi$ -network shown in Figure 4.1.58.
- (9) Determine the star-network equivalent of the delta-network shown in Figure 4.1.59.
- (10) Determine the delta-network equivalent of the star-network shown in Figure 4.1.60.

## Networks of complex impedances

Until now, we have only considered networks of pure resistance and DC sources in our explanation of the basic circuit theorems. However, many real circuits use AC sources and contain a mixture of resistance and reactance

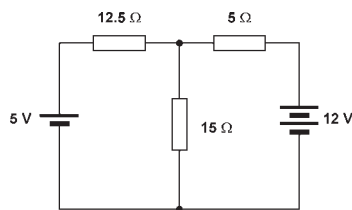


**Figure 4.1.56** See Questions 4.1.1

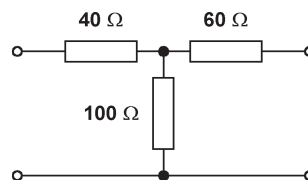


**Figure 4.1.57** See Questions 4.1.1

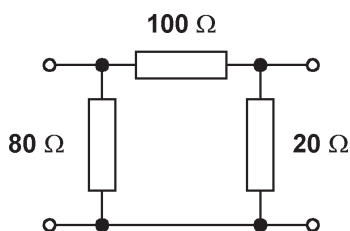


**Figure 4.1.58** See Questions 4.1.1

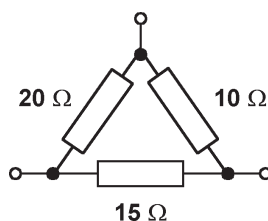


**Figure 4.1.59** See Questions 4.1.1

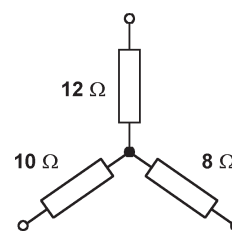


**Figure 4.1.60** See Questions 4.1.1

in which case it is necessary to consider the effects of *impedance* rather than resistance. Before we begin, we shall introduce another technique for representing complex impedances which will save us a great deal of time and effort later on!

### Another way of representing complex impedances

In Chapter 3 we showed how a complex impedance could be represented using the  $j$ -operator. In particular you should recall that any complex impedance can be represented by the relationship:

$$Z = (R \pm jX),$$

where  $Z$  represents impedance,  $R$  represents resistance, and  $X$  represents reactance.

You should also recall that a positive  $j$ -term ( $+jX$ ) indicates an inductive reactance whilst a negative  $j$ -term ( $-jX$ ) indicates a capacitive reactance. Since we use a set of rectangular axes (an Argand diagram) to represent  $Z = (R \pm jX)$  graphically, this form of notation is sometimes referred to as rectangular notation.

An alternative way of representing a complex quantity is that of using a notation based on a polar, rather than a rectangular notation. Two quantities are required to specify an impedance using polar notation;  $Z$  and  $\theta$  (where  $Z$  is the *modulus*, and  $\theta$  is the angle, or *argument*).

Figure 4.1.61 shows the relationship between the rectangular and polar methods of representing a complex impedance.

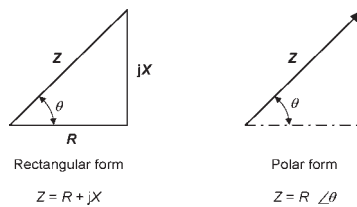


**Figure 4.1.61** Relationship between rectangular and polar methods of representing complex impedances

From Figure 4.1.61,  $Z = \sqrt{R^2 + X^2}$  and  $\theta = \arctan(X/R)$ .

Notice also that  $R = Z \cos \theta$  and  $X = Z \sin \theta$ .

Using the foregoing relationships it is relatively simple to convert from rectangular form to polar form, and vice versa. Here are a few examples of some identical impedances:

| Rectangular form  | Polar form               |
|-------------------|--------------------------|
| 10 or $(10 + j0)$ | $10 \angle 0^\circ$      |
| $+j10$            | $10 \angle 90^\circ$     |
| $-j10$            | $10 \angle -90^\circ$    |
| $10 + j10$        | $14.14 \angle 45^\circ$  |
| $10 - j10$        | $14.14 \angle -45^\circ$ |

#### Example 4.1.12

A circuit comprises a  $30 \Omega$  resistor connected in series with an inductor having an inductive reactance of  $40 \Omega$ . Determine the impedance of the circuit and express your answer in both rectangular and polar form.

Now  $Z = R \pm jX$

where  $R = 30 \Omega$  and  $X = +j40 \Omega$ .

Thus  $Z = (30 + j40) \Omega$  (rectangular form).



Since  $Z = \sqrt{R^2 + X^2}$  (we sometimes refer to this as *modulus-Z* or  $|Z|$ )

$$Z = \sqrt{R^2 + X^2} = \sqrt{30^2 + 40^2} = \sqrt{2500} = 50 \Omega$$

and  $\theta = \arctan(X/R) = \arctan(40/30) = 53.13^\circ$ .

Thus  $Z = 50 \angle 53.13^\circ \Omega$  (polar form)

(note that  $30 + j40 \equiv 50 \angle 53.13^\circ$ ).

Multiplying and dividing impedances expressed in the rectangular ( $R \pm jX$ ) form can be quite tedious and error prone. To find the product of two impedances we have to multiply out the brackets, rearrange the terms and replace any  $j^2$  by  $-1$ . Having done that we must 'tidy up' the result by grouping the terms together and writing the result in the  $R \pm jX$  form. This is how it is done:

$$\text{Let } Z_1 = R_1 + jX_1 \text{ and } Z_2 = R_2 + jX_2.$$

$$\begin{aligned} \text{Then } Z_1 \times Z_2 &= (R_1 + jX_1)(R_2 + jX_2) \\ &= R_1R_2 + R_1jX_2 + R_2jX_1 + j^2X_1X_2 \end{aligned}$$

$$\begin{aligned} \text{or } Z_1 \times Z_2 &= R_1R_2 + j(R_1X_2 + R_2X_1) + (-1)X_1X_2 \\ &= R_1R_2 + j(R_1X_2 + R_2X_1) - X_1X_2 \end{aligned}$$

$$\text{hence } Z_1 \times Z_2 = (R_1R_2 - X_1X_2) + j(R_1X_2 + R_2X_1).$$

The division of impedances expressed in complex form is even more complicated as it requires us to multiply the top and bottom of the expression by the *complex conjugate* of the denominator. This is how it is done:

$$\text{Once again, let } Z_1 = R_1 + jX_1 \text{ and } Z_2 = R_2 + jX_2.$$

$$\begin{aligned} \text{Then } \frac{Z_1}{Z_2} &= \frac{R_1 + jX_1}{R_2 + jX_2} = \frac{(R_1 + jX_1)(R_2 - jX_2)}{(R_2 + jX_2)(R_2 - jX_2)} \\ &= \frac{(R_1R_2 + X_1X_2) + j(R_2X_1 - R_1X_2)}{R_2^2 + R_2jX_2 - R_2jX_2 - j^2X_2^2} \end{aligned}$$

$$\begin{aligned} \text{or } \frac{Z_1}{Z_2} &= \frac{(R_1R_2 + X_1X_2) + j(R_2X_1 - R_1X_2)}{R_2^2 - j^2X_2^2} \\ &= \frac{(R_1R_2 + X_1X_2) + j(R_2X_1 - R_1X_2)}{R_2^2 - (-1)X_2^2} \end{aligned}$$

$$\begin{aligned} \text{or } \frac{Z_1}{Z_2} &= \frac{(R_1R_2 + X_1X_2) + j(R_2X_1 - R_1X_2)}{R_2^2 + X_2^2} \\ &= \frac{R_1R_2 + X_1X_2}{R_2^2 + X_2^2} + \frac{j(R_2X_1 - R_1X_2)}{R_2^2 + X_2^2}. \end{aligned}$$

Multiplication and division of impedances is very much simpler using polar notation, since:

$$Z_1 \angle \theta_1 \times Z_2 \angle \theta_2 = Z_1 Z_2 \angle (\theta_1 + \theta_2) \quad \text{and} \quad \frac{Z_1 \angle \theta_1}{Z_2 \angle \theta_2} = \frac{Z_1}{Z_2} \angle (\theta_1 - \theta_2)$$

where  $Z_1 \angle \theta_1$  is  $Z_1$  expressed in polar form and  $Z_2 \angle \theta_2$  is  $Z_2$  expressed in polar form.

## Mathematics in action

### *Multiplication, division and powers of complex quantities expressed in polar form*

The rules for multiplying and dividing complex quantities expressed in polar form are as follows:

Consider two complex numbers expressed in polar form,  $Z_1 \angle \theta_1$  and  $Z_2 \angle \theta_2$ .

Multiplying these numbers gives:

$$Z_1 \angle \theta_1 \times Z_2 \angle \theta_2 = Z_1 Z_2 \angle (\theta_1 + \theta_2)$$

$$\text{whilst dividing these numbers gives: } \frac{Z_1 \angle \theta_1}{Z_2 \angle \theta_2} = \frac{Z_1}{Z_2} \angle (\theta_1 - \theta_2).$$

$$\text{Now } (Z_1 \angle \theta_1)^2 = Z_1 \angle \theta_1 \times Z_1 \angle \theta_1 = Z_1 Z_1 \angle (\theta_1 + \theta_1) = Z_1^2 \angle (2\theta_1)$$

$$\begin{aligned} \text{similarly } (Z_2 \angle \theta_2)^2 &= Z_2 \angle \theta_2 \times Z_2 \angle \theta_2 = Z_2 Z_2 \angle (\theta_2 + \theta_2) \\ &= Z_2^2 \angle (2\theta_2). \end{aligned}$$

$$\text{Thus } (Z \angle \theta)^2 = Z^2 \angle 2\theta$$

$$\text{also } (Z \angle \theta)^3 = Z^3 \angle 3\theta \text{ whilst } (Z \angle \theta)^4 = Z^4 \angle 4\theta.$$

$$\text{Hence } (Z \angle \theta)^n = Z^n \angle n\theta.$$

Now let us use polar notation to solve a problem and compare this method with rectangular notation:

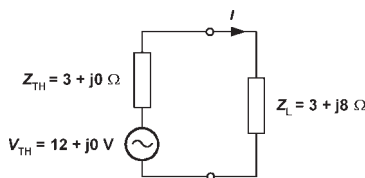


**Figure 4.1.62** See Example 4.1.13

#### Example 4.1.13

Determine the current flowing in the load,  $Z_L$ , in the circuit shown in Figure 4.1.62.

From Figure 4.1.62:

Thévenin equivalent circuit components are:

$$V_{TH} = 12 + j0 \text{ V and } Z_{TH} = 3 + j0 \Omega \text{ whilst } Z_L = 3 + j8 \Omega.$$

Firstly, using the rectangular form of impedance:

$$\begin{aligned} \text{The total impedance present, } Z_T &= Z_{TH} + Z_L \\ &= 3 + j0 + 3 + j8 = 6 + j8. \end{aligned}$$

The current flowing can be determined from,  $I = V_{TH}/Z_T$

$$\begin{aligned}\text{thus } I &= \frac{V_{TH}}{Z_T} = \frac{12 + j0}{6 + j8} = \frac{(12 + j0)(6 - j8)}{(6 + j8)(6 - j8)} \\ &= \frac{72 - j96}{36 + 64} = \frac{72 - j96}{100} = 0.72 - j0.96 \text{ A.}\end{aligned}$$

Secondly, using the polar form of impedance (see Example 4.1.12):

$$V_{TH} = 12\angle 0^\circ \text{ V and } Z_{TH} = 3\angle 0^\circ \Omega \text{ and } Z_L = 8.54\angle 69.4^\circ \Omega$$

$$\begin{aligned}\text{Once again, } I &= \frac{V_{TH}}{Z_T} = \frac{12\angle 0^\circ}{10\angle 53.1^\circ} \\ &= \frac{12}{10} \angle (0^\circ - 53.1^\circ) = 1.2\angle -53.1^\circ \text{ A.}\end{aligned}$$

(note that  $0.72 - j0.960 \equiv 1.2\angle -53.1^\circ$ ).

### Using the network theorems

The examples that follow illustrate how the network theorems can be applied to networks of complex impedances. All we need to do is replace resistances by impedances – expressed in either rectangular or polar form, whichever is more convenient – and remember that voltages and currents will also take a complex form (which may also be expressed in either rectangular or polar format).

The first example can be solved quite easily using the rectangular form. The remainder are solved by converting from rectangular to polar form, and vice versa, as required. When deciding which form to use, just recall that, in general, it is easier to:

- Add impedances when they are represented in rectangular form.
- Multiply and divide impedances when they are represented in polar form.

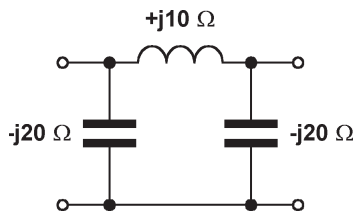


**Figure 4.1.63** See Example 4.1.14

### Example 4.1.14

Determine the T-network equivalent of the  $\pi$ -network shown in Figure 4.1.63.

The network components are:  $Z_A = Z_B = -j20 \Omega$  whilst  $Z_C = +j10 \Omega$ .

$$\text{Now } Z_1 = \frac{Z_A Z_C}{Z_A + Z_B + Z_C}$$

$$\text{hence } Z_1 = \frac{(-j20) \times (j10)}{-j20 - j20 + j10} = \frac{-j^2 200}{-j30} = \frac{j200}{30} = j6.67 \Omega$$

and since the network is symmetrical,  $Z_2 = Z_1 = j6.67 \Omega$ .

$$\text{Finally, } Z_3 = \frac{Z_A Z_B}{Z_A + Z_B + Z_C}$$

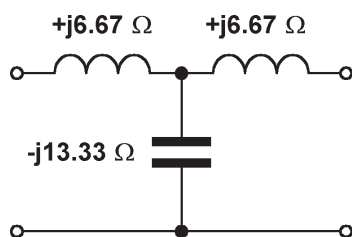


**Figure 4.1.64** *T-network equivalent of the  $\pi$ -network shown in Figure 4.1.63*

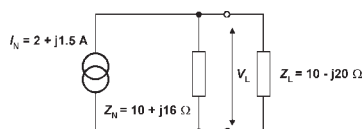


**Figure 4.1.65** *See Example 4.1.15*

$$\begin{aligned}\text{thus } Z_3 &= \frac{(-j20) \times (-j20)}{-j20 - j20 + j10} = \frac{j^2 400}{-j30} = \frac{j400}{-30} \\ &= -j13.33 \Omega.\end{aligned}$$

The T-equivalent of the  $\pi$ -network shown in Figure 4.1.63 is shown in Figure 4.1.64.

### Example 4.1.15

Determine the voltage developed across the load,  $Z_L$ , in the circuit shown in Figure 4.1.65.

From Figure 4.1.62, the Norton equivalent circuit components are:

$$I_N = 2 + j1.5 \text{ A and } Z_N = 10 + j16 \Omega \text{ whilst } Z_L = 10 - j20 \Omega.$$

Writing these in polar form (see Example 4.1.12) gives:

$$I_N = 2.5 \angle 36.9^\circ \text{ A, } Z_N = 18.87 \angle 60^\circ \Omega \text{ and } Z_L = 22.36 \angle -63.4^\circ \Omega.$$

The total impedance,  $Z_T$  in parallel with  $Z_L$ , is given by:

$$Z_T = \frac{Z_N \times Z_L}{Z_N + Z_L} \quad (\text{i})$$

Note that it is worth evaluating  $Z_N + Z_L$  in rectangular form and converting the result into polar form before entering values in Equation (i).

$$\begin{aligned}\text{Now } Z_N + Z_L &= 10 + j16 + 10 - j20 \\ &= 20 - j4 = 20.4 \angle 11.3^\circ \Omega.\end{aligned}$$

$$\begin{aligned}\text{Thus } Z_T &= \frac{Z_N \times Z_L}{Z_N + Z_L} = \frac{18.87 \angle 60^\circ \times 22.36 \angle -63.4^\circ}{20.4 \angle 11.3^\circ} \\ &= \frac{18.87 \times 22.36}{20.4} \angle (60^\circ - 63.4^\circ - 11.3^\circ)\end{aligned}$$

$$\text{or } Z_T = 20.68 \angle -14.7^\circ \Omega.$$

$$\begin{aligned}\text{Now } V_L &= I_N \times Z_T = 2.5 \angle 36.9^\circ \times 20.68 \angle -14.7^\circ \\ &= 2.5 \times 20.68 \angle (36.9^\circ - 14.7^\circ)\end{aligned}$$

$$\text{or } V_T = 51.7 \angle -22.2^\circ \text{ V.}$$

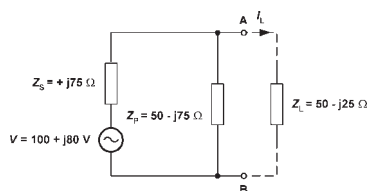


**Figure 4.1.66** *See Example 4.1.16*

### Example 4.1.16

Determine the Thévenin equivalent of the circuit shown in Figure 4.1.66. Hence determine the current developed in a load impedance of  $(50 - j25) \Omega$  connected between terminals A and B.

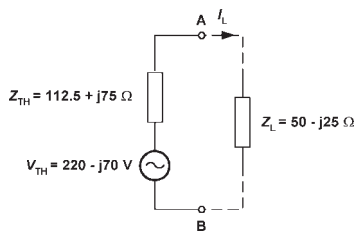


**Figure 4.1.67** Thévenin equivalent of Figure 4.1.66

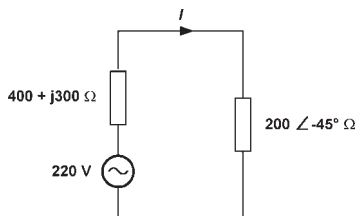


**Figure 4.1.68** See Questions 4.1.2

In order to determine the Thévenin equivalent of the circuit we need to find the open-circuit voltage developed between terminals A and B. This is the Thévenin equivalent voltage,  $V_{TH}$ .

$$V_{TH} = V \times \frac{Z_P}{Z_P + Z_S} = (100 + j80) \left( \frac{50 - j75}{50 - j75 + j75} \right) \\ = \frac{(100 + j80)(50 - j75)}{50}$$

Thus  $V_{TH} = 220 - j70$  V.

Converting this into polar form gives:

$$V_{TH} = 230.87 \angle -17.65^\circ \text{ V.}$$

Next we need to find the impedance seen looking into the network between A and B with the voltage source replaced by a short-circuit. This is the Thévenin equivalent impedance,  $Z_{TH}$ .

$$Z_{TH} = \frac{Z_S \times Z_P}{Z_P + Z_S} = \frac{(j75)(50 - j75)}{50 - j75 + j75} = \frac{(j75)(50 - j75)}{50}$$

Thus  $Z_{TH} = 112.5 + j75 \Omega$ .

The Thévenin equivalent circuit is shown in Figure 4.1.67. Now, to find the current flowing when terminals A and B are linked by a load impedance of  $(50 - j25) \Omega$  connected between terminals A and B, we need calculate the total impedance present,  $Z_T$ .

Now  $Z_T = Z_{TH} + Z_L$ .

Thus  $Z_T = 112.5 + j75 + 50 - j25 = 162.5 + j50 \Omega$ .

Converting the result into polar form gives:

$$Z_T = 170.02 \angle 17.1^\circ \Omega.$$

$$\text{Finally, } I = \frac{V_{TH}}{Z_T} = \frac{230.87 \angle -17.65^\circ}{170.02 \angle 17^\circ} = 1.36 \angle -34.7^\circ \text{ A.}$$

Converting this to rectangular form gives:

$$I = 1.12 - j0.77 \text{ A.}$$

### Questions 4.1.2

- (1) Express the following impedances in polar form:
  - (a)  $Z = 15 + j30 \Omega$ ; (b)  $Z = j12 \Omega$ ; (c)  $Z = 0.3 - j0.4 \Omega$ ;
  - (d)  $Z = -j1000 \Omega$ .
- (2) Express the following impedances in rectangular form:
  - (a)  $Z = 25 \angle 60^\circ \Omega$ ; (b)  $Z = 125 \angle -45^\circ \Omega$ ;
  - (c)  $Z = 650 \angle 85^\circ \Omega$ ;
- (3) Determine the current flowing in the circuit shown in Figure 4.1.68.

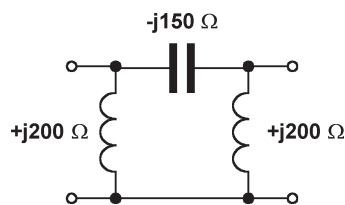


**Figure 4.1.69** See Questions 4.1.2

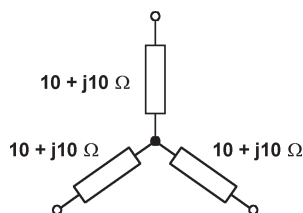


**Figure 4.1.70** See Questions 4.1.2

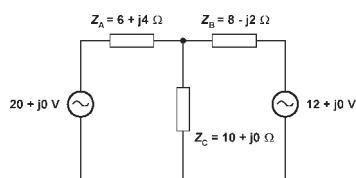


**Figure 4.1.71** See Questions 4.1.2

- (4) An AC source of  $(100 + j50)$  V and internal impedance  $(20 + j0) \Omega$  is connected, in turn, to each of the following loads:  
 (a)  $Z = 80 + j0$ ; (b)  $Z = 60 + j80$ ; (c)  $Z = 20 - j100$ .  
 Determine the current that will flow in each case.
- (5) Determine the T-network equivalent of the  $\pi$ -network shown in Figure 4.1.69.
- (6) Determine the delta-network equivalent of the star-network shown in Figure 4.1.70.
- (7) Use the superposition theorem to determine the current in  $Z_C$  in the figure shown in Figure 4.1.71.

## Maximum power transfer theorem

The ability to transfer maximum power from a circuit to a load is crucial in a number of electrical and electronic applications. The relationship between the internal resistance of a source, the resistance of the load to which it is connected, and the power dissipated in that load can be easily illustrated by considering a simple example:

### Example 4.1.17

A source consists of a 100 V battery having an internal resistance of  $100 \Omega$ . What value of load resistance connected to this source will receive maximum power?

The equivalent circuit of this arrangement is shown in Figure 4.1.72.

Now  $P = V^2/R$  therefore we might consider what power would appear in resistor  $R_L$  for different values of  $R_L$ . The power in  $R_L$  can be determined from:

$$P_L = V_L^2/R_L = (V \times R_L / (R_L + R_S))^2 / R_L = V^2 \times R_L / (R_L + R_S)^2.$$

Now, when  $R_L = 100 \Omega$ :

$$\begin{aligned} P_L &= (100^2 \times 100) / (100 + 100)^2 \\ &= 8.26 \text{ W.} \end{aligned}$$

This process can be repeated for other values of  $R_L$  (in, say, the range 0–200  $\Omega$ ) and a graph can be plotted from which the maximum value can be located, as shown in Figure 4.1.73.

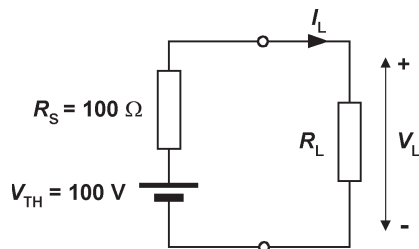


**Figure 4.1.72** See Example 4.1.17

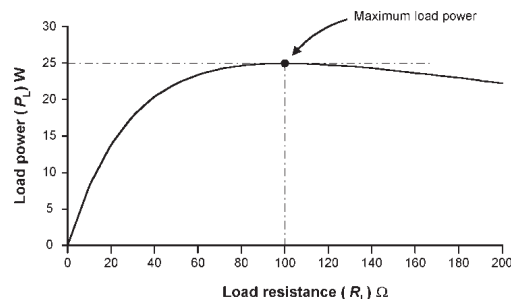


**Figure 4.1.73** See Example 4.1.17

Taking the maximum value for  $P_L$  from the graph, we can conclude that the value of  $R_L$  in which maximum power would be dissipated is  $100\ \Omega$ .

From the foregoing, it should be noted that maximum power is dissipated in the load when its resistance is equal to that of the internal resistance of the source, that is, when  $R_L = R_S$ . Note also that, in the extreme cases (i.e. when  $R_L = 0$  and also when  $R_L = \infty$ ) the power in the load,  $P_L = 0$ .

A more powerful approach would be to apply differential calculus to prove that the maximum value for  $P_L$  occurs when  $R_L = R_S$ . The method is as follows:

$$P_L = V^2 \times \frac{R_L}{(R_L + R_S)^2} = \frac{V^2 R_L}{(R_L + R_S)^2}.$$

Applying the quotient rule gives:

$$\frac{dP_L}{dR_L} = \frac{d}{dR_L} \left( \frac{V^2 R_L}{(R_L + R_S)^2} \right) = \frac{V^2 [(R_L + R_S)^2 - 2R_L(R_L + R_S)]}{(R_L + R_S)^4}.$$

For maximum power in RL the numerator must be zero. Hence:

$$(R_L + R_S)^2 - 2R_L(R_L + R_S) = 0.$$

$$\text{Thus } (R_L + R_S)^2 = 2R_L(R_L + R_S)$$

$$\text{or } R_L + R_S = 2R_L$$

$$\text{or } R_L = R_S.$$

The maximum power transfer theorem states that, in the case of DC circuits:

*Maximum power will be dissipated in a load when the resistance of the load is equal to the internal resistance of the source (i.e. when  $R_L = R_S$ ).*

In the case of AC circuits in which the source and load impedances are not purely resistive:

*Maximum power will be dissipated in a load when the impedance of the load is the conjugate of the internal impedance of the source (i.e. when  $Z_L = R + jX$  and  $Z_S = R - jX$ , and vice versa).*

Note that, in the case of an AC circuit in which the load impedance is purely resistive, maximum power will be dissipated in a load when  $R_L = \sqrt{R_S^2 + X_S^2}$ . Furthermore, when both the source and load impedances are purely resistive, maximum power will be dissipated in a load when  $R_L = R_S$ .

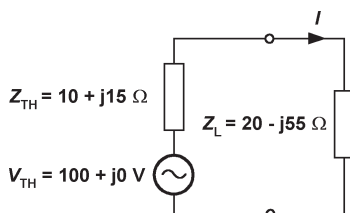


**Figure 4.1.74** See Example 4.1.18

### Example 4.1.18

Determine the power delivered to the load in the circuit shown in Figure 4.1.74. Also find the impedance that will receive maximum power.

The total impedance in the circuit,  $Z_T$ , will be given by:

$$Z_T = Z_{TH} + Z_L = 10 + j15 + 20 - j55 = 30 - j40 \Omega$$

and the resistive part of this impedance,  $R = 30 \Omega$ .

The current flowing will be given by:

$$I = \frac{V_{TH}}{Z_T} = \frac{100 + j0}{30 - j40} = 2 \angle 53^\circ \text{ A.}$$

The power dissipated in the load will be given by  $P_L = I^2 R$ .

$$\text{Thus } P_L = 2^2 \times 30 = 120 \text{ W.}$$

The impedance that will receive maximum power will simply be the conjugate of the impedance of the source. Since  $Z_{TH} = 10 + j15 \Omega$  the value of  $Z_L$  for maximum power transfer will be  $Z_L = 10 - j15 \Omega$ . In this condition, the power dissipated in the load will be:

$$P_L = 5^2 \times 10 = 250 \text{ W.}$$

When maximum power is transferred between a source and the load to which it is connected, they are said to be *matched*. The process of *matching* a source to a load is important in applications such as measurement, instrumentation, telecommunications, and data transmission.

Matching can often be achieved by adding resistance or reactance to a circuit in order to satisfy the criteria of the maximum power transfer theorem. The following example shows how this is done:

### Example 4.1.19

A transmitter having an output impedance of  $(40 - j15) \Omega$  is to be matched to an aerial system having an impedance of  $(50 + j35) \Omega$ . Determine the impedance required in a matching network that will be connected between the transmitter and the aerial system.

In order to obtain a perfect match between the aerial system and the transmitter, the optimum value of  $Z_S$  would be  $(50 - j35) \Omega$ . Hence the impedance of the series-connected matching network should comprise of a resistance of  $10 \Omega$  together with a reactance of  $-j20 \Omega$ .

The new source impedance,  $Z'_S$  then becomes:

$$Z'_S = (40 - j15) + (10 - j25) = 50 - j35 \Omega$$

(note that  $Z'_S$  is now equal to the conjugate of the impedance of the aerial system).

Figure 4.1.75 illustrates this arrangement.

Unfortunately, the problem with introducing additional impedances (sometimes referred to as *pads*) into a circuit in order to achieve a matched condition is that power will be lost in any resistive matching components. In some applications, this loss of power can be tolerated



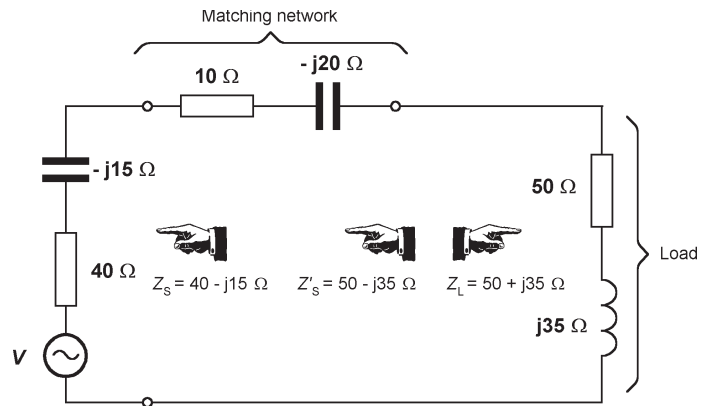


Figure 4.1.75 See Example 4.1.19

however, in many applications power loss can be undesirable and an alternative method of matching is required.

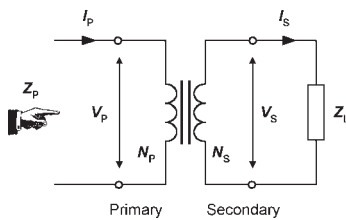


Figure 4.1.76 Transformer matching

### Transformer matching

Later, in this chapter we will explain the principle of the transformer but, for the moment, we will simply assume that a transformer is 'ideal' and can provide us with a virtually loss-free method of coupling an AC from one circuit to another. Transformers also provide us with a means of matching a source to a load. This is because an impedance connected to the secondary of a transformer becomes 'reflected' into its primary circuit. Consider the circuit shown in Figure 4.1.76.

The impedance seen 'looking into' the transformer's primary winding,  $Z_P$ , will be given by:

$$Z_P = V_P / I_P \quad (i)$$

whereas the load impedance connected to the secondary winding of the transformer,  $Z_L$ , will be given by:

$$Z_L = V_S / I_S. \quad (ii)$$

Now the voltages, currents and 'turns ratio',  $N_P / N_S$ , are connected by the relationship:

$$\frac{V_P}{V_S} = \frac{I_S}{I_P} = \frac{N_P}{N_S} \quad (iii)$$

$$\text{Thus } V_P = \frac{N_P}{N_S} \times V_S.$$

$$\text{From (i) and (iii), } Z = \frac{N_P}{N_S} \times \frac{V_S}{I_P}.$$

$$\text{But } V_S = I_S \times Z_L \text{ from (ii)}$$

$$\text{thus } Z = \frac{N_P}{N_S} \times \frac{I_S Z_L}{I_P}.$$

$$\text{Now } \frac{I_S}{I_P} = \frac{N_P}{N_S} \text{ from (iii).}$$

### Another view

It is worth noting that, when maximum power is transferred from a source to a load, the efficiency is only 50% since equal amounts of power are dissipated internally (i.e. in  $R_S$ ) and externally (i.e. in  $R_L$ ). In many practical applications this condition may not be desirable and we often take steps to ensure that the impedance of a source (e.g. a mains voltage outlet) is very much less than that of the load to which it is connected (e.g. an electric heater).

$$\text{Thus } Z = \frac{N_P}{N_S} \times \frac{N_P}{N_S} \times Z_L = \left( \frac{N_P}{N_S} \right)^2 \times Z_L$$

$$\text{hence } Z = \left( \frac{N_P}{N_S} \right)^2 \times Z_L \quad \text{or} \quad Z_L = \left( \frac{N_S}{N_P} \right)^2 Z.$$

### Example 4.1.20

A transformer having a turns ratio of 20:1 has its secondary connected to a load having an impedance of  $(6 + j8) \Omega$ . Determine the impedance seen looking into the primary winding.

$$\text{Now } Z = \left( \frac{N_P}{N_S} \right)^2 \times Z_L$$

where  $N_P/N_S = 20$  and  $Z_L = 6 + j8 \Omega$

thus  $Z = 20^2 (6 + j8) = 2400 + j3200 = 4000 \angle 53^\circ \Omega$ .

### Example 4.1.21

A transformer is required to match a load having an impedance of  $(600 + j150) \Omega$  to a source having an impedance of  $(90 + j10) \Omega$ . Determine the required turns ratio.

$$\text{Now } Z = \left( \frac{N_P}{N_S} \right)^2 \times Z_L$$

$$\text{thus } \frac{N_P}{N_S} = \sqrt{\frac{Z}{Z_L}} = \sqrt{\frac{600 + j150}{90 + j10}} = \sqrt{\frac{618.5}{90.6}} = 2.613.$$

The required turns ratio is 2.6:1.

### Questions 4.1.3

- (1) A transformer is required to match a loudspeaker of impedance  $(4 + j0) \Omega$  to the output stage of an amplifier which has an impedance of  $2 \text{ k}\Omega$ . Determine the required turns ratio of the transformer and the primary voltage required to produce an output power of 10 W in the loudspeaker.
- (2) An aerial matching transformer having 20 primary turns and 4 secondary turns is connected between a transmitter and an aerial. If the impedance seen looking into the primary of the transformer is  $(200 - j50) \Omega$  determine the impedance of the aerial.

## Mesh current analysis

In mesh current analysis, we identify a number of closed loops (or *meshes*) in which currents flow in the circuit and then use these to develop a set of equations based on Kirchhoff's voltage law. Consider the circuit with two voltage sources,  $V_1$  and  $V_2$ , shown in Figure 4.1.77.

In mesh A:  $V_1 = I_1(Z_1 + Z_2) - I_2Z_2$ .

In mesh B:  $0 = I_2(Z_1 + Z_2 + Z_3) - I_1Z_2 - I_3Z_4$ .

In mesh C:  $V_2 = -I_3(Z_4 + Z_5) + I_2Z_4$

where  $I_1$ ,  $I_2$  and  $I_3$  are the *mesh currents*. Note that the true current in  $Z_1$  is  $I_1$ , in  $Z_2$  it is  $(I_1 - I_2)$ , in  $Z_3$  it is  $I_2$ , in  $Z_4$  it is  $(I_2 - I_3)$ , and in  $Z_5$  it is  $I_3$ .

The use of mesh current analysis is best illustrated by an example:

### Example 4.1.22

Determine the current in each component in the circuit of Figure 4.1.77 if the components have the following values:

$$V_1 = 20 + j0, V_2 = 10 + j0, Z_1 = 10 + j0, Z_2 = 20 + j0, \\ Z_3 = 5 + j10, Z_4 = 10 + j0, \text{ and } Z_5 = 5 + j0.$$

Forming the mesh current equations for loops A, B and C gives:

In mesh A:  $V_1 = I_1(Z_1 + Z_2) - I_2Z_2$   
 $20 + j0 = I_1(10 + j0 + 20 + j0) - I_2(20 + j0)$   
 $20 = 30I_1 - 20I_2$  (i)

In mesh B:  $0 = I_2(Z_1 + Z_2 + Z_3) - I_1Z_2 - I_3Z_4$   
 $0 = I_2(10 + j0 + 20 + j0 + 5 + j10)$   
 $\quad - I_1(20 + j0) - I_3(10 + j0)$   
 $0 = I_2(35 + j10) - I_120 - I_310$  (ii)

In mesh C:  $V_2 = -I_3(Z_4 + Z_5) + I_2Z_4$   
 $10 + j0 = -I_3(10 + j0 + 5 + j0) + I_2(10 + j0)$   
 $10 = -I_315 + I_210$  (iii)

From (i)  $I_1 = (20 + 20I_2)/30 = 0.667 + 0.667I_2$ .

From (iii)  $I_3 = 0.667I_2 - 0.667$ .

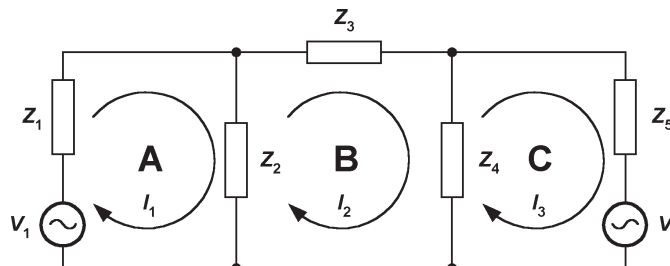


**Figure 4.1.77** Mesh current analysis

Substituting for  $I_1$  and  $I_3$  in (ii) gives:

$$0 = I_2(35 + j10) - 20(0.667 + 0.667I_2) - 10(0.667I_2 - 0.667)$$

$$0 = (35 + j10)I_2 - 13.34 - 13.33I_2 - 6.67I_2 + 6.67$$

$$13.33 - 6.67 = (35 - 13.33 - 6.67 + j10)I_2$$

$$\text{thus } 6.67 = (15 + j10)I_2$$

$$\text{thus } I_2 = 6.67/(15 + j10) = 5.55 - j12.02.$$

$$\text{From (i) } I_1 = 0.667 + 0.667(5.55 - j12.02)$$

$$\text{thus } I_1 = 0.667 + 3.7 - j8.02 = 4.37 - j8.02.$$

$$\text{From (iii) } I_3 = 0.667I_2 - 0.667$$

$$\begin{aligned} \text{thus } I_3 &= 0.667(5.55 - j12.02) - 0.667 \\ &= 3.7 - j8.02 - 0.667 = 3.03 - j8.02. \end{aligned}$$

Current in  $Z_1$  is  $I_1 = 4.37 - j8.02$  A.

$$\begin{aligned} \text{Current in } Z_2 \text{ is } I_1 - I_2 &= 4.37 - j8.02 - 5.55 + j12.02 \\ &= 11.18 + j4 \text{ A.} \end{aligned}$$

Current in  $Z_3$  is  $I_2 = 5.55 - j12.02$  A.

$$\begin{aligned} \text{Current in } Z_4 \text{ is } I_2 - I_3 &= 5.55 - j12.02 - 3.03 + j8.02 \\ &= 2.52 - j4 \text{ A.} \end{aligned}$$

Current in  $Z_5$  is  $I_3 = 3.03 - j8.02$  A.

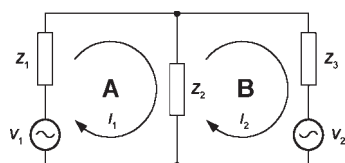


**Figure 4.1.78** See Example 4.1.23

The simultaneous equations in mesh current analysis can often be most easily solved by using matrices, as the following example shows:

### Example 4.1.23

Determine the current in each component, and the voltage dropped across  $Z_2$ , in the circuit shown in Figure 4.1.78. The components have the following values:

$$V_1 = 10 + j0, V_2 = 4 + j0, Z_1 = 6 + j0, Z_2 = 2 + j6, Z_3 = 2 + j0.$$

Forming the mesh current equations for loops A and B gives:

$$\begin{aligned} \text{In mesh A: } V_1 &= I_1 Z_1 + (I_1 - I_2) Z_2 \\ V_1 &= I_1 Z_1 + I_1 Z_2 - I_2 Z_2 \\ \text{or } V_1 &= I_1 (Z_1 + Z_2) - I_2 Z_2. \end{aligned} \tag{i}$$

$$\begin{aligned} \text{In mesh B: } V_2 &= -I_2 Z_3 + (I_1 - I_2) Z_2 \\ V_2 &= -I_2 Z_3 + I_1 Z_2 - I_2 Z_2 \\ \text{or } V_2 &= I_1 Z_2 - I_2 (Z_2 + Z_3). \end{aligned} \tag{ii}$$

$$\begin{aligned} \text{From (i) } 10 + j0 &= I_1(6 + j0 + 2 + j6) - I_2(2 + j6) \\ 10 &= I_1(8 + j6) - I_2(2 + j6). \end{aligned} \tag{iii}$$

$$\begin{aligned}\text{From (ii)} \quad 4 &= I_1(2 + j6) - I_2(2 + j6 + 2 + j0) \\ 4 &= I_1(2 + j6) - I_2(4 + j6).\end{aligned}\tag{iv}$$

$$\text{From (iii)} \quad 10 = (8 + j6) I_1 - (2 + j6) I_2.$$

$$\text{From (iv)} \quad 4 = (2 + j6) I_1 - (4 + j6) I_2.$$

Writing these simultaneous equations in matrix form gives:

$$\begin{bmatrix} (8 + j6) & -(2 + j6) \\ (2 + j6) & -(4 + j6) \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \end{bmatrix}.$$

Hence:

$$\frac{I_1}{\begin{vmatrix} 10 & -(2 + j6) \\ 4 & -(4 + j6) \end{vmatrix}} = \frac{I_2}{\begin{vmatrix} (8 + j6) & 10 \\ (2 + j6) & 4 \end{vmatrix}} = \frac{1}{\begin{vmatrix} (8 + j6) & -(2 + j6) \\ (2 + j6) & -(4 + j6) \end{vmatrix}}$$

$$\frac{I_1}{-10(4 + j6) + 4(2 + j6)} = \frac{I_2}{4(8 + j6) - 10(2 + j6)}$$

$$= \frac{1}{-(4 + j6)(8 + j6) + (2 + j6)^2}$$

$$\frac{I_1}{-32 - j36} = \frac{I_2}{12 - j36} = \frac{I}{-28 - j48}.$$

Since:

$$\frac{I_1}{-32 - j36} = \frac{1}{-28 - j48}$$

$$I_1 = \frac{-32 - j36}{-28 - j48} = \frac{32 + j36}{28 + j48} = 0.85 - j0.17 = 0.87 \angle -11^\circ \text{ A}.$$

Similarly:

$$\frac{I_2}{12 - j36} = \frac{1}{-28 - j48}$$

$$I_2 = \frac{12 - j36}{-28 - j48} = -\left( \frac{12 - j36}{28 + j48} \right)$$

$$= -(-0.45 - j0.51) = 0.45 + j0.51 = 0.68 \angle 48.6^\circ \text{ A}.$$

The current in  $Z_2$  is  $I_1 - I_2 = (0.85 - j0.17) - (0.45 + j0.51)$

or  $I_1 - I_2 = 0.4 - j0.68 = 0.79 \angle -59.5^\circ \text{ A}.$

The voltage dropped across  $Z_2$  will be given by:

$$\begin{aligned}(I_1 - I_2) \times Z_2 &= (0.79 \angle -59.5^\circ) \times (6.32 \angle 71.6^\circ) \\ &= 4.99 \angle 12.1^\circ = 4.88 + j1.05 \text{ V}.\end{aligned}$$

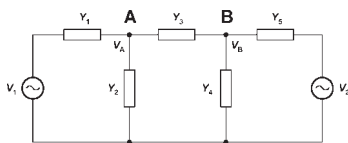


**Figure 4.1.79** Node voltage analysis

## Node voltage analysis

In node voltage analysis, we identify a number of junctions (or *nodes*) at which currents divide and then use these to develop a set of equations based on Kirchhoff's current law. Consider the circuit with two voltage sources,  $V_1$  and  $V_2$ , shown in Figure 4.1.79.

At node A:  $0 = (V_A - V_1)Y_1 + V_A Y_2 + (V_A - V_B)Y_3$ .

At node B:  $0 = (V_B - V_A)Y_3 + V_B Y_4 + (V_B - V_2)Y_5$ .

The use of node voltage analysis is best illustrated by an example:

### Example 4.1.24

Determine the voltage at each node in Figure 4.1.79 if the components have the following values:

$V_1 = 20 + j0$ ,  $V_2 = 4 + j0$ ,  $Z_1 = 10 + j0$ ,  $Z_2 = 2 + j6$ ,  
 $Z_3 = 2 - j2$ ,  $Z_4 = 4 + j3$ , and  $Z_5 = 2 + j0$ .

Applying Kirchhoff's current law at nodes A and B gives:

$$0 = (V_A - V_1)Y_1 + V_A Y_2 + (V_A - V_B)Y_3 \quad (\text{i})$$

and

$$0 = (V_B - V_A)Y_3 + V_B Y_4 + (V_B - V_2)Y_5. \quad (\text{ii})$$

Now:

$$Y_1 = \frac{1}{Z_1} = \frac{1}{10 + j0} = 0.1$$

$$Y_2 = \frac{1}{Z_2} = \frac{1}{2 + j6} = 0.05 - j0.15$$

$$Y_3 = \frac{1}{Z_3} = \frac{1}{2 - j2} = 0.25 + j0.25$$

$$Y_4 = \frac{1}{Z_4} = \frac{1}{4 + j3} = 0.16 - j0.12$$

$$Y_5 = \frac{1}{Z_5} = \frac{1}{2 + j0} = 0.5.$$

From (i):

$$0 = 0.1(V_A - (20 + j0)) + V_A(0.05 - j0.15) \\ + (V_A - V_B)(0.25 + j0.25)$$

$$0 = 0.1V_A - 2 + j0.05V_A - j0.15V_A + 0.25V_A + j0.25V_A \\ - 0.25V_B - j0.25V_B$$

$$0 = 0.4V_A - j0.1V_A - 0.25V_B - j0.25V_B - 2. \quad (\text{iii})$$

From (ii):

$$0 = (V_B - V_A)(0.25 + j0.25) + V_B(0.16 - j0.12) \\ + 0.5(V_B - (4 + j0))$$

$$0 = 0.25V_B + j0.25V_B - 0.25V_A - j0.25V_A + 0.16V_B \\ - j0.12V_B + 0.5V_B - 2$$

$$0 = -0.25V_A - j0.25V_A + 0.91V_B + j0.13V_B - 2. \quad (\text{iv})$$

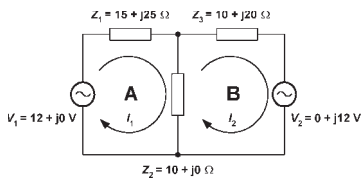


**Figure 4.1.80** See Questions 4.1.4

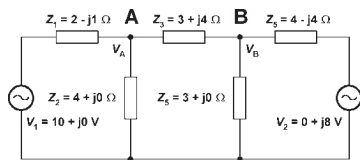


**Figure 4.1.81** See Questions 4.1.4

From (iii):

$$2 = (0.4 - j0.1)V_A - (0.25V_B + j0.25V_B).$$

From (iv)

$$2 = -(0.25 + j0.25)V_A + (0.91 + j0.13)V_B.$$

Writing these two simultaneous equations in matrix form gives:

$$\begin{bmatrix} (0.4 - j0.1) & -(0.25 + j0.25) \\ -(0.25 + j0.25) & (0.91 + j0.13) \end{bmatrix} \begin{bmatrix} V_A \\ V_B \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Hence:

$$\frac{V_A}{\begin{vmatrix} 2 & -(0.25 + j0.25) \\ 2 & (0.91 + j0.13) \end{vmatrix}} = \frac{V_B}{\begin{vmatrix} (0.4 - j0.1) & 2 \\ -(0.25 + j0.25) & 2 \end{vmatrix}} = \frac{1}{\begin{vmatrix} (0.4 - j0.1) & -(0.25 + j0.25) \\ -(0.25 + j0.25) & (0.91 + j0.13) \end{vmatrix}}$$

$$\frac{V_A}{\frac{2(0.91 + j0.13) + 2(0.25 + j0.25)}{1}} = \frac{V_B}{\frac{2(0.4 - j0.1) + 2(0.25 + j0.25)}{1}} = \frac{1}{(0.4 - j0.1)(0.91 + j0.13) - (0.25 + j0.25)^2}$$

$$\frac{V_A}{2.32 + j0.76} = \frac{V_B}{1.3 + j0.3} = \frac{1}{0.377 - j0.747}$$

$$\frac{V_A}{2.32 + j0.76} = \frac{1}{0.377 - j0.747}$$

$$\text{Thus } V_A = \frac{2.32 + j0.76}{0.377 - j0.747} = 0.438 + j2.88 = 2.92 \angle 81.36^\circ$$

$$\frac{V_B}{1.3 + j0.3} = \frac{1}{0.377 - j0.747}.$$

$$\text{Thus } V_B = \frac{1.3 + j0.3}{0.377 - j0.747} = 0.38 + j1.55 = 1.59 \angle 76.2^\circ.$$

### Questions 4.1.4

- (1) Determine the currents flowing in mesh A and B in Figure 4.1.80.
- (2) Determine the voltage that appears at nodes A and B in Figure 4.1.81.

## Coupled magnetic circuits

Any circuit in which a change of magnetic flux is produced by a change of current is said to possess *self-inductance*. In turn, the change of flux will produce an e.m.f. across the terminals of the circuit. We refer to this as an *induced e.m.f.*

When the flux produced by one coil links with another we say that the two coils exhibit *mutual inductance*,  $M$ . The more flux that links the two coils the greater the amount of mutual inductance that exists between them. This important principle underpins the theory of the transformer.

### Coefficient of coupling

Let us assume that two coils, having self-inductances,  $L_1$  and  $L_2$ , are placed close together. The fraction of the total flux from one coil that links with the other coil is called the *coefficient of coupling*,  $k$ .

If all of the flux produced by  $L_1$ , links with,  $L_2$ , then the magnetic coupling between them is perfect and we can say that the *coefficient of coupling*,  $k$ , is unity (i.e.  $k = 1$ ). Alternatively, if only half the flux produced by  $L_1$  links with  $L_2$ , then  $k = 0.5$ .

$$\text{Thus } k = \Phi_2/\Phi_1$$

where  $\Phi_1$  is the flux in the first coil ( $L_1$ ) and  $\Phi_2$  is the flux in the second coil ( $L_2$ ).

#### Example 4.1.25

A coil,  $L_1$ , produces a magnetic flux of 80 mWb. If 20 mWb appears in a second coil,  $L_2$ , determine the coefficient of coupling,  $k$ .

Now  $k = \Phi_2/\Phi_1$   
 where  $\Phi_1 = 80 \text{ mWb}$  and  $\Phi_2 = 20 \text{ mWb}$ .  
 Thus  $k = 20/80 = 0.25$ .

### Another view

In practice, inductive circuits are *coils*. By placing the coils close together or by winding them on the same closed magnetic core it is possible to obtain high values of mutual inductance. In this condition, we say that the two coils are *close coupled* or *tightly coupled*. Where the two coils are not in close proximity or are wound on separate magnetic cores, values of mutual inductance will be quite small. Under these circumstances we say that the two coils are *loose coupled*.

### Mutual inductance

We said earlier that, when two inductive circuits are coupled together by magnetic flux, they are said to possess *mutual inductance*. The unit of mutual inductance is the same as that for self-inductance and two circuits are said to have a mutual inductance of 1 Henry (H) if an e.m.f. of 1 V is produced in one circuit when the current in the other circuit varies at a uniform rate of 1 ampere per second (A/s). Hence, if two circuits have a mutual inductance,  $M$ , and the current in one circuit (the *primary*) varies at a rate  $di/dt$  (A/s) the e.m.f. induced in the second circuit (the *secondary*) will be given by:

$$e = -M \times (\text{rate of change of current}) \text{ or } e = -M \times \frac{di}{dt} \text{ V.} \quad (\text{i})$$

Note the minus sign which indicates that the e.m.f. in the second circuit opposes the increase in flux produced by the first circuit.

The e.m.f. induced in an inductor is given by:

$$e = -N \times (\text{rate of change of flux}) \text{ or } e = -N \times \frac{d\Phi}{dt} \text{ V}$$



where  $N$  is the number of turns and  $\Phi$  is the flux (in Wb).

Hence the e.m.f. induced in the secondary coil of two magnetically coupled inductors will be given by:

$$e = -N_2 \times \frac{d\Phi}{dt} \text{ V} \quad (\text{ii})$$

where  $N_2$  is the number of secondary turns. Combining Equations (i) and (ii) gives:

$$-M \times \frac{di}{dt} = -N_2 \times \frac{d\Phi}{dt}.$$

$$\text{Thus } M = N_2 \times \frac{d\Phi}{di} \times \frac{dt}{dt} = N_2 \times \frac{d\Phi}{di}.$$

If the relative permeability of the magnetic circuit remains constant, the rate of change of flux with current will also be a constant. Thus we can conclude that:

$$M = N_2 \times \frac{d\Phi}{di} = N_2 \times \frac{\Phi_2}{I_1} \quad (\text{iii})$$

where  $\Phi_2$  is the flux linked with the secondary circuit and  $I_1$  is the primary current.

Expression (iii) provides us with an alternative definition of mutual inductance. In this case,  $M$  is expressed in terms of the secondary flux and the primary current.

### Example 4.1.26

Two coils have a mutual inductance of 500 mH. Determine the e.m.f. produced in one coil when the current in the other coil is increased at a uniform rate from 0.2 A to 0.8 A in a time of 10 ms.

$$\text{Now } e = -M \times \frac{di}{dt}.$$

Since the current changes from 0.2 A to 0.8 A in 10 ms the rate of change of current with time,  $\frac{di}{dt}$ , is  $\frac{0.8 - 0.2}{0.01} = 60 \text{ A/s}$ .

$$\text{Thus } e = -0.5 \times 60 = -30 \text{ V}.$$

### Example 4.1.27

An e.m.f. of 40 mV is induced in a coil when the current in a second coil is varied at a rate of 3.2 A/s. Determine the mutual inductance of the two coils.

$$\text{Now } e = -M \times \frac{di}{dt} \quad \text{thus } M = -\frac{e}{\left(\frac{di}{dt}\right)} = \frac{0.04}{3.2} = 12.5 \text{ mH}.$$

### Coupled circuits

Consider the case of two coils that are coupled together so that all of the flux produced by one coil links with the other. These two coils are, in effect, perfectly coupled. Since the product of inductance ( $L$ ) and current ( $I$ ) is equal to the product of the number of turns ( $N$ ) and the flux ( $\Phi$ ), we can deduce that:

$$L_1 I_1 = N_1 \Phi \text{ or } L_1 = N_1 \Phi / I_1 \quad (\text{i})$$

Similarly:

$$L_2 I_2 = N_2 \Phi \text{ or } L_2 = N_2 \Phi / I_2 \quad (\text{ii})$$

where  $L_1$  and  $L_2$  represent the inductance of the first and second coils, respectively,  $N_1$  and  $N_2$  represent the turns of the first and second coils, respectively, and  $\Phi$  is the flux shared by the two coils.

Since the product of reluctance ( $S$ ) and flux ( $\Phi$ ) is equal to the product of the number of turns ( $N$ ) and the current ( $I$ ), we can deduce that:

$$S \Phi = N_1 I_1 = N_2 I_2 \text{ or } \Phi = N_1 I_1 / S = N_2 I_2 / S.$$

$$\text{Hence, from (i): } L_1 = N_1 \Phi / I_1 = N_1 (N_1 I_1) / I_1 S = N_1^2 / S$$

$$\text{and from (ii): } L_2 = N_2 \Phi / I_2 = N_2 (N_2 I_2) / I_2 S = N_2^2 / S.$$

$$\text{Now } L_1 \times L_2 = (N_1^2 / S) \times (N_2^2 / S) = N_1^2 N_2^2 / S^2. \quad (\text{iii})$$

$$\text{Now } M = N_2 \times \frac{\Phi}{I_1} \text{ (see Equation (iii) on page 318).}$$

Multiplying top and bottom by  $N_1$  gives:

$$M = N_1 N_2 \Phi / N_1 I_1 = N_1 N_2 / S$$

Combining Equations (iii) and (iv) gives:

$$M = \sqrt{L_1 L_2}.$$

In deriving the foregoing expression we have assumed that the flux is perfectly coupled between the two coils. In practice, there will always be some leakage of flux. Furthermore, there may be cases when we do not wish to couple two inductive circuits tightly together. The general equation is:

$$M = k \sqrt{L_1 L_2}$$

where  $k$  is the coefficient of coupling (note that  $k$  is always less than 1).

#### Example 4.1.28

Two coils have self-inductances of 10 mH and 40 mH. If the two coils exhibit a mutual inductance of 5 mH, determine the coefficient of coupling.

$$\text{Now } M = k \sqrt{L_1 L_2} \quad \text{thus } k = \frac{M}{\sqrt{L_1 L_2}} = \frac{5}{\sqrt{400}} = 0.25.$$

### Series connection of coupled coils

There are two ways of series connecting inductively coupled coils, either in *series aiding* (so that the mutual inductance adds to the combined self-inductance of the two coils) or in *series opposing* (in which case the mutual inductance subtracts from the combined self-inductance of the two coils).

In the former, *series aiding* case, the effective inductance is given by:

$$L = L_1 + L_2 + 2M$$

whilst in the latter, *series opposing* case, the effective inductance is given by:

$$L = L_1 + L_2 - 2M$$

where  $L_1$  and  $L_2$  represent the inductance of the two coils and  $M$  is their mutual inductance.

#### Example 4.1.29

When two coils are connected in series, their effective inductance is found to be 1.2 H. When the connections to one of the two coils are reversed, the effective inductance is 0.8 H. If the coefficient of coupling is 0.5, find the inductance of each coil and also determine the mutual inductance. In the first, series aiding case:

$$L = L_1 + L_2 + 2M \text{ or } 1.2 = L_1 + L_2 + 2M.$$

In the second, series opposing case:

$$L = L_1 + L_2 - 2M \text{ or } 0.8 = L_1 + L_2 - 2M.$$

$$\text{Now } M = k\sqrt{L_1 L_2} \text{ thus } M = 0.5\sqrt{L_1 L_2}.$$

Hence:

$$1.2 = L_1 + L_2 + 2(0.5\sqrt{L_1 L_2}) = L_1 + L_2 + \sqrt{L_1 L_2} \quad (\text{i})$$

and

$$0.8 = L_1 + L_2 - 2(0.5\sqrt{L_1 L_2}) = L_1 + L_2 - \sqrt{L_1 L_2}. \quad (\text{ii})$$

Adding (i) and (ii) gives:

$$1.2 + 0.8 = L_1 + L_2 + \sqrt{L_1 L_2} + L_1 + L_2 - \sqrt{L_1 L_2}.$$

Hence:

$$2 = 2(L_1 + L_2) \text{ or } 1 = L_1 + L_2 \text{ thus } L_1 = 1 - L_2.$$

Substituting for  $L_1$  in Equation (i) gives:

$$1.2 = (1 - L_2) + L_2 + \sqrt{(1 - L_2) \times L_2}$$

$$1.2 = 1 + \sqrt{L_2 - L_2^2}$$

$$0.2 = \sqrt{L_2 - L_2^2}$$

$$0.04 = L_2 - L_2^2$$

$$0 = -0.04 + L_2 - L_2^2$$

$$0 = 0.04 - L_2 + L_2^2.$$

Solving this quadratic equation gives:

$$L_2 = 0.959 \text{ H or } 0.041 \text{ H}$$

thus

$$L_1 = 0.041 \text{ H or } 0.959 \text{ H.}$$

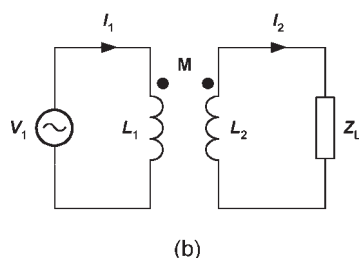
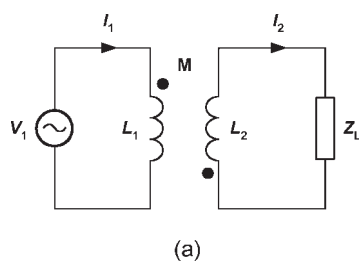


**Figure 4.1.82** Dot notation

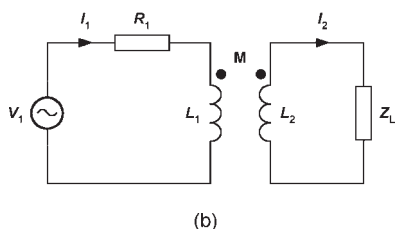
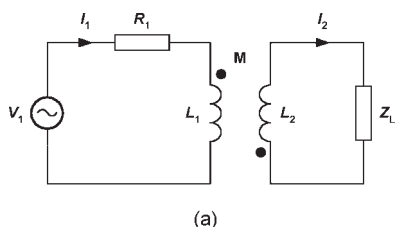


**Figure 4.1.83** Effect of primary winding resistance

### Dot notation

When two coupled circuits are drawn on a circuit diagram it is impossible to tell from the drawing the direction of the voltage induced in the second circuit (called the *secondary*) when a changing current is applied to the first circuit (known as the *primary*). When the direction of this induced voltage is important, we can mark the circuit with dots to indicate the direction of the currents and induced voltages, as shown in Figure 4.1.82.

In Figure 4.1.82(a), the total magneto-motive force produced will be the sum of the magneto-motive force produced by each individual coil. In this condition, the mutual inductance,  $M$ , will be *positive*.

In Figure 4.1.82(b), the total magneto-motive force produced will be the difference of the magneto-motive force produced by each individual coil. In this condition, the mutual inductance,  $M$ , will be *negative*.

Assuming that the primary coil in Figure 4.1.82(a) has negligible resistance, the primary voltage,  $V_1$ , will be given by:

$$V_1 = (I_1 \times j\omega L_1) + (I_2 \times j\omega M).$$

Conversely, in Figure 4.1.82(b), the primary voltage,  $V_1$ , will be given by:

$$V_1 = (I_1 \times j\omega L_1) - (I_2 \times j\omega M).$$

where  $I_1$  and  $I_2$  are the primary and secondary currents, respectively,  $L_1$  and  $L_2$  are the primary and secondary inductances, respectively,  $\omega$  is the angular velocity of the current, and  $M$  is the mutual inductance of the two circuits.

If the primary coil has resistance, as shown in Figure 4.1.83(a), the primary voltage,  $V_1$ , will be given by:

$$V_1 = I_1 (R_1 + j\omega L_1) + (I_2 \times j\omega M).$$

Conversely, in Figure 4.1.83(b), the primary voltage,  $V_1$ , will be given by:

$$V_1 = I_1 (R_1 + j\omega L_1) - (I_2 \times j\omega M).$$

## Transformers

### The e.m.f. equation for a transformer

If a sine wave current is applied to the primary winding of a transformer, the flux will change from its negative maximum value,  $-\Phi_M$ , to its positive

maximum value,  $+\Phi_M$ , in a time equal to half the periodic time of the current,  $t/2$ . The total change in flux over this time will be  $2\Phi_M$ .

The average rate of change of flux in a transformer is thus given by:

$$\text{Total flux change/Time} = 2\Phi_M/(t/2) = 4\Phi_M/t.$$

Now  $f = 1/t$  thus:

$$\text{Average rate of change of flux} = 4f\Phi_M.$$

The average e.m.f. induced per turn will be given by:

$$E/N = 4f\Phi_M.$$

Hence for a primary winding with  $N_1$  turns, the average induced primary voltage will be given by:

$$E_1 = 4N_1f\Phi_M.$$

Similarly, for a secondary winding with  $N_2$  turns, the average induced secondary voltage will be given by:

$$E_2 = 4N_2f\Phi_M.$$

(Note that we have assumed that there is perfect flux linkage between the primary and secondary windings, i.e.  $k = 1$ .)

For a sinusoidal voltage, the effective or rms value is 1.11 times the average value. Thus:

$$E_1 = 4.44N_1f\Phi_M \quad (\text{i})$$

and

$$E_2 = 4.44N_2f\Phi_M. \quad (\text{ii})$$

From Equations (i) and (ii):

$$E_1/E_2 = N_1/N_2.$$

Assuming that the transformer is 'loss-free', the power delivered to the secondary circuit will be the same as that in the primary circuit. Thus:

$$P_1 = P_2 \text{ or } E_1 \times I_1 = E_2 \times I_2.$$

Thus:

$$E_1/E_2 = I_2/I_1 = N_1/N_2.$$

### Example 4.1.30

A transformer operates with a 220 V 50 Hz supply and has 800 primary turns. Determine the maximum value of flux present.

Now  $E_1 = 220$  V,  $N_1 = 800$ , and  $f = 50$  Hz.

Since  $E_1 = 4.44N_1f\Phi_M$  V

$$\begin{aligned} \Phi_M &= E_1/4.44N_1f = 220/(4.44 \times 800 \times 50) = 220/177\,600 \\ &= 1.24 \text{ mWb.} \end{aligned}$$

### Example 4.1.31

A transformer has 455 primary turns and 66 secondary turns. If the primary winding is connected to a 110 V AC supply and

the secondary is connected to an  $8\ \Omega$  load, determine:

- the current that will flow in the secondary;
- the supply current.

You may assume that the transformer is 'loss-free'.

Now  $E_1 = 110\text{ V}$ ,  $N_1 = 455$  and  $N_2 = 66$

$E_1/E_2 = I_2/I_1 = N_1/N_2$  thus

$$E_2 = E_1 \times (N_2/N_1) = 110 \times 66/455 = 16\text{ V}.$$

The current delivered to an  $8\ \Omega$  load will thus be:

$$I_2 = E_2/R = 16/8 = 2\text{ A}.$$

$$I_1 = I_2 \times (N_2/N_1) = 2 \times 66/455 = 0.29\text{ A}.$$

### Equivalent circuit of a transformer

An *ideal transformer* would have coil windings that have negligible resistance and a magnetic core that is perfect. Furthermore, all the flux produced by the primary winding would be coupled into the secondary winding and no flux would be lost in the space surrounding the transformer.

In practice, a real transformer suffers from a number of imperfections. These are summarised below:

#### Leakage flux

Not all of the magnetic flux produced by the primary winding of a transformer is coupled into its secondary winding. This is because some flux is lost into the space surrounding the transformer (even though this has a very much higher reluctance than that of the magnetic core).

The leakage flux increases with the primary and secondary current and its effect is the same as that produced by an inductive reactance connected in series with the primary and secondary windings. We can take this into account by including this additional inductance to the simplified equivalent circuit of a transformer shown in Figure 4.1.84.

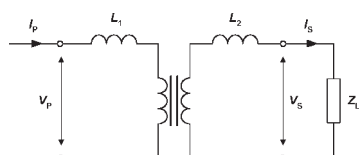


**Figure 4.1.84** Transformer equivalent circuit showing the effect of leakage flux

#### Winding resistance

The windings of a real transformer exhibit both inductance and resistance. The primary resistance,  $R_1$ , effectively appears in series with the primary inductance,  $L_1$ , whilst the secondary resistance,  $R_2$ , effectively appears in series with the secondary inductance,  $L_2$  (see Figure 4.1.85).

It is sometimes convenient to combine the primary and secondary resistances and inductances into a single pair of components connected in series with the primary circuit. We can do this by *referring* the secondary resistance and inductance to the primary circuit.

The amount of secondary resistance referred into the primary,  $R'_2$ , is given by:

$$R'_2 = R_2 \left( \frac{V_1}{V_2} \right)^2$$

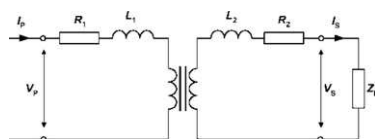


**Figure 4.1.85** Transformer equivalent circuit showing the effect of leakage flux and winding resistance

thus the effective total primary resistance is given by:

$$R_e = R_1 + R_2 \left( \frac{V_1}{V_2} \right)^2.$$

Similarly, the amount of secondary inductance referred into the primary,  $L'_2$ , is given by:

$$L'_2 = L_2 \left( \frac{V_1}{V_2} \right)^2$$

thus the effective total primary inductance is given by:

$$L_e = L_1 + L_2 \left( \frac{V_1}{V_2} \right)^2.$$

The effective impedance,  $Z_e$ , appearing in series with the primary circuit (with both  $R_2$  and  $L_2$  referred into the primary) is thus given by:

$$Z_e = R_e + j\omega L_e.$$

(Note that  $|Z_e| = \sqrt{(R_e^2 + (\omega L_e)^2)}$  – see Example 4.1.12 on page 301).

### Magnetising current

In a real transformer, a small current is required in order to magnetise the transformer core. This current is present regardless of whether or not the transformer is connected to a load. This magnetising current is equivalent to the current that would flow in an inductance,  $L_m$ , connected in parallel with the primary winding.

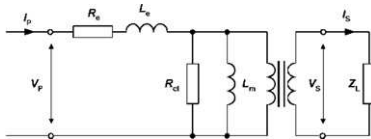


**Figure 4.1.86** Transformer equivalent circuit showing the effect of leakage flux, winding resistance, and core losses

### Core losses

A real transformer also suffers from losses in its core. These losses are attributable to *eddy currents* (small currents induced in the laminated core material) and *hysteresis* (energy loss in the core due to an imperfect  $B$ – $H$  characteristic). These losses can be combined together and represented by a resistance,  $R_{cl}$ , connected in parallel with the primary winding. The complete equivalent circuit for a transformer is shown in Figure 4.1.86.

### Example 4.1.32

A transformer designed for operation at 50 Hz has 220 primary turns and 110 secondary turns. The primary and secondary resistances are  $1.5 \Omega$  and  $0.5 \Omega$ , respectively, whilst the primary and secondary leakage inductances are 20 mH and 10 mH, respectively. Assuming core losses are negligible, determine the equivalent impedance referred to the primary circuit.

$$\text{Now } Z_e = \sqrt{R_e^2 + (\omega L_e)^2},$$

$$\text{where } R_e = R_1 + R_2 \left( \frac{V_1}{V_2} \right)^2 \quad \text{and} \quad L_e = L_1 + L_2 \left( \frac{V_1}{V_2} \right)^2.$$

$$\text{Also } \left( \frac{V_1}{V_2} \right)^2 = \left( \frac{N_1}{N_2} \right)^2 = \left( \frac{220}{110} \right)^2 = 4.$$

$$\text{Thus } R_e = 1.5 + (0.5 \times 4) = 3.5 \, \Omega \text{ and}$$

$$L_e = 0.02 + (0.01 \times 4) = 0.06 \, \text{H}.$$

$$\begin{aligned} \text{Finally, } Z_e &= R_e + j\omega L_e = 3.5 + j(2\pi \times 50 \times 0.06) \\ &= 3.5 + j18.84 \, \Omega \end{aligned}$$

$$|Z_e| = \sqrt{3.5^2 + 18.84^2} = 19.16 \, \Omega.$$

### Questions 4.1.5

- (1) A flux of  $60 \, \mu\text{Wb}$  is produced in a coil. A second coil is wound over the first coil. Determine the flux that will be induced in this coil if the coefficient of coupling is 0.3.
- (2) Two coils exhibit a mutual inductance of  $40 \, \text{mH}$ . Determine the e.m.f. produced in one coil when the current in the other coil is increased at a uniform rate from  $100 \, \text{mA}$  to  $600 \, \text{mA}$  in  $0.50 \, \text{ms}$ .
- (3) An e.m.f. of  $0.2 \, \text{V}$  is induced in a coil when the current in a second coil is varied at a rate of  $1.5 \, \text{A/s}$ . Determine the mutual inductance that exists between the two coils.
- (4) Two coils have self-inductances of  $60 \, \text{mH}$  and  $100 \, \text{mH}$ . If the two coils exhibit a mutual inductance of  $20 \, \text{mH}$ , determine the coefficient of coupling.
- (5) Two coupled coils have negligible resistance and mutual inductance of  $20 \, \text{mH}$ . If a supply of  $22 \angle 90^\circ \, \text{V}$  is applied to one coil and a current of  $0.5 \, \text{A}$  flows in the other coil, determine the current in the first coil if it is found to have an inductance of  $60 \, \text{mH}$ .
- (6) A transformer operates from a  $110 \, \text{V}$   $60 \, \text{Hz}$  supply. If the transformer has 400 primary turns, determine the maximum flux present. By how much will the flux be increased if the transformer was to be used on a  $50 \, \text{Hz}$  supply?
- (7) A 'loss-free' transformer has 900 primary turns and 225 secondary turns. If the primary winding is connected to a  $220 \, \text{V}$  supply and the secondary is connected to an  $11 \, \Omega$  resistive load, determine:
  - (a) the secondary voltage;
  - (b) the secondary current;
  - (c) the supply power.

## Resonant circuits

In Chapter 3 we introduced series and parallel resonance and derived the expressions for resonant frequency, dynamic resistance, and  $Q$ -factor. This chapter considers the behaviour of resonant circuits at frequencies above and below resonance and the effects of loading on tuned circuit performance.



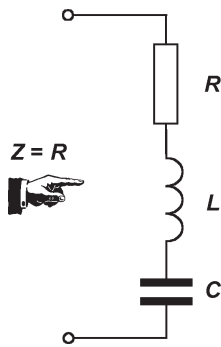


**Figure 4.1.87** A series  $R$ – $L$ – $C$  circuit

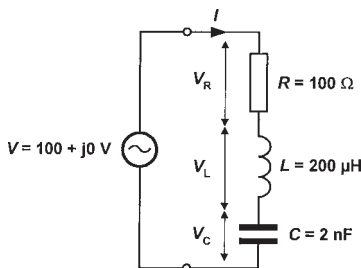


**Figure 4.1.88** See Example 4.1.33

The impedance of the  $R$ – $L$ – $C$  series circuit shown in Figure 4.1.87 is given by the expression:

$$Z = R + j\omega L - j/\omega C = R + j(\omega L - 1/\omega C).$$

At resonance,  $\omega L = 1/\omega C$ , thus  $Z = R + j(0) = R$ .

In this condition, the circuit behaves like a pure resistor and thus the supply current and voltage are in-phase.

At any frequency,  $f$ , we can determine the current flowing in a series resonant circuit and also the voltage dropped across each of the circuit elements present, as shown in the following example:

### Example 4.1.33

Determine the voltage dropped across each component in the circuit shown in Figure 4.1.88.

At 100 kHz, the inductive reactance is given by:

$$j\omega L = j \times 2\pi f \times L = j \times 6.28 \times 100 \times 10^3 \times 200 \times 10^{-6} \\ = j125.6 \Omega.$$

Similarly, the capacitive reactance is given by:

$$-\frac{j}{\omega C} = -j/(2\pi f \times C) \\ = -j/(6.28 \times 100 \times 10^3 \times 2 \times 10^{-9}) = -j796.2 \Omega.$$

The impedance of the circuit is given by:

$$Z = R + j(\omega L - 1/\omega C) = 100 + j(125.6 - 796.2) \\ = (100 - j670.6) \Omega.$$

The current flowing in the circuit will be given by:

$$I = V/Z = (100 + j0)/(100 - j670.6) = 0.022 + j0.146 \text{ A}.$$

The voltage dropped across the capacitor will be given by:

$$V_C = I \times \frac{j}{\omega C} = (0.022 + j0.146) \times -j796.2 \\ = (116.25 - j17.52) \text{ V}.$$

The voltage dropped across the inductor will be given by:

$$V_L = I \times j\omega L = (0.022 + j0.146) \times j125.6 \\ = (-18.33 + j2.76) \text{ V}.$$

The voltage dropped across the resistor will be given by:

$$V_R = I \times R = (0.022 + j0.146) \times 100 = (2.2 + j14.6) \text{ V}.$$

(You might like to check these answers by adding the three individual component voltages to see if they are equal to the supply voltage – they should be!)

### The universal resonance curve

The resonance curves shown in Figure 4.1.89 are graphs that show the ratio of actual current ( $I$ ) to current at resonance ( $I_0$ ) plotted against the proportion of the resonant frequency (i.e. the ratio  $f/f_0$ ) for a set of different

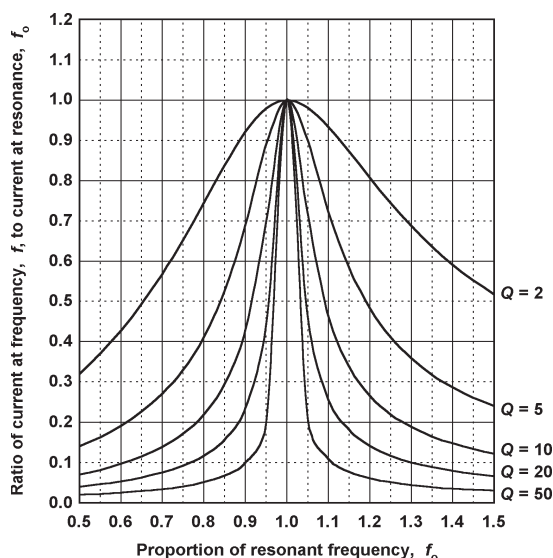


**Figure 4.1.89** Resonance curves for various  $Q$ -factors

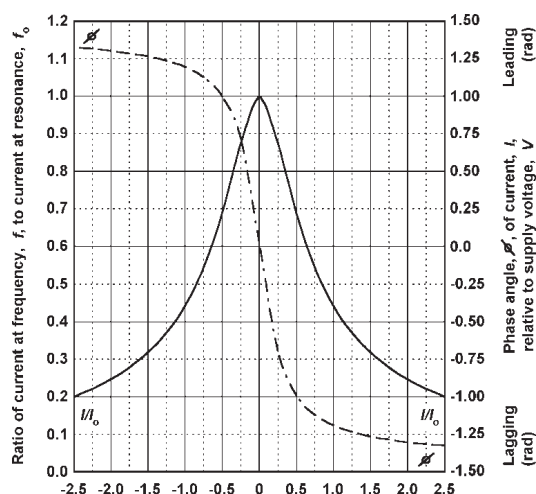


**Figure 4.1.90** Universal resonance curve

$Q$ -factors. Provided that we know the  $Q$ -factor of a tuned circuit *and it just happened to be the same* as one of those shown in Figure 4.1.89 we could use one of the graphs to determine the current flowing in the circuit at any frequency.

An alternative type of resonance curve (see Figure 4.1.90) can be used for any value of  $Q$ -factor. In this case, the horizontal axis shows values of  $\alpha$  where:

$$\alpha = Q \times \frac{\Delta f}{f_0}$$

where  $Q$  is the  $Q$ -factor of the circuit ( $Q = \omega L/R$ ),  $\Delta f$  is the deviation from the resonant frequency,  $f_0$ . This graph is known as a *universal resonance curve* since it can be made to apply to any resonant circuit.

### Example 4.1.34

A tuned circuit has a  $Q$ -factor of 10 and a resonant frequency of 100 kHz. If the current at resonance is 100 mA, use the

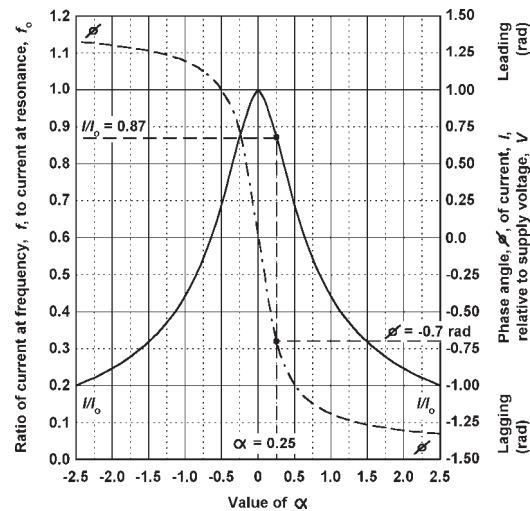


**Figure 4.1.91** See Example 4.1.34

universal resonance curve to determine the current and phase angle at 102.5 kHz.

At 102.5 kHz,  $\Delta f = f - f_0 = 102.5 - 100 = 2.5$  kHz.

Since  $Q = 10$  and  $f_0 = 100$  kHz,  $\alpha = 10 \times (2.5/100) = 0.25$ .

Referring to the universal resonance curve shown in Figure 4.1.91 shows that when  $\alpha = 0.25$ ,  $I/I_0 = 0.87$  and  $\phi = -0.7$  rad. Thus the current flowing,  $I$ , will be given by:  $I = 100 \times 0.87 = 87$  mA and the phase angle,  $\phi = -0.7$  rad  $= -40.1^\circ$ .

### Example 4.1.35

At resonance, a current of 2.5 A flows in a circuit that is series resonant at a frequency of 400 Hz. At 440 Hz a current of 1.25 A flows in the same circuit. Use the universal resonance curve to determine the  $Q$ -factor of the circuit.

Now,  $I/I_0 = 1.25/2.5 = 0.5$ .

From the universal resonance curve (see Figure 4.1.90), when  $I/I_0 = 0.5$ ,  $\alpha = 0.7$ .

Since  $\alpha = Q \times \frac{\Delta f}{f_0}$ ,  $Q = \alpha \times \frac{f_0}{\Delta f}$ .

Now  $f_0 = 400$  Hz and  $\Delta f = 440 - 400 = 40$  Hz.

Thus  $Q = 0.7 \times \frac{400}{40} = 7$ .

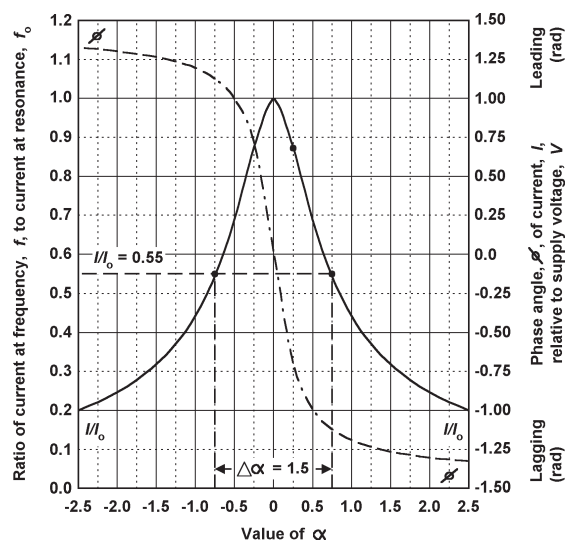


**Figure 4.1.92** See Example 4.1.36

### Example 4.1.36

Use the universal resonance curve to determine the bandwidth of a series tuned circuit having a  $Q$ -factor of 20 and a resonant frequency of 1 MHz at the points on the response curve where the current has fallen to 55% of its maximum.

Now,  $I/I_0 = 0.55$ .

From the universal resonance curve (see Figure 4.1.92), the points at which  $I/I_0 = 0.55$  correspond to values of  $\alpha$  of  $+0.75$  and  $-0.75$ .

$$\text{Since } \alpha = Q \times \frac{\Delta f}{f_0}, \Delta f = \alpha \times \frac{f_0}{Q}.$$

$$\text{Thus } \Delta f = 0.75 \times \frac{1 \text{ MHz}}{20} = 37.5 \text{ kHz}.$$

The total bandwidth is thus equal to  $2 \times 37.5 = 75 \text{ kHz}$ .

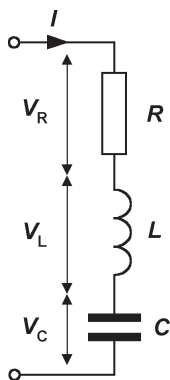


**Figure 4.1.93** Series resonant circuit

### Q-factor

In Chapter 3 we introduced the notion of the 'goodness' of a resonant circuit and we showed how this was related to the amount of 'magnification' that was produced by a circuit. We shall now develop this idea a little further.

For the series resonant circuit shown in Figure 4.1.93,  $Q$ -factor is the *voltage magnification factor* and it can be defined as:

$$Q = \frac{\text{voltage across } L}{\text{voltage across } R} = \frac{V_L}{V_R} = \frac{IX_L}{IR} = \frac{\omega L}{R},$$

where  $V_L$  and  $V_R$  are the voltages dropped across  $L$  and  $R$  at resonance and  $I$  is the current flowing in the circuit at resonance.

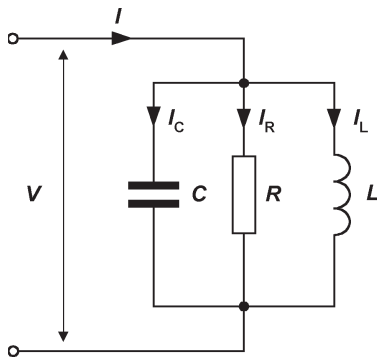


**Figure 4.1.94** Parallel resonant circuit

For the parallel resonant circuit shown in Figure 4.1.94,  $Q$ -factor is the current magnification factor and it can be defined as:

$$Q = \frac{\text{current in } L}{\text{current in } R} = \frac{I_L}{I_R} = \frac{V/X_L}{V/R} = \frac{VR}{VX_L} = \frac{R}{X_L} = \frac{R}{\omega L}$$

where  $I_L$  and  $I_R$  are the currents in  $L$  and  $R$  at resonance and  $V$  is the voltage developed across the circuit.

### Example 4.1.37

A parallel resonant circuit consists of  $L = 10 \mu\text{H}$ ,  $C = 20 \text{ pF}$ , and  $R = 10 \text{ k}\Omega$ . Determine the  $Q$ -factor of the circuit at resonance.

First we need to find the resonant frequency from:

$$\begin{aligned} f_0 &= \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{10 \times 10^{-6} \times 20 \times 10^{-12}}} \\ &= \frac{1}{2\pi\sqrt{200 \times 10^{-18}}} \end{aligned}$$

$$f_0 = \frac{0.159}{1.414 \times 10^{-8}} = 11.24 \times 10^6 = 11.24 \text{ MHz.}$$

At resonance,

$$X_L = 2\pi f_0 L = 6.28 \times 11.24 \times 10^6 \times 10 \times 10^{-6} = 705.9 \Omega.$$

$$\text{Now } Q = R/\omega L = 10\,000/705.9 = 14.2.$$

### Loading and damping

Loading of a series tuned circuit occurs whenever another circuit is coupled to it. This causes the tuned circuit to be *damped* and the result is a reduction in  $Q$ -factor together with a corresponding increase in bandwidth (recall that bandwidth  $= f_0/Q$ ).

In some cases, we might wish to deliberately introduce damping into a circuit in order to make it less selective, reducing the  $Q$ -factor and increasing the bandwidth. For a series tuned circuit, we must introduce the damping resistance in series with the existing components ( $L$ ,  $C$ , and  $R$ ) whereas, for a parallel tuned circuit, the damping resistance must be connected in parallel with the existing components.

### Questions 4.1.6

- (1) A series tuned circuit consists of  $L = 450 \mu\text{H}$ ,  $C = 1 \text{ nF}$ , and  $R = 30 \Omega$ . Determine:
  - (a) The frequency at which maximum current will flow in the circuit.

- (b) The  $Q$ -factor of the circuit at resonance.
- (c) The bandwidth of the tuned circuit.
- (2) If the tuned circuit in Question 1, above, is connected to a voltage source of  $(2 + j0)$  V at a frequency of 500 kHz, determine the current that will flow and the voltage that will be developed across each component.
- (3) A series tuned circuit has a  $Q$ -factor of 15 and is resonant at a frequency of 45 kHz. If a current of 50 mA flows in the circuit at resonance, use the universal resonance curve to determine:
  - (a) The current that will flow at a frequency of 51 kHz.
  - (b) The frequency at which the phase angle will be  $72^\circ$  (leading).
- (4) A current of 10 mA flows in a series  $L$ – $C$ – $R$  circuit at its resonant frequency of 455 kHz. If the current falls to 3.7 mA at a frequency of 480 kHz, determine the  $Q$ -factor of the tuned circuit.
- (5) A parallel resonant circuit comprises  $L = 200 \mu\text{H}$ ,  $C = 700 \text{ pF}$ , and  $R = 75 \text{ k}\Omega$ .
  - (a) Determine the resonant frequency,  $Q$ -factor and bandwidth of the tuned circuit.
  - (b) If the bandwidth is to be increased to 5 kHz, determine the value of additional parallel damping resistance required.

## 4.2 NETWORKS

In the previous section we introduced some simple networks (see Figure 4.1.37 on page 294). We also showed how it was possible to replace one type of network with another without affecting its electrical characteristics. In this section we shall develop this work further so that you can gain an understanding of the effect that networks have on signals.

### Basic network models

The simple networks that we met in the last section can exist in two basic forms, *unbalanced and balanced*. In the latter case, none of the network's input and output terminals are connected directly to common or ground. The unbalanced and balanced forms of the basic T- and  $\pi$ -networks are shown in Figure 4.2.1.

The networks shown in Figure 4.2.1 all have two *ports*. One port (i.e. pair of terminals) is connected to the input whilst the other is connected to the output. For convenience, many *two-port networks* are made symmetrical and they perform exactly the same function and have the same characteristics regardless of which way round they are connected.

### Characteristic impedance

It is often convenient to analyse the behaviour of a signal transmission path in terms of a number of identical series-connected networks. One important feature of any network is that, when an infinite number of identical symmetrical networks are connected in series, the resistance (or impedance) seen looking into the network will have a definite value. This value is known as the *characteristic impedance* of the network.

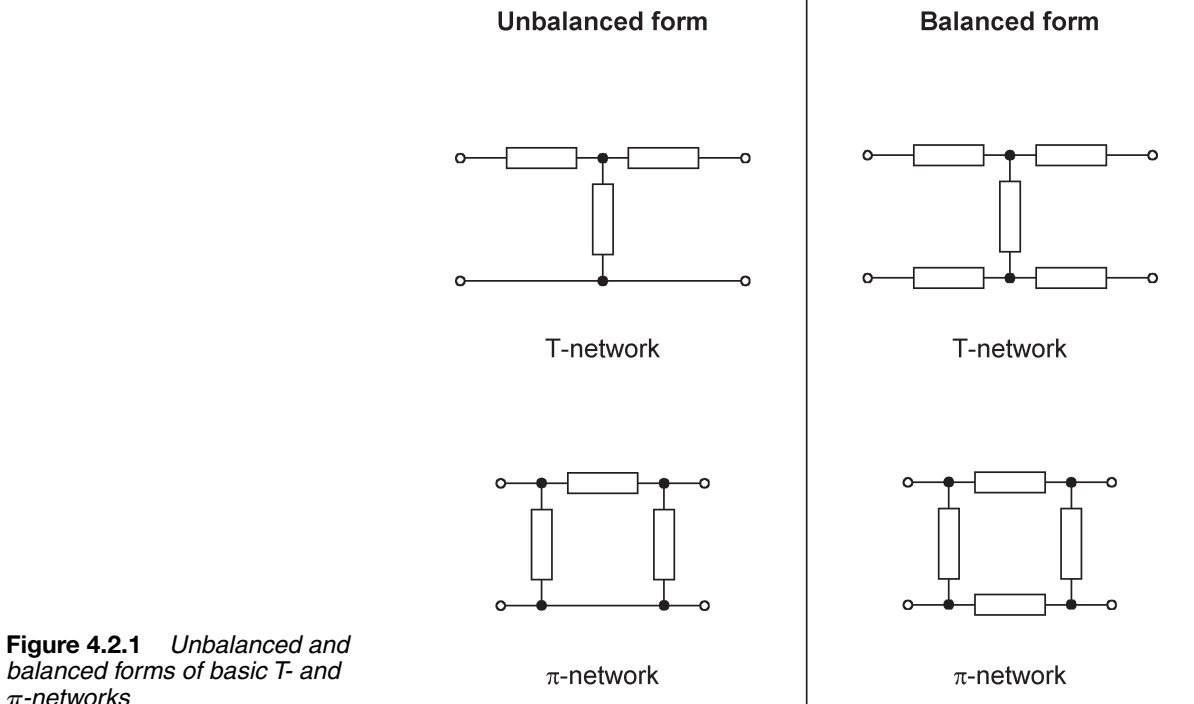


**Figure 4.2.1** Unbalanced and balanced forms of basic T- and  $\pi$ -networks

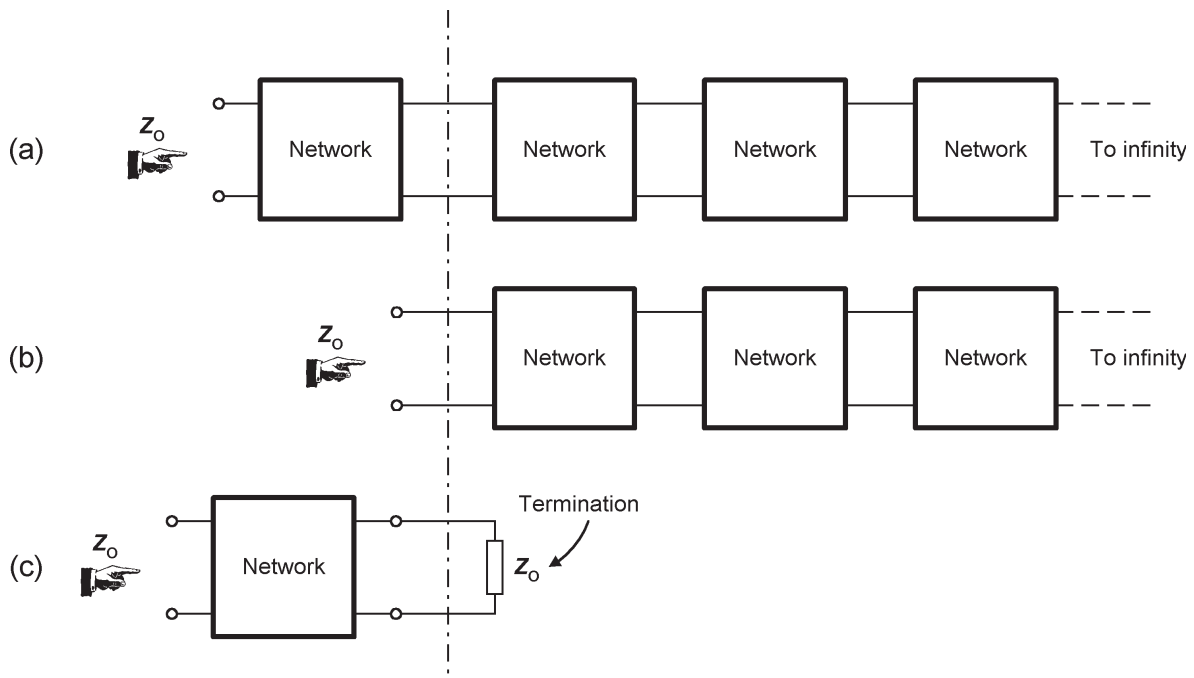


**Figure 4.2.2** Characteristic impedance seen looking into an infinite number of series-connected networks

Take a look at Figure 4.2.2. In Figure 4.2.2(a) an infinite number of identical networks are connected in series. By definition, the impedance seen looking into this arrangement will be equal to the characteristic impedance,  $Z_0$ . Now suppose that we remove the first network in the chain, as shown in Figure 4.2.2(b). To all intents and purposes, we will still be looking into an infinite number of series-connected networks. Thus, once again, we will see an impedance equal to  $Z_0$  when we look into the network.

Finally, suppose that we place an impedance of  $Z_0$  across the output terminals of the single network that we removed earlier. This *terminated* network (see Figure 4.2.2(c)) will behave *exactly* the same way as the arrangement in Figure 4.2.2(a). In other words, by correctly terminating the network in its characteristic impedance, we have made one single network section appear the same as a series of identical networks stretching to infinity.

It should be fairly obvious that the characteristic impedance of a network is determined by the values of resistance (or impedance) within the network. We shall now prove that this is the case and determine values of  $Z_0$  for the T- and  $\pi$ -networks that we met in the previous section.

### The T-network

Take a look at the network shown in Figure 4.2.3.

Now  $Z_0 = Z_1 + (Z_2 \parallel (Z_1 + Z_0))$  (where  $\parallel$  means 'in parallel with').

$$\text{Thus } Z_0 = Z_1 + \frac{Z_2(Z_1 + Z_0)}{Z_0 + Z_1 + Z_2}$$

$$Z_0 = Z_1 + \frac{Z_1 Z_2 + Z_0 Z_2}{Z_0 + Z_1 + Z_2} = Z_1 + \frac{Z_1 Z_2}{Z_0 + Z_1 + Z_2} + \frac{Z_0 Z_2}{Z_0 + Z_1 + Z_2}$$

$$\begin{aligned} Z_0 &= \frac{Z_1(Z_0 + Z_1 + Z_2) + Z_1 Z_2 + Z_0 Z_2}{Z_0 + Z_1 + Z_2} \\ &= \frac{Z_1 Z_0 + Z_1^2 + Z_1 Z_2 + Z_1 Z_2 + Z_0 Z_2}{Z_0 + Z_1 + Z_2} \end{aligned}$$

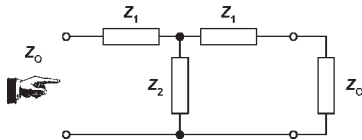


Figure 4.2.3 T-network

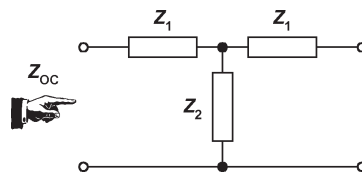


Figure 4.2.4 Impedance seen looking into a T-network with the output terminals left open-circuit

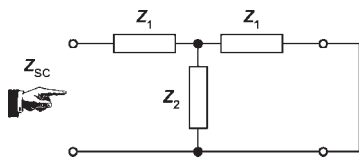


Figure 4.2.5 Impedance seen looking into a T-network with the output terminals short-circuited

$$Z_0(Z_0 + Z_1 + Z_2) = Z_1^2 + 2Z_1 Z_2 + Z_1 Z_0 + Z_0 Z_2$$

$$\text{or } Z_0^2 + Z_0 Z_1 + Z_0 Z_2 = Z_1^2 + 2Z_1 Z_2 + Z_1 Z_0 + Z_0 Z_2.$$

$$\text{Hence } Z_0^2 = Z_1^2 + 2Z_1 Z_2 \text{ or } Z_0 = \sqrt{Z_1^2 + 2Z_1 Z_2}. \quad (\text{i})$$

The impedance seen looking into the T-network with the output terminals left open-circuit (see Figure 4.2.4) is simply given by:

$$Z_{OC} = Z_1 + Z_2.$$

The impedance seen looking into the T-network with the output terminals linked together by a short-circuit (see Figure 4.2.5) is given by:

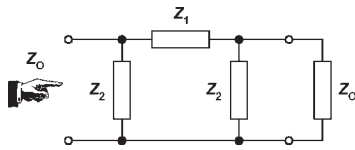
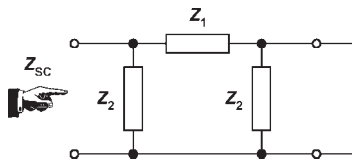
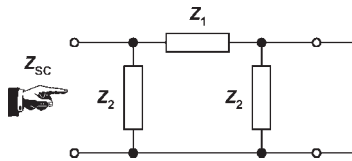
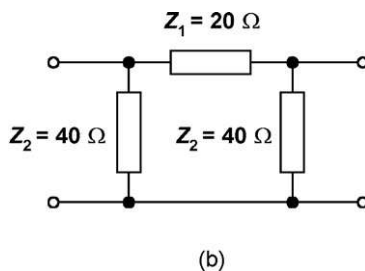
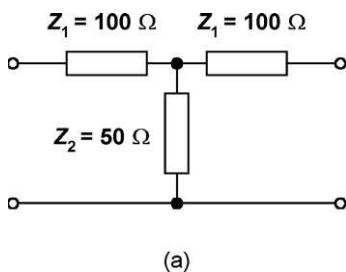
$$Z_{SC} = Z_1 + (Z_1 \parallel Z_2) \text{ (where } \parallel \text{ means 'in parallel with').}$$

$$\begin{aligned} Z_{SC} &= Z_1 + \frac{Z_1 Z_2}{Z_1 + Z_2} = Z_1 \frac{(Z_1 + Z_2)}{(Z_1 + Z_2)} + \frac{Z_1 Z_2}{(Z_1 + Z_2)} \\ &= \frac{Z_1^2 + Z_1 Z_2 + Z_1 Z_2}{(Z_1 + Z_2)} = \frac{Z_1^2 + 2Z_1 Z_2}{(Z_1 + Z_2)}. \end{aligned}$$

$$\text{Now } Z_{OC} \times Z_{SC} = (Z_1 + Z_2) \frac{Z_1^2 + 2Z_1 Z_2}{(Z_1 + Z_2)} = Z_1^2 + 2Z_1 Z_2. \quad (\text{ii})$$

Combining Equations (i) and (ii) gives:  $Z_0 = \sqrt{Z_{OC} Z_{SC}}.$



**Figure 4.2.6**  $\pi$ -network**Figure 4.2.7** Impedance seen looking into a  $\pi$ -network with the output terminal left open-circuit**Figure 4.2.8** Impedance seen looking into a  $\pi$ -network with the output terminal short-circuited**Figure 4.2.9** See Example 4.2.1

### The $\pi$ -network

Take a look at the network shown in Figure 4.2.6.

$$\text{Now } Z_0 = Z_2 \parallel \left( Z_1 + \left( \frac{Z_0 Z_2}{Z_0 + Z_2} \right) \right) \text{ (where } \parallel \text{ means 'in parallel with')}$$

$$\text{or } Z_0 = Z_2 \parallel \left( \frac{Z_1 Z_0 + Z_1 Z_2 + Z_0 Z_2}{Z_0 + Z_2} \right)$$

$$\text{thus } Z_0 = \frac{Z_2 (Z_1 Z_0 + Z_1 Z_2 + Z_0 Z_2) / (Z_0 + Z_2)}{Z_2 + ((Z_1 Z_0 + Z_1 Z_2 + Z_0 Z_2) / (Z_0 + Z_2))}$$

$$\text{From which } Z_0^2 = \frac{Z_1 Z_2^2}{Z_1 + 2Z_2} \text{ or } Z_0 = \sqrt{\frac{Z_1 Z_2^2}{Z_1 + 2Z_2}}. \quad (\text{i})$$

The impedance seen looking into the  $\pi$ -network with the output terminals left open-circuit (see Figure 4.2.7) is given by:

$$Z_{OC} = Z_2 \parallel (Z_1 + Z_2) \text{ (where } \parallel \text{ means 'in parallel with')}$$

$$Z_{OC} = \frac{Z_2 (Z_1 + Z_2)}{Z_2 + Z_1 + Z_2} = \frac{Z_2 (Z_1 + Z_2)}{Z_1 + 2Z_2}.$$

The impedance seen looking into the T-network with the output terminals linked together by a short-circuit (see Figure 4.2.8) is given by:

$$Z_{SC} = Z_2 \parallel Z_1 \text{ (where } \parallel \text{ means 'in parallel with')}$$

$$Z_{SC} = \frac{Z_1 Z_2}{Z_1 + Z_2}.$$

$$\text{Now } Z_{OC} \times Z_{SC} = \left( \frac{Z_2 (Z_1 + Z_2)}{Z_1 + 2Z_2} \right) \left( \frac{Z_1 Z_2}{Z_1 + Z_2} \right) = \frac{Z_1 Z_2^2}{Z_1 + 2Z_2}. \quad (\text{ii})$$

Combining Equations (i) and (ii) once again gives:  $Z_0 = \sqrt{Z_{OC} Z_{SC}}.$

### Example 4.2.1

Determine the characteristic impedance of each of the networks shown in Figure 4.2.9.

Figure 4.2.9(a) shows a T-network in which  $Z_1 = 100 \Omega$  and  $Z_2 = 50 \Omega$ .

$$\text{For a T-network, } Z_0 = \sqrt{Z_1^2 + 2Z_1 Z_2}.$$

$$\begin{aligned}\text{Thus } Z_0 &= \sqrt{100^2 + (2 \times 100 \times 50)} \\ &= \sqrt{20\,000} = 141.4 \, \Omega.\end{aligned}$$

Figure 4.2.9(b) shows a  $\pi$ -network in which  $Z_1 = 20 \, \Omega$  and  $Z_2 = 40 \, \Omega$ .

$$\text{For a } \pi\text{-network, } Z_0 = \sqrt{\frac{Z_1 Z_2^2}{Z_1 + 2Z_2}}.$$

$$\begin{aligned}\text{Thus } Z_0 &= \sqrt{\frac{20 \times 40^2}{20 + (2 \times 40)}} = \sqrt{\frac{20 \times 1600}{100}} \\ &= \sqrt{320} = 17.9 \, \Omega.\end{aligned}$$

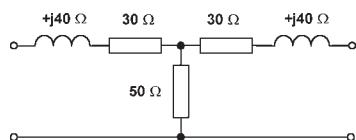


**Figure 4.2.10** See Example 4.2.2

### Example 4.2.2

Determine the characteristic impedance of the T-network shown in Figure 4.2.10.

Comparing this with the T-network shown in Figure 4.2.3 reveals that:

$$Z_1 = 30 + j40 \, \Omega \text{ and that } Z_2 = 50 \, \Omega.$$

$$\text{For a T-network, } Z_0 = \sqrt{Z_1^2 + 2Z_1 Z_2}.$$

$$\begin{aligned}\text{Thus } Z_0 &= \sqrt{(30 + j40)^2 + (2 \times (30 + j40) \times 50)} \\ &= \sqrt{8000 + j6400} \, \Omega.\end{aligned}$$

In order to find the square root of this complex quantity we shall convert it to polar form:

$$\text{Thus } Z_0 = \sqrt{10\,245 \angle 38.66^\circ} = 101.2 \angle 19.33 \, \Omega.$$

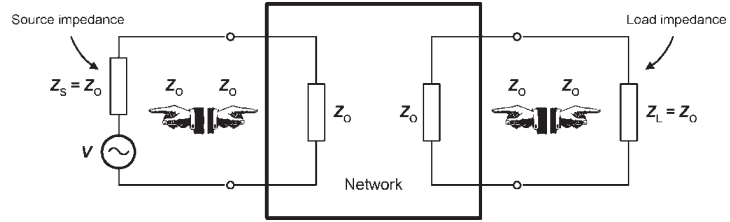
Converting this back to rectangular form gives:

$$Z_0 = 95.5 + j33.5 \, \Omega.$$

## Attenuators

One of the most common applications of the basic T- and  $\pi$ -networks is to reduce the amplitude of signals present in a matched system by a precise amount. The network is then said to *attenuate* the signal and the circuit is referred to as an *attenuator*. In order to work correctly (i.e. provide the required amount of attenuation) an attenuator needs to be *matched* to the system in which it is used. This simply means ensuring that the impedance of the source, as well as that of the load, matches the characteristic impedance of the attenuator. In this condition, we say that an attenuator is correctly *terminated*. Figure 4.2.11 illustrates this concept.

**Figure 4.2.11** Correctly matched attenuator



Before we take a look at the operation of two simple forms of attenuator, it is worth pointing out that the impedances used in attenuators are always pure resistances. The reason for this is that an attenuator must provide the same attenuation at all frequencies and the inclusion of reactive components (inductors and/or capacitors) would produce a non-linear attenuation/frequency characteristic.

### The T-network attenuator

The circuit of a correctly terminated T-network attenuator is shown in Figure 4.2.12.

From Figure 4.2.12:

$$V_3 = V_1 - I_1 R_1$$

$$\text{but } I_1 = \frac{V_1}{R_0}, \text{ thus } V_3 = V_1 - \left( \frac{V_1}{R_0} \right) \times R_1$$

$$\text{or } V_3 = V_1 \left( 1 - \frac{R_1}{R_0} \right). \quad (i)$$

$$\text{Also } V_2 = V_3 \times \frac{R_0}{R_0 + R_1} \text{ or } V_3 = V_2 \left( \frac{R_0 + R_1}{R_0} \right). \quad (ii)$$

$$\text{Equating (i) and (ii) gives: } V_1 \left( 1 - \frac{R_1}{R_0} \right) = V_2 \left( \frac{R_0 + R_1}{R_0} \right).$$

$$\text{Thus } V_1 \left( \frac{R_0 - R_1}{R_0} \right) = V_2 \left( \frac{R_0 + R_1}{R_0} \right).$$

The attenuation,  $\alpha$ , provided by the attenuator is given by  $V_1/V_2$ , hence:

$$\alpha = \frac{V_1}{V_2} = \left( \frac{R_0}{R_0 - R_1} \right) \left( \frac{R_0 + R_1}{R_0} \right) = \frac{R_0 + R_1}{R_0 - R_1}.$$

$$\text{Thus } \alpha = \frac{R_0 + R_1}{R_0 - R_1}.$$

$$\text{Now } R_1 + R_0 = \alpha(R_0 - R_1) = \alpha R_0 - \alpha R_1.$$

$$\text{Thus } R_1 + \alpha R_1 = \alpha R_0 - R_0$$

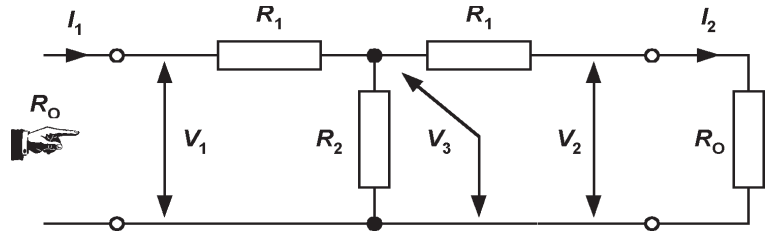


**Figure 4.2.12** Correctly terminated T-network attenuator

$$\text{and } R_1(1 + \alpha) = R_0(\alpha - 1)$$

$$\text{thus } R_1 = R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right).$$

$$\text{From Equation (i) on page 333, } R_0 = \sqrt{R_1^2 + 2R_1R_2}.$$

$$\text{Thus } R_0 = \sqrt{\left( R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right) \right)^2 + 2R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right) R_2}$$

$$\begin{aligned} \text{and } R_0^2 &= \left( R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right) \right)^2 + 2R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right) R_2 \\ &= R_0^2 \left( \frac{\alpha - 1}{\alpha + 1} \right)^2 + 2R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right) R_2 \end{aligned}$$

$$\text{thus } R_0^2 = R_0^2 \left( \frac{\alpha - 1}{\alpha + 1} \right)^2 + 2R_0R_2 \left( \frac{\alpha - 1}{\alpha + 1} \right).$$

Dividing both sides by  $R_0$  gives:

$$R_0 = R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right)^2 + 2R_2 \left( \frac{\alpha - 1}{\alpha + 1} \right)$$

$$\text{thus } 2R_2 \left( \frac{\alpha - 1}{\alpha + 1} \right) = R_0 - R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right)^2$$

$$\begin{aligned} \text{or } 2R_2 &= R_0 \left( \frac{\alpha + 1}{\alpha - 1} \right) - R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right) \left( \frac{\alpha + 1}{\alpha - 1} \right) \\ &= R_0 \left( \frac{\alpha + 1}{\alpha - 1} \right) - R_0 \left( \frac{\alpha - 1}{\alpha + 1} \right). \end{aligned}$$

$$\text{Hence } 2R_2 = R_0 \left( \left( \frac{\alpha + 1}{\alpha - 1} \right) - \left( \frac{\alpha - 1}{\alpha + 1} \right) \right)$$

$$\begin{aligned} \text{and } R_2 &= \frac{R_0}{2} \left( \left( \frac{\alpha + 1}{\alpha - 1} \right) - \left( \frac{\alpha - 1}{\alpha + 1} \right) \right) \\ &= \frac{R_0}{2} \left( \frac{(\alpha + 1)(\alpha + 1) - (\alpha - 1)(\alpha - 1)}{(\alpha + 1)(\alpha - 1)} \right) \end{aligned}$$

$$\begin{aligned}\text{thus } R_2 &= \frac{R_0}{2} \left( \frac{(\alpha^2 + 2\alpha + 1) - (\alpha^2 - 2\alpha + 1)}{(\alpha^2 - 1)} \right) \\ &= \frac{R_0}{2} \left( \frac{4\alpha}{\alpha^2 - 1} \right)\end{aligned}$$

$$\text{hence } R_2 = R_0 \left( \frac{2\alpha}{\alpha^2 - 1} \right).$$

### The $\pi$ -network attenuator

The circuit of a correctly terminated  $\pi$ -network attenuator is shown in Figure 4.2.13.

From Figure 4.2.13:

$$I_1 = V_1/R_0, I_1 = I_3 + I_4 \text{ and } I_3 = I_2 + I_5.$$

$$\text{Thus } I_1 = I_2 + I_4 + I_5$$

$$\text{and since } I_2 = V_2/R_0, I_4 = V_0/R_2, \text{ and } I_5 = V_2/R_2$$

$$\frac{V_1}{R_0} = \frac{V_2}{R_0} + \frac{V_1}{R_2} + \frac{V_2}{R_2}.$$

$$\text{As before, } \alpha = V_1/V_2.$$

$$\text{Thus } V_2 = V_1/\alpha.$$

$$\text{Hence } \frac{V_1}{R_0} = \frac{V_1}{\alpha R_0} + \frac{V_1}{R_2} + \frac{V_1}{\alpha R_2}$$

$$\text{or } \frac{1}{R_0} - \frac{1}{\alpha R_0} = \frac{1}{R_2} + \frac{1}{\alpha R_2}$$

$$\text{and } \frac{1}{R_0} \left( 1 - \frac{1}{\alpha} \right) = \frac{1}{R_2} \left( 1 + \frac{1}{\alpha} \right)$$

$$\text{thus } \frac{1}{R_0} \left( \frac{\alpha - 1}{\alpha} \right) = \frac{1}{R_2} \left( \frac{\alpha + 1}{\alpha} \right)$$

$$\text{from which } R_2 = R_0 \left( \frac{\alpha + 1}{\alpha - 1} \right). \quad (i)$$

Now  $I_1 = I_3 + I_4$  thus:

$$\frac{V_1}{R_0} = \frac{V_1 - V_2}{R_1} + \frac{V_1}{R_2}$$

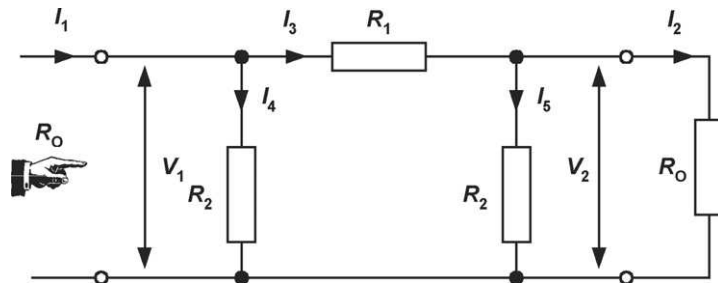


**Figure 4.2.13** Correctly terminated  $\pi$ -network attenuator

$$\text{or } \frac{V_1}{R_0} = \frac{V_1}{R_2} - \frac{V_2}{R_1} + \frac{V_1}{R_1}$$

$$\text{or } \frac{V_1}{R_0} = \frac{V_1}{R_2} - \frac{V_1}{\alpha R_1} + \frac{V_1}{R_1}$$

$$\text{or } \frac{1}{R_0} = \frac{1}{R_2} - \frac{1}{\alpha R_1} + \frac{1}{R_1}$$

$$\text{and } \frac{1}{R_0} - \frac{1}{R_2} = \frac{1}{R_1} \left( 1 - \frac{1}{\alpha} \right).$$

Substituting  $R_2$  from Equation (i) gives:

$$\frac{1}{R_0} - \frac{1}{R_0} \left( \frac{\alpha - 1}{\alpha + 1} \right) = \frac{1}{R_1} \left( \frac{\alpha - 1}{\alpha} \right)$$

$$\text{from which } \frac{1}{R_0} \left( 1 - \left( \frac{\alpha - 1}{\alpha + 1} \right) \right) = \frac{1}{R_1} \left( \frac{\alpha - 1}{\alpha} \right)$$

$$\text{and } \frac{1}{R_0} \left( \frac{(\alpha + 1) - (\alpha - 1)}{(\alpha + 1)} \right) = \frac{1}{R_1} \left( \frac{\alpha - 1}{\alpha} \right)$$

$$\text{or } \frac{1}{R_0} \left( \frac{2}{\alpha + 1} \right) = \frac{1}{R_1} \left( \frac{\alpha - 1}{\alpha} \right)$$

$$\text{Hence } R_1 = R_0 \left( \frac{\alpha - 1}{\alpha} \right) \left( \frac{\alpha + 1}{2} \right)$$

$$\text{or } R_1 = R_0 \left( \frac{\alpha^2 - 1}{2\alpha} \right).$$

### Propagation coefficient

The infinite series of symmetrical networks that we met on page 333 has another important characteristic. This is the relationship that exists between the input and output currents and voltages. Take a look at the arrangement shown in Figure 4.2.14.

Since each section is identical we can infer that the ratio of input to output current will be the same for each network section.

Thus  $I_0/I_1 = I_1/I_2 = I_2/I_3 = \text{a constant}$ .

The value of the constant will be determined by amount of attenuation and phase shift produced by each network section. It is convenient to

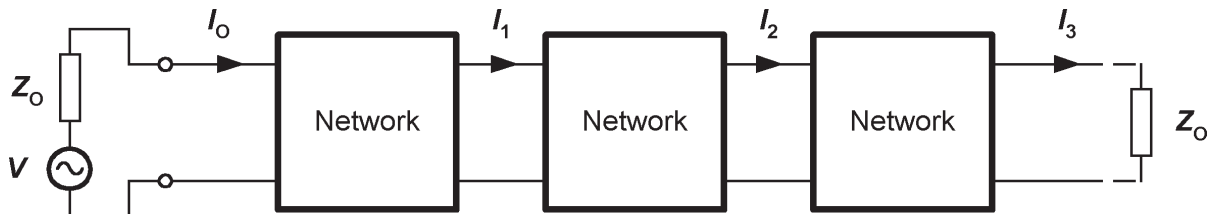


**Figure 4.2.14** Decreasing current in an infinite number of series-connected networks

express this as a complex quantity,  $e^\gamma$ , where  $\gamma$  is known as the *propagation coefficient*.

Thus, for a single section:

$$e^\gamma = I_0/I_1 = I_1/I_2 = I_2/I_3$$

hence for the  $n$ th section,  $e^\gamma = I_{n-1}/I_n$ .

For two, correctly terminated, sections the relationship between the output and input currents will be given by:

$$I_0/I_2 = (I_1/I_2) \times (I_0/I_1) = e^\gamma \times e^\gamma = e^{2\gamma}.$$

Similarly, for three, correctly terminated, sections the relationship between the output and input currents will be given by:

$$I_0/I_3 = (I_2/I_3) \times (I_1/I_2) \times (I_0/I_1) = e^\gamma \times e^\gamma \times e^\gamma = e^{3\gamma}.$$

Thus, for  $n$  sections correctly terminated we can infer that the current  $I_0/I_n = e^{n\gamma}$ .

Since  $\gamma$  is a complex number, we can infer that it has both real and imaginary parts. Thus:

$$\gamma = \alpha + j\beta$$

where  $\alpha$  is the *attenuation coefficient* and  $\beta$  is the *phase change coefficient*.  $\alpha$  is measured in *nepers* and  $\beta$  is measured in *radians*.

For a single network section,  $I_0/I_1 = e^\gamma$ .

Ignoring phase shift and, the real part of the current ratio (i.e. the *modulus*) will be given by:

$$|I_0/I_1| = e^\gamma.$$

Taking logs to the base  $e$  gives:

$$\ln |I_0/I_1| = \alpha.$$

The attenuation produced by  $n$  sections will thus be given by:

$$\ln |I_n/I_0| = n\alpha \text{ nepers.}$$

Finally, the phase shift produced by  $n$  sections will simply be given by  $n \times \beta$  rad.

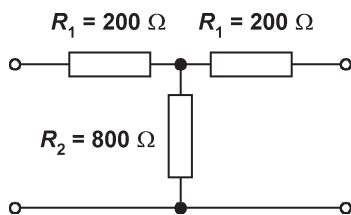


**Figure 4.2.15** See Example 4.2.3

### Example 4.2.3

Determine the characteristic impedance and attenuation provided by the T-network attenuator shown in Figure 4.2.15.

Comparing this with the T-network shown in Figure 4.2.3 reveals that:

$$Z_1 = R_1 = 200 \Omega \text{ and that } Z_2 = R_2 = 800 \Omega.$$

$$\text{For a T-network, } Z_0 = \sqrt{Z_1^2 + 2Z_1Z_2}.$$

$$\text{Thus } Z_0 = \sqrt{200^2 + (2 \times 200 \times 800)} = \sqrt{3600} = 600 \Omega.$$

$$\text{For a T-network attenuator, } \alpha = \frac{R_0 + R_1}{R_0 - R_1}.$$

$$\text{Thus } \alpha = \frac{600 + 200}{600 - 200} = \frac{800}{400} = 2.$$

### Relationship between nepers and decibels

In Chapter 3 we introduced the use of decibels (dB) for specifying voltage, current and power ratios. You should recall that decibels are calculated using logarithms to a base of 10. Nepers, on the other hand, are determined using the natural logarithm of a ratio of either voltage, current or power.

If the resistive components of the impedances at the input and output of the network are equal (this would be the case for a symmetrical network), attenuation can be easily converted from nepers to decibels and vice versa.

For a given current ratio,  $I_0/I_1$ , the attenuation,  $m$ , expressed using decibels would be given by:

$$m = 20 \log_{10}(I_0/I_1).$$

Using standard laws of logarithms gives:

$$m = 20 \log_e(I_0/I_1) \times \log_{10} e$$

$$m = 20 \times \log_{10} e \times \log_e(I_0/I_1)$$

$$m = 20 \times 0.4343 \times \log_e(I_0/I_1) = 8.686 \times \log_e(I_0/I_1).$$

But  $\log_e(I_0/I_1)$  is the attenuation expressed in nepers. Thus, we can infer the following rules:

- to convert from nepers to decibels, multiply the value in nepers by 8.686;
- to convert from decibels to nepers, multiply the value in decibels by 0.1151.

## Mathematics in action

### Changing the base of logarithms

The change of base formula can be proved using the following simple steps:

If  $y = \log_a(x)$  by definition,  $x = a^y$

Now  $\log_b(x) = \log_b(a^y) = y \log_b(a)$ .

Thus  $\log_b(x) = \log_a(x) \times \log_b(a)$ .

### Questions 4.2.1

- (1) Determine the characteristic impedance of each of the networks shown in Figure 4.2.16.
- (2) Design a T-network attenuator that will have a characteristic impedance of  $50 \Omega$  and an attenuation of 10.
- (3) Design a  $\pi$ -network attenuator that will have a characteristic impedance of  $600 \Omega$  and an attenuation of 13 dB.



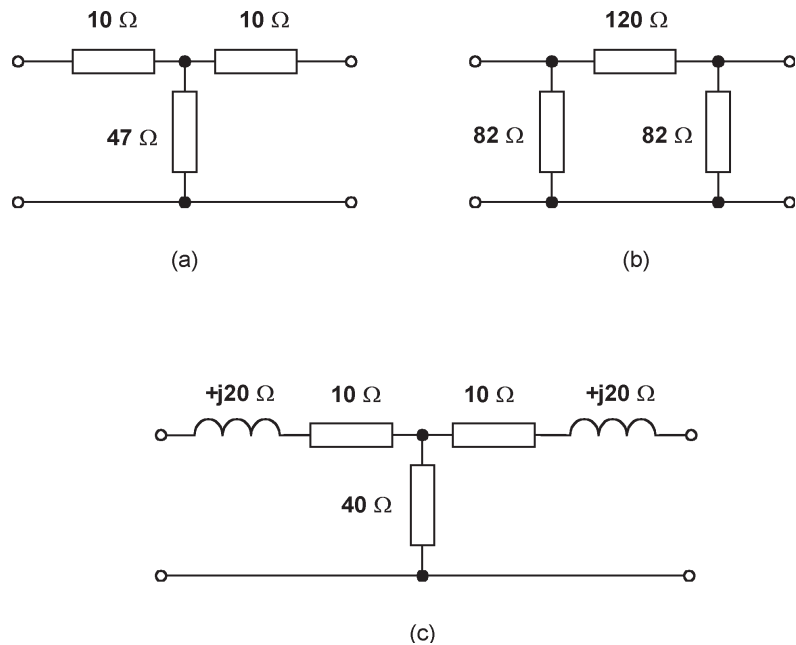


**Figure 4.2.16** See Questions 4.2.1

- (4) Three identical symmetrical two-port networks provide a total 270 Electrical and electronic principles attenuation of 7.5 and a phase change of  $+5^\circ$ . Determine the propagation coefficient for a single two-port network.
- (5) The input and output voltage of a correctly terminated attenuator are 1.8 V and 225 mV, respectively. Determine the output current supplied to a matched  $75\ \Omega$  load and the attenuation expressed in both decibels and nepers.

## 4.3 COMPLEX WAVES

In Chapter 3 we introduced complex waveforms and showed how a complex wave could be expressed by an equation of the form:

$$v = V_1 \sin(\omega t) + V_2 \sin(2\omega t \pm \phi_2) + V_3 \sin(3\omega t \pm \phi_3) + V_4 \sin(4\omega t \pm \phi_4) + \dots \quad (\text{i})$$

In this section we shall develop this idea further by introducing a powerful technique called Fourier analysis.

### Fourier analysis

Fourier analysis is based on the concept that all waveforms, whether continuous or discontinuous, can be expressed in terms of a convergent series of the form:

$$v = A_0 + A_1 \cos t + A_2 \cos 2t + A_3 \cos 3t + \dots + B_1 \sin t + B_2 \sin 2t + B_3 \sin 3t + \dots, \quad (\text{ii})$$

where  $A_0, A_1, \dots, B_1, B_2, \dots$  are the amplitudes of the individual components that make up the complex waveform. Essentially, this formula is the same as the simplified relationship in Equation (i) but with the introduction of cosine as well as sine components. Note also that we have made the assumption that one complete cycle occurs in a time equal to  $2\pi$  s (recall that  $\omega = 2\pi f$  and one complete cycle requires  $2\pi$  rad).

## Another view

If you are puzzled by Equations (iv)–(vi), it may help to put into words what the values of the Fourier coefficients actually represent:

- $A_0$  is the mean value of  $v$  over one complete cycle of the waveform (i.e. from  $-\pi$  to  $+\pi$  or from 0 to  $2\pi$ ).
- $A_1$  is twice the mean value of  $v \cos t$  over one complete cycle.
- $B_1$  is twice the mean value of  $v \sin t$  over one complete cycle.
- $A_2$  is twice the mean value of  $v \cos 2t$  over one complete cycle.
- $B_2$  is twice the mean value of  $v \sin 2t$  over one complete cycle, and so on. We shall use this fact later when we describe an alternative method of determining the Fourier series for a particular waveform.

Another way of writing Equation (ii) is:

$$v = A_0 + \sum_{n=1}^{n=\infty} (A_n \sin nt + B_n \cos nt). \quad (\text{iii})$$

For the range  $-\pi$  to  $+\pi$ , the value of the constant terms (the amplitudes of the individual components),  $A_0, A_1, \dots B_1, B_2, \dots$ , etc. can be determined as follows:

$$A_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} v \, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( A_0 + \sum_{n=1}^{n=\infty} (A_n \cos nt + B_n \sin nt) \right) dt \quad (\text{iv})$$

$$A_n = \frac{1}{\pi} \int_{-\pi}^{\pi} v \cos nt \, dt \quad \text{where } n = 1, 2, 3, \dots \quad (\text{v})$$

$$B_n = \frac{1}{\pi} \int_{-\pi}^{\pi} v \sin nt \, dt \quad \text{where } n = 1, 2, 3, \dots \quad (\text{vi})$$

The values,  $A_0, A_1, \dots B_1, B_2, \dots$ , etc. are called the *Fourier coefficients* of the *Fourier series* defined by Equation (iii).

Just in case this is all beginning to sound a little too complex in terms of mathematics, let us take a look at some examples based on waveforms that you should immediately recognise:

### Example 4.3.1

Determine the Fourier coefficients of the sine wave superimposed on the constant DC level shown in Figure 4.3.1.

By inspection, the equation of the wave (over the range 0– $2\pi$ ) is:

$$v = 10 + 5 \sin t.$$

By comparison with Equation (ii), we can deduce that:

$A_0 = 10$  (in other words, the mean voltage over the range 0– $2\pi$  is 10 V).

$B_1 = 5$  (this is simply the coefficient of  $\sin t$ ).

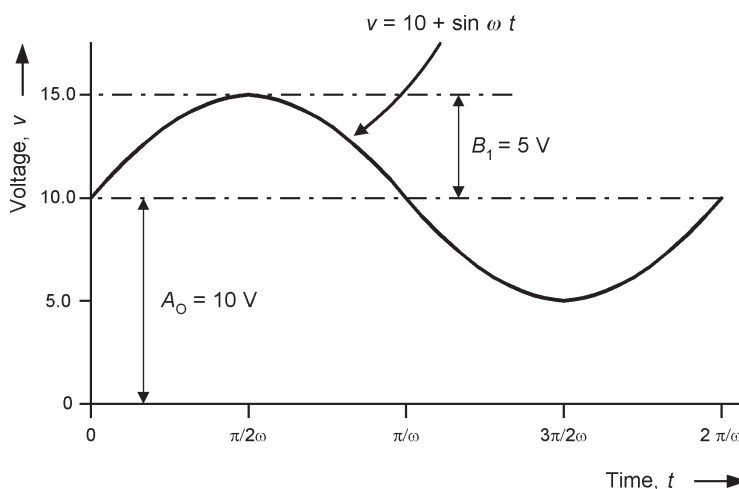


**Figure 4.3.1** See Example 4.3.1

There are no other terms and thus all other Fourier coefficients ( $A_1$ ,  $A_2$ ,  $B_2$ , etc.) are all zero.

Now let us see if we can arrive at the same results using Equations (iv)–(vi):

The value of  $A_0$  is found from:

$$A_0 = \frac{1}{2\pi} \int_0^{2\pi} v \, dt = \frac{1}{2\pi} \int_0^{2\pi} (10 + 5 \sin t) \, dt$$

$$A_0 = \frac{1}{2\pi} [10t]_{t=0}^{t=2\pi} + \frac{1}{2\pi} \int_0^{2\pi} 5 \sin t \, dt$$

$$A_0 = \left[ \frac{20\pi}{2\pi} - 0 \right] + 0 = 10.$$

The value of  $A_1$  is found from:

$$A_1 = \frac{1}{\pi} \int_0^{2\pi} v \cos t \, dt = \frac{1}{\pi} \int_0^{2\pi} (10 + 5 \sin t) \cos t \, dt$$

$$A_1 = \frac{1}{\pi} \int_0^{2\pi} (10 \cos t + 5 \sin t \cos t) \, dt$$

$$A_1 = \frac{1}{\pi} \int_0^{2\pi} 10 \cos t \, dt + \frac{1}{\pi} \int_0^{2\pi} 5 \sin t \cos t \, dt$$

$$A_1 = \frac{1}{\pi} \int_0^{2\pi} 10 \cos t \, dt + 0$$

$$A_1 = \frac{1}{\pi} [10 \sin t]_{t=0}^{t=2\pi} + 0 = 0.$$

The value of  $B_1$  is found from:

$$B_1 = \frac{1}{\pi} \int_0^{2\pi} v \sin t \, dt = \frac{1}{\pi} \int_0^{2\pi} (10 + 5 \sin t) \sin t \, dt$$

$$B_1 = \frac{1}{\pi} \int_0^{2\pi} (10 \sin t + 5 \sin^2 t) \, dt$$

$$B_1 = \frac{1}{\pi} \int_0^{2\pi} 10 \sin t \, dt + \frac{1}{\pi} \int_0^{2\pi} 5 \sin^2 t \, dt$$

$$B_1 = 0 + \frac{1}{\pi} \left[ \frac{5}{2} \left( t - \frac{1}{2} \sin 2t \right) \right]_{t=0}^{t=2\pi}$$

$$B_1 = 0 + \frac{10\pi}{2\pi} = 5.$$

**Example 4.3.2**

Determine the Fourier series for the rectangular pulse shown in Figure 4.3.2.

Unlike the waveform in the previous example, this voltage is discontinuous and we must therefore deal with it in two parts; that from 0 to  $1.5\pi$  and that from  $1.5\pi$  to  $2\pi$ .

The equation of the voltage is as follows:

$$v = 2 \text{ (over the range } 0 < t < 1.5\pi)$$

$$v = 0 \text{ (over the range } 1.5\pi < t < 2\pi).$$

To find  $A_0$  we simply need to find the mean value of the waveform over the range 0– $2\pi$ . By considering the area under the waveform it should be obvious that this is 1.5.

Hence  $A_0 = 1.5$ .

Next we need to find the value of  $A_n$  for  $n = 1, 2, 3, \dots$ , and so on. From Equation (v):

$$A_n = \frac{1}{\pi} \int_0^{2\pi} v \cos nt \, dt.$$

But since there are two distinct parts to the waveform, we need to integrate separately over the two time periods, 0– $1.5\pi$  and  $1.5\pi$ – $2\pi$ :

$$A_n = \frac{1}{\pi} \int_0^{1.5\pi} 2 \cos nt \, dt + \frac{1}{\pi} \int_{1.5\pi}^{2\pi} 0 \cos nt \, dt$$

$$A_n = \frac{1}{\pi} \int_0^{1.5\pi} 2 \cos nt \, dt + 0$$

$$A_n = \frac{2}{n\pi} [\sin nt]_{t=0}^{t=1.5\pi} + 0 = \frac{2}{n\pi} [\sin 1.5n\pi - \sin 0].$$

$$\text{Thus } A_n = \frac{2}{n\pi} \sin 1.5n\pi.$$

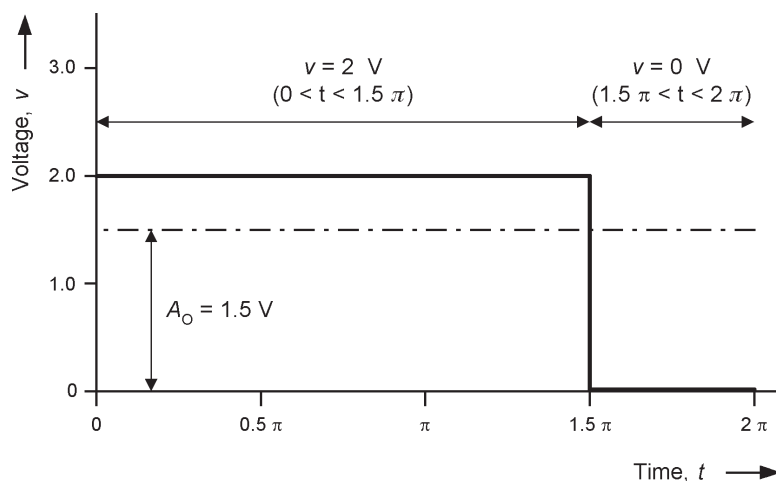


**Figure 4.3.2** See Example 4.3.2

Now, when  $n = 1$ ,  $A_1 = \frac{2}{\pi} \sin 1.5\pi = -\frac{2}{\pi}$

when  $n = 2$ ,  $A_2 = \frac{1}{\pi} \sin 3\pi = 0$

when  $n = 3$ ,  $A_3 = \frac{2}{3\pi} \sin 4.5\pi = +\frac{2}{3\pi}$

when  $n = 4$ ,  $A_4 = \frac{1}{2\pi} \sin 6\pi = 0$

when  $n = 5$ ,  $A_5 = \frac{2}{5\pi} \sin 7.5\pi = -\frac{2}{5\pi}$ .

The cosine terms in the Fourier series will thus be:

$$-\frac{2}{\pi} \cos t + \frac{2}{3\pi} \cos 3t - \frac{2}{5\pi} \cos 5t \dots$$

The sine terms can be similarly found:

$$B_n = \frac{1}{\pi} \int_0^{1.5\pi} 2 \sin nt \, dt + \frac{1}{\pi} \int_{1.5\pi}^{2\pi} 0 \sin nt \, dt$$

$$B_n = \frac{1}{\pi} \int_0^{1.5\pi} 2 \sin nt \, dt + 0$$

$$B_n = \frac{-2}{n\pi} [\cos nt]_{t=0}^{t=1.5\pi} + 0 = \frac{-2}{n\pi} [\cos 1.5n\pi - \cos 0].$$

$$\text{Thus } B_n = \frac{-2}{n\pi} (\cos 1.5n\pi - 1).$$

Now, when  $n = 1$ ,  $B_1 = \frac{-2}{\pi} (\cos 1.5\pi - 1) = \frac{2}{\pi}$

when  $n = 2$ ,  $B_2 = \frac{-2}{2\pi} (\cos 3\pi - 1) = \frac{2}{\pi}$

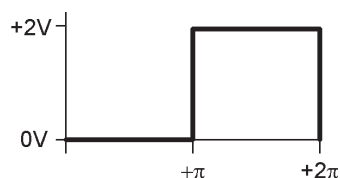
when  $n = 3$ ,  $B_3 = \frac{-2}{3\pi} (\cos 4.5\pi - 1) = \frac{2}{3\pi}$

when  $n = 4$ ,  $B_4 = \frac{-2}{4\pi} (\cos 6\pi - 1) = 0$

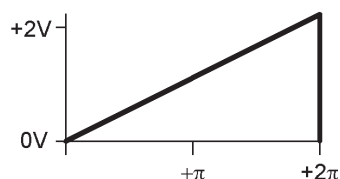
when  $n = 5$ ,  $B_5 = \frac{-2}{5\pi} (\cos 7.5\pi - 1) = \frac{2}{5\pi}$ .

Thus, the sine terms in the Fourier series will be:

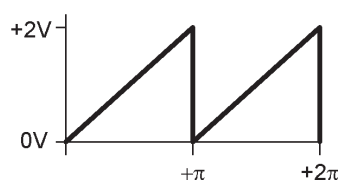
$$\frac{2}{\pi} \sin t + \frac{2}{\pi} \sin 2t + \frac{2}{3\pi} \sin 3t + \frac{2}{5\pi} \sin 5t \dots$$



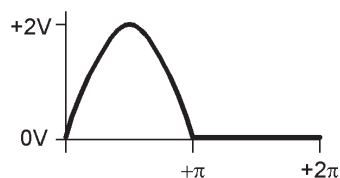
(a)



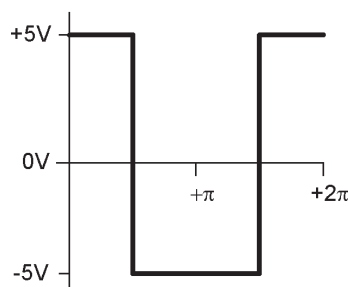
(b)



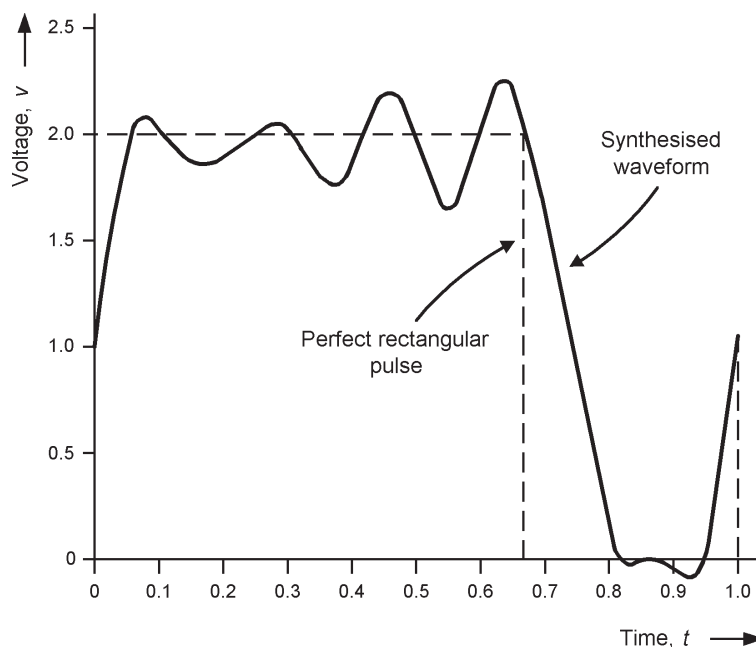
(c)



(d)



(e)

**Figure 4.3.4** See Question 4.3.1**Figure 4.3.3** A pulse waveform synthesised from harmonic components up to the ninth harmonic

We are now in a position to develop the expression for the pulse. This will be similar to that shown in Equation (ii) but with:

$$A_0 = 1.5, A_1 = -\frac{2}{\pi}, A_2 = 0, A_3 = \frac{2}{3\pi}, \text{ and so on.}$$

The expression for the pulse is thus:

$$v = 1.5 - \frac{2}{\pi} \cos t + \frac{2}{3\pi} \cos 3t - \frac{2}{5\pi} \cos 5t \dots \\ + \frac{2}{\pi} \sin t + \frac{2}{\pi} \sin 2t + \frac{2}{3\pi} \sin 3t + \frac{2}{5\pi} \sin 5t \dots$$

Using this function in a standard spreadsheet program (up to and including the ninth harmonic terms) produces the *synthesised* pulse shown in Figure 4.3.3.

### Question 4.3.1

Derive the Fourier series for each of the waveforms shown in Figure 4.3.4.

### An alternative method

An alternative tabular method can be used to determine Fourier coefficients. This method is based on an application of the trapezoidal rule and it is particularly useful when a number of voltage readings are available at regular time intervals over the period of one complete cycle of a

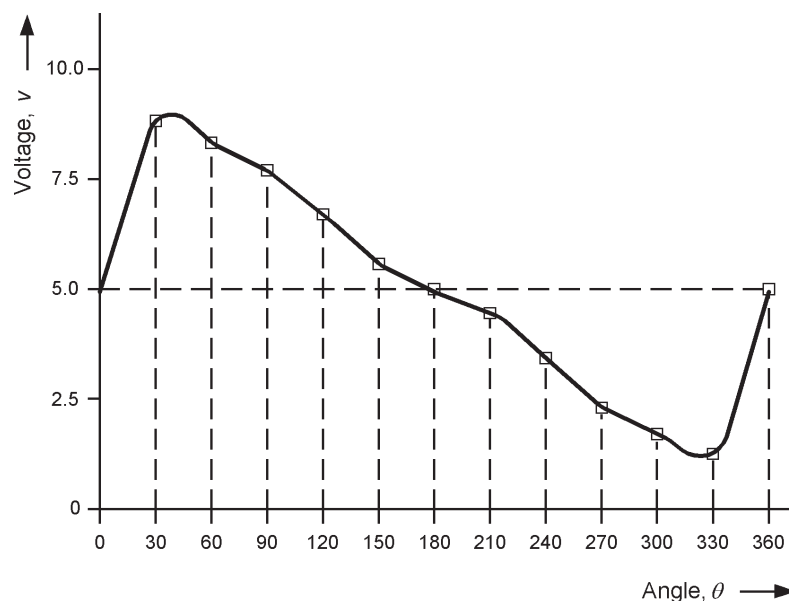


**Figure 4.3.5** See Example 4.3.3

waveform. Once again, we shall show how this method works by using an example:

### Example 4.3.3

The values of voltage over a complete cycle of a waveform (see Figure 4.3.5) are as follows:

|                 |     |     |     |     |     |     |     |     |     |     |     |     |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Angle, $\theta$ | 30  | 60  | 90  | 120 | 150 | 180 | 210 | 240 | 270 | 300 | 330 | 360 |
| Voltage, $v$    | 8.9 | 8.1 | 7.5 | 6.5 | 5.5 | 5.0 | 4.5 | 3.5 | 2.5 | 1.9 | 1.1 | 5   |

First we need to determine the value of  $A_0$ . We can do this by calculating the mean value of voltage over the complete cycle. Applying the trapezoidal rule (i.e. adding together the values of voltage and dividing by the number of intervals) gives:

$$A_0 = (8.9 + 8.1 + 7.5 + 6.5 + \dots + 5)/12 = 5.$$

Similarly, to determine the value of  $A_1$  we multiply each value of voltage by  $\cos \theta$ , before adding them together and dividing by half the number of intervals (recall that  $A_1$  is twice the mean value of  $v \cos t$  over one complete cycle). Thus:

$$A_1 = (7.707 + 4.049 - 0.002 - 3.252 + \dots + 5)/6 = 0.$$

The same method can be used to determine the remaining Fourier coefficients. The use of a spreadsheet for this calculation is highly recommended (see Figure 4.3.6).

Having determined the Fourier coefficients, we can write down the expression for the complex waveform. In this case it is:

$$v = 5 + 2.894 \sin \theta + 1.444 \sin 2\theta + 0.635 \sin 3\theta + 0.522 \sin 4\theta + \dots$$

(note that there are no terms in  $\cos \theta$  in the Fourier series for this waveform).

## Fourier analysis - tabular method

| $y$      | $\theta$ | $v$ | $\cos \theta$ | $v \cos \theta$ | $\sin \theta$ | $v \sin \theta$ | $\cos 2\theta$ | $v \cos 2\theta$ | $\sin 2\theta$ | $v \sin 2\theta$ | $\cos 3\theta$ | $v \cos 3\theta$ | $\sin 3\theta$ | $v \sin 3\theta$ | $\cos 4\theta$ | $v \cos 4\theta$ | $\sin 4\theta$ | $v \sin 4\theta$ |
|----------|----------|-----|---------------|-----------------|---------------|-----------------|----------------|------------------|----------------|------------------|----------------|------------------|----------------|------------------|----------------|------------------|----------------|------------------|
| $y_1$    | 30       | 8.9 | 0.866         | 7.707           | 0.500         | 4.451           | 0.500          | 4.449            | 0.866          | 7.708            | 0.000          | -0.002           | 1.000          | 8.900            | -0.500         | -4.452           | 0.866          | 7.706            |
| $y_2$    | 60       | 8.1 | 0.500         | 4.049           | 0.866         | 7.015           | -0.500         | -4.052           | 0.866          | 7.014            | -1.000         | -8.100           | 0.000          | -0.003           | -0.500         | -4.046           | -0.866         | -7.017           |
| $y_3$    | 90       | 7.5 | 0.000         | -0.002          | 1.000         | 7.500           | -1.000         | -7.500           | 0.000          | -0.003           | 0.001          | 0.005            | -1.000         | -7.500           | 1.000          | 7.500            | 0.001          | 0.006            |
| $y_4$    | 120      | 6.5 | -0.500        | -3.252          | 0.866         | 5.628           | -0.500         | -3.247           | -0.866         | -5.631           | 1.000          | 6.500            | 0.001          | 0.005            | -0.501         | -3.256           | 0.865          | 5.626            |
| $y_5$    | 150      | 5.5 | -0.866        | -4.764          | 0.500         | 2.748           | 0.501          | 2.753            | -0.866         | -4.761           | -0.001         | -0.006           | 1.000          | 5.500            | -0.499         | -2.744           | -0.867         | -4.767           |
| $y_6$    | 180      | 5   | -1.000        | -5.000          | 0.000         | -0.002          | 1.000          | 5.000            | 0.001          | 0.004            | -1.000         | -5.000           | -0.001         | -0.006           | 1.000          | 5.000            | 0.002          | 0.008            |
| $y_7$    | 210      | 4.5 | -0.866        | -3.896          | -0.500        | -2.252          | 0.499          | 2.246            | 0.867          | 3.899            | 0.001          | 0.006            | -1.000         | -4.500           | -0.502         | -2.257           | 0.865          | 3.893            |
| $y_8$    | 240      | 3.5 | -0.500        | -1.748          | -0.866        | -3.032          | -0.501         | -1.753           | 0.865          | 3.029            | 1.000          | 3.500            | 0.002          | 0.006            | -0.498         | -1.743           | -0.867         | -3.035           |
| $y_9$    | 270      | 2.5 | 0.001         | 0.002           | -1.000        | -2.500          | -1.000         | -2.500           | -0.001         | -0.003           | -0.002         | -0.005           | 1.000          | 2.500            | 1.000          | 2.500            | 0.002          | 0.006            |
| $y_{10}$ | 300      | 1.9 | 0.501         | 0.951           | -0.866        | -1.645          | -0.499         | -0.948           | -0.867         | -1.647           | -1.000         | -1.900           | -0.002         | -0.004           | -0.502         | -0.954           | 0.865          | 1.643            |
| $y_{11}$ | 330      | 1.1 | 0.866         | 0.953           | -0.499        | -0.549          | 0.501          | 0.551            | -0.865         | -0.952           | 0.002          | 0.002            | -1.000         | -1.100           | -0.497         | -0.547           | -0.868         | -0.954           |
| $y_{12}$ | 360      | 5   | 1.000         | 5.000           | 0.001         | 0.004           | 1.000          | 5.000            | 0.002          | 0.008            | 1.000          | 5.000            | 0.002          | 0.012            | 1.000          | 5.000            | 0.003          | 0.016            |
| Sum      | 60       |     | Sum           | 0.001           | Sum           | 17.367          | Sum            | 0.000            | Sum            | 8.666            | Sum            | 0.001            | Sum            | 3.810            | Sum            | 0.000            | Sum            | 3.131            |
| $A_0 =$  | 5        |     | $A_1 =$       | 0.000           | $B_1 =$       | 2.894           | $A_2 =$        | 0.000            | $B_2 =$        | 1.444            | $A_3 =$        | 0.000            | $B_3 =$        | 0.635            | $A_4 =$        | 0.000            | $B_4 =$        | 0.522            |

**Figure 4.3.6** Spreadsheet table of Fourier coefficients

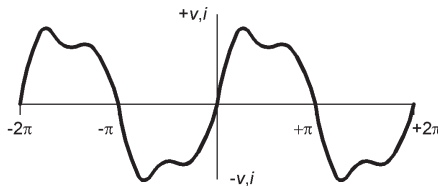
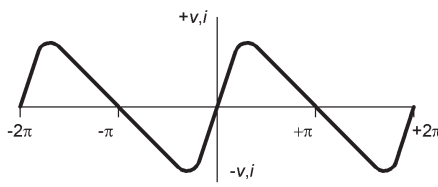
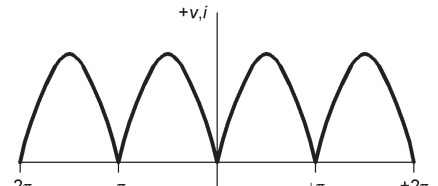
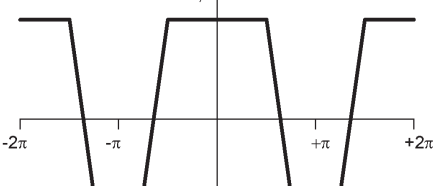
| Fourier coefficient                                                                                                                        | $A_0$       | $A_n$               | $B_n$               |
|--------------------------------------------------------------------------------------------------------------------------------------------|-------------|---------------------|---------------------|
| Complex waveform                                                                                                                           | (DC offset) | (cosine terms)      | (sine terms)        |
|  <p>Alternate half-cycles inverted</p>                  | No          | Yes (odd $n$ only)  | Yes (odd $n$ only)  |
|  <p>Alternate half-cycles reflected</p>                | No          | No                  | Yes                 |
|  <p>Half-cycles repeated</p>                           | Yes         | Yes (even $n$ only) | Yes (even $n$ only) |
|  <p>Wave symmetrical about the <math>y</math>-axis</p> | Yes         | Yes                 | No                  |

**Figure 4.3.7** Effect of symmetry and reflection on Fourier coefficients for different types of waveform



## Mathematics in action

### *The trapezoidal rule*

In order to determine the mean value of  $y = f(t)$  over a given period, we can divide the period into a number of smaller intervals and determine the  $y$  ordinate value for each.

If the  $y$  values are  $y_1, y_2, y_3, \dots, y_m$ , and there are  $m$  equal values in a time interval,  $t$ , the mean value of  $y$ ,  $\bar{y}$ , will be given by:

$$\bar{y} = \frac{y_1 + y_2 + y_3 + \dots + y_m}{m} = \frac{\sum_{n=0}^{n=m} y_n}{m}.$$

### Question 4.3.2

The following voltages are taken over one complete cycle of a complex waveform:

|                 |   |     |     |     |     |     |     |     |     |     |      |      |     |
|-----------------|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|-----|
| Angle, $\theta$ | 0 | 30  | 60  | 90  | 120 | 150 | 180 | 210 | 240 | 270 | 300  | 330  | 360 |
| Voltage, $v$    | 0 | 3.5 | 8.7 | 9.1 | 9.5 | 9.5 | 9.1 | 8.7 | 3.5 | 0   | -3.5 | -3.5 | 0   |

Use a tabular method to determine the Fourier series for the waveform.

In Chapter 3 we discussed the effect of adding harmonics to a fundamental on the shape of the complex wave produced. Revisiting this in the light of what we now know about Fourier series allows us to relate wave-shape to the corresponding Fourier series. Figure 4.3.7 shows the effect symmetry (and reflected symmetry) has on the values of  $A_0, A_1, A_2, \dots, B_1, B_2$ , etc. This information can allow us to take a few short-cuts when analysing a waveform (there is no point trying to evaluate Fourier coefficients that are just not present!).

## The root mean square value of a waveform

The root mean square (rms) value of a waveform is the effective value of the current or voltage concerned. It is defined as the value of DC or voltage that would produce the same power in a pure resistive load.

Assume that we are dealing with a voltage given by:

$$v = V \cos \omega t.$$

If this voltage is applied to a pure resistance,  $R$ , the current flowing will be given by:

$$i = \frac{V}{R} \cos \omega t.$$

The instantaneous power dissipated in the resistor,  $p$ , will be given by the product of  $i$  and  $v$ . Thus:

$$p = iv = \frac{V}{R} \cos \omega t \times V \cos \omega t = \frac{V^2}{R} \cos^2 \omega t = \frac{V^2}{R} (0.5 + 0.5 \cos 2\omega t).$$

## Mathematics in action

### Trigonometric identities

In the previous text, we used a trigonometric identity to replace the  $\cos^2 \omega$  term with something that is easier to cope with. The reasoning is as follows:

$$\cos(A + B) = \cos A \cos B - \sin A \sin B.$$

$$\text{Thus } \cos 2A = \cos^2 A - \sin^2 A.$$

$$\text{But } \sin^2 A + \cos^2 A = 1 \text{ thus } \sin^2 A = 1 - \cos^2 A.$$

$$\text{Thus } \cos 2A = \cos^2 A - (1 - \cos^2 A) = 2 \cos^2 A - 1.$$

$$\text{Hence } \cos^2 A = 0.5(1 + \cos 2A) = 0.5 + 0.5 \cos 2A.$$

$$\text{Hence, } p = 0.5 \frac{V^2}{R} + 0.5 \frac{V^2}{R} \cos 2\omega t$$

$$\text{or } p = 0.5 \frac{V^2}{R} + 0.5 \frac{V^2}{R} \cos 2\omega t = \frac{V^2}{2R} + \frac{V^2}{2R} \cos 2\omega t. \quad (\text{i})$$

This is an important result. From Equation (i), we can infer that the power waveform (i.e. the waveform of  $p$  plotted against  $t$ ) will comprise a cosine waveform at *twice* the frequency of the voltage (see Figure 4.3.8). Furthermore, if we apply our recently acquired knowledge of Fourier series, we can infer that the mean value of the power waveform (over a complete cycle of the voltage or current) will be the same as its amplitude. The value of the Fourier coefficients being:

$$A_0 = \frac{V^2}{2R} \quad (\text{i.e. the mean value of the power waveform over one } 2R \text{ cycle of voltage})$$

$$A_2 = \frac{V^2}{2R} \quad (\text{i.e. the amplitude of the term in } \cos 2t)$$

(note that no other components are present).

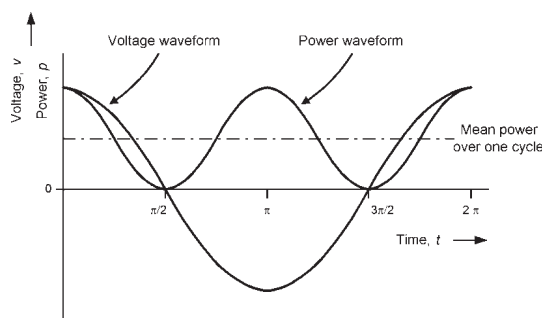


**Figure 4.3.8** Voltage and corresponding power waveform

The mean power, over one cycle, is thus given by  $A_0$ .

$$\text{Thus } P = \frac{V^2}{2R}.$$

$$\text{and } V = \sqrt{2PR}.$$

Now, let  $V_{\text{RMS}}$  be the equivalent DC voltage that will produce the same power.

$$\text{Thus } P = \frac{V_{\text{RMS}}^2}{R}$$

$$\text{and } V_{\text{RMS}} = \sqrt{PR}.$$

$$\text{Hence } V = \sqrt{2V_{\text{RMS}}}.$$

In other words, a waveform given by  $v = 1.414 \cos \omega t$  V will produce the same power in a load as a direct voltage of 1 V. A similar relationship can be obtained for current:

$$I = \sqrt{2I_{\text{RMS}}}$$

where  $I$  is the amplitude of the current wave.

When a wave is complex, it follows that its rms value can be found by adding together the values of each individual harmonic component present. Thus:

$$V_{\text{RMS}} = \sqrt{V_0^2 + \frac{V_1^2 + V_2^2 + V_3^2 + \dots + V_m^2}{2}} = \sqrt{V_0^2 + \frac{\sum_{n=1}^{n=m} V_n^2}{2}}$$

where  $V_0$  is the value of any DC component that may be present, and  $V_1, V_2, V_3, \dots, V_m$  are the amplitudes of the harmonic components present in the wave.

### Power factor

When harmonic components are present, the power factor can be determined by adding together the power supplied by each harmonic before dividing by the product of rms voltage and current. Hence:

$$\text{Power factor} = \frac{\text{total power supplied}}{\text{true power}} = \frac{P_{\text{TOT}}}{V_{\text{RMS}} I_{\text{RMS}}}$$

where  $P_{\text{TOT}}$  is found by adding the power due to each harmonic component present. Thus:

$$\begin{aligned} \text{Power factor} &= \frac{\frac{V_1}{\sqrt{2}} \frac{I_1}{\sqrt{2}} \cos \phi_1 + \frac{V_2}{\sqrt{2}} \frac{I_2}{\sqrt{2}} \cos \phi_2 + \frac{V_3}{\sqrt{2}} \frac{I_3}{\sqrt{2}} \cos \phi_3 + \dots + \frac{V_n}{\sqrt{2}} \frac{I_n}{\sqrt{2}} \cos \phi_n}{V_{\text{RMS}} I_{\text{RMS}}} \end{aligned}$$

Power factor

$$= \frac{\frac{V_1 I_1}{2} \cos \phi_1 + \frac{V_2 I_2}{2} \cos \phi_2 + \frac{V_3 I_3}{2} \cos \phi_3 + \dots + \frac{V_n I_n}{2} \cos \phi_n}{V_{\text{RMS}} I_{\text{RMS}}}.$$

$$\text{Power factor} = \frac{\sum_{n=1}^{n=m} \frac{V_n I_n \cos \phi_n}{2}}{V_{\text{RMS}} I_{\text{RMS}}}.$$

### Example 4.3.4

Determine the effective voltage and average power in a  $10 \Omega$  resistor when the following voltage is applied to it:

$$v = 10 \sin \omega t + 5 \sin 2\omega t + 2 \sin 3\omega t.$$

$$\text{Now } V_{\text{RMS}} = \sqrt{V_0^2 + \frac{V_1^2 + V_2^2 + V_3^2 + \dots + V_n^2}{2}}.$$

In this case,  $V_0 = 0$  (there is no DC component),  $V_1 = 10$ ,  $V_2 = 5$ , and  $V_3 = 2$ .

$$\text{Thus } V_{\text{RMS}} = \sqrt{\frac{10^2 + 5^2 + 2^2}{2}} = \sqrt{\frac{129}{2}} = 8.03 \text{ V}$$

$$P = \frac{8.03^2}{10} = 6.45 \text{ W}.$$

### Example 4.3.5

The voltage and current present in an AC circuit is as follows:

$$v = 50 \sin \omega t + 10 \sin(3\omega t + \pi/2)$$

$$i = 3.54 \sin(\omega t + \pi/4) + 0.316 \sin(3\omega t + 0.321).$$

Determine the total power supplied and the power factor. The total power supplied,  $P_{\text{TOT}}$ , is given by:

$$P_{\text{TOT}} = \frac{50 \times 3.54}{2} \cos(-\pi/4) + \frac{10 \times 0.316}{2} \cos(\pi/2 - 0.321)$$

$$P_{\text{TOT}} = (88.5 \times 0.707) + (1.58 \times 0.32) = 62.6 + 0.51 = 63.11 \text{ W}.$$

The rms voltage,  $V_{\text{RMS}}$ , is given by:

$$V_{\text{RMS}} = \sqrt{\frac{50^2 + 10^2}{2}} = \sqrt{\frac{2600}{2}} = 36.1 \text{ V}.$$

The rms current,  $I_{\text{RMS}}$ , is given by:

$$I_{\text{RMS}} = \sqrt{\frac{3.54^2 + 0.316^2}{2}} = \sqrt{\frac{12.63}{2}} = 2.51 \text{ A.}$$

The true power,  $P$ , is thus:

$$P = V_{\text{RMS}} I_{\text{RMS}} = 36.1 \times 2.51 = 90.6 \text{ W.}$$

Finally, to find the power factor we simply divide the total power,  $P_{\text{TOT}}$ , by the true power,  $P$ :

$$\text{Power factor} = \frac{P_{\text{TOT}}}{P} = \frac{63.11}{90.6} = 0.7.$$

### Questions 4.3.3

- (1) A voltage is given by the expression:

$$v = 100 \sin \omega t + 40 \sin(2\omega t - \pi/2) + 20 \sin(4\omega t - \pi/2).$$

If this voltage is applied to a  $50 \Omega$  resistor:

- determine the power in the resistor due to the fundamental and each harmonic component;
  - derive an expression for the current flowing in the resistor;
  - calculate the rms voltage and current;
  - determine the total power dissipated.
- (2) The voltage,  $v$ , and current,  $i$ , in a circuit are defined by the following expressions:

$$v = 30 \sin \omega t + 7.5 \sin(3\omega t + \pi/4)$$

$$i = \sin \omega t + 0.15 \sin(3\omega t - \pi/12).$$

Determine:

- the rms voltage;
  - the rms current;
  - the total power;
  - the true power;
  - the power factor of the circuit.
- (3) A voltage given by the expression:
- $$v = 100 \sin 314t + 50 \sin 942t - 40 \sin 1570t$$
- is applied to an impedance of  $(40 + j30) \Omega$ . Determine:
- an expression for the instantaneous current flowing;
  - the effective (rms) voltage;
  - the effective (rms) current;
  - the total power supplied;
  - the true power;
  - the power factor.

## 4.4 TRANSIENTS IN R-L-C CIRCUITS

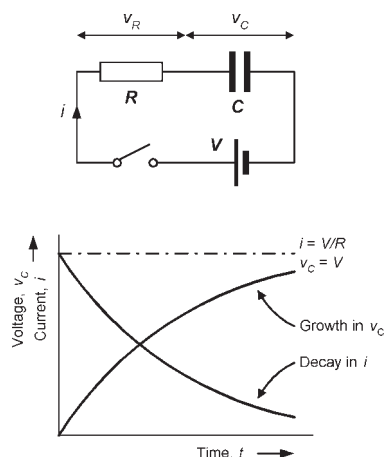


**Figure 4.4.1** Time response of a series C-R network

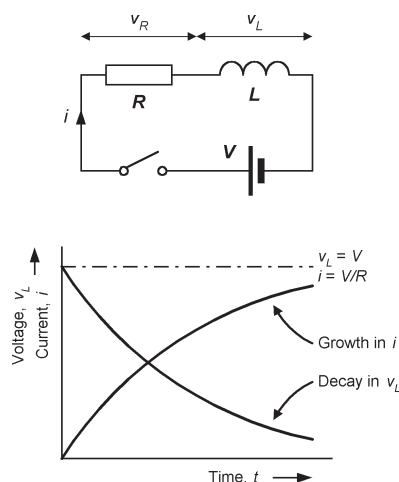


**Figure 4.4.2** Time response of a series L-R network

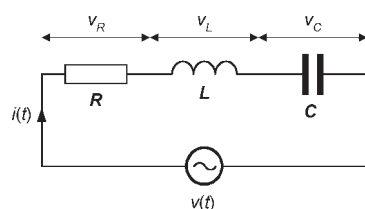


**Figure 4.4.3** Series L-C-R circuit showing forcing function (voltage) and forced function (current)

Earlier in this chapter, and in Chapter 3, we described how circuits containing resistance, inductance, and capacitance, behave when steady AC are applied to them. In this final section of this chapter, we investigate the behaviour of R-L-C circuits when subjected to a sudden change in voltage or current, known as a *transient*. To develop an understanding of how these circuits behave, we shall treat them as systems.

There are two types of system that we shall be concerned with; first-order and second-order systems:

- A *first-order system* is one that can be modelled using a first-order differential equation. First-order electrical systems involve combinations of C and R or L and R.
- A *second-order system* is one that can be modelled using a second-order differential equation. Second-order electrical systems involve combinations of all three types of component; L, C, and R.

You are probably already familiar with the time response of simple C-R and L-R systems. In the first case, the voltage developed across the capacitor in a C-R circuit will grow exponentially (and the current will decay exponentially) as the capacitor is charged from a constant voltage supply (see Figure 4.4.1). In the second case, the inductor voltage will decay exponentially (and the current will grow exponentially) as the inductor is charged from a constant voltage supply (see Figure 4.4.2).

The mathematical relationships between voltage, current, and time for the simple series C-R and L-R circuits are as follows:

For Figure 4.4.1:

$$i = \frac{V}{R} e^{-t/CR} \quad \text{and} \quad v_C = V(1 - e^{-t/CR}).$$

For Figure 4.4.2:

$$i = \frac{V}{R}(1 - e^{-Rt/L}) \quad \text{and} \quad v_L = V(1 - e^{-Rt/L}).$$

The voltage impressed on a first or second-order circuit is often referred to as the *forcing function* whilst the current is known as the *forced response*. In a simple case (as in Figures 4.4.1 and 4.4.2), once the switch has been closed, the forcing function is constant. This is not always the case, however, as the forcing function may *itself* be a function of time. This added degree of complexity requires a different approach to the solution of these, apparently simple, circuits.

Consider the following type of voltage that we met earlier in this chapter:

$$v = 100 + 50 \sin \omega t \text{ V.}$$

When used as a forcing function, this voltage can be thought of as constant step function (of amplitude 100 V) onto which is superimposed a sinusoidal variation (of amplitude 50 V).

In general, we can describe a forcing function as a voltage,  $v(t)$ , where the brackets and the  $t$  remind us that the voltage is a *function of time*. The forced function,  $i(t)$ , would similarly be a function of time (again denoted by the brackets and the  $t$ ).

Now take a look at Figure 4.4.3. Here a forcing function,  $v(t)$ , is applied to a simple L-C-R circuit. By applying Kirchhoff's voltage law we can deduce that:

$$v(t) = v_R + v + v_C$$

$$\text{where } v_R = i(t)R$$

$$\text{and } v_L = L \times (\text{rate of change of current with time}) = L \frac{di(t)}{dt}$$

$$\text{and } v_C = -(\text{area under the current - time graph}) = -\int i(t)dt.$$

$$\text{Thus } v(t) = i(t)R + L \frac{di(t)}{dt} + \frac{1}{C} \int i(t)dt.$$

This is an important relationship and you should have spotted that it is a second-order differential equation. Solving the equation can be problematic but is much simplified by using a technique called the Laplace transform.

## Laplace transforms

The Laplace transform provides us with a means of transforming differential equations into straightforward algebraic equations, thus making the solution of expressions involving several differential or integral terms relatively simple. The technique is as follows:

- (1) Write down the basic expression for the circuit in terms of voltage, current, and component values.
- (2) Transform the basic equation using the table of standard Laplace transforms (each term in the expression is transformed separately).
- (3) Simplify the transformed expression as far as possible. Insert initial values. Also insert component values where these are provided.
- (4) When you simplify the expression, try to arrange it into the same form as one of the standard forms in the table of standard Laplace transforms. Failure to do this will prevent you from performing the inverse transformation (i.e. converting the expression back into the time domain). Note that you may have to use partial fractions to produce an equation containing terms that conform to the standard form.
- (5) Use the table of standard Laplace transforms in reverse to obtain the inverse Laplace transform.

Do not panic if this is beginning to sound very complicated – the three examples that follow will take you through the process on a step by step basis!

## Mathematics in action

### *Laplace transforms*

The Laplace transform of a function of time,  $f(t)$ , is found by multiplying the function by  $e^{-st}$  and integrating the product between the limits of zero and infinity. The result (if it exists) is known as the Laplace transform of  $f(t)$ .

$$\text{Thus: } F(s) = \mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t)dt.$$

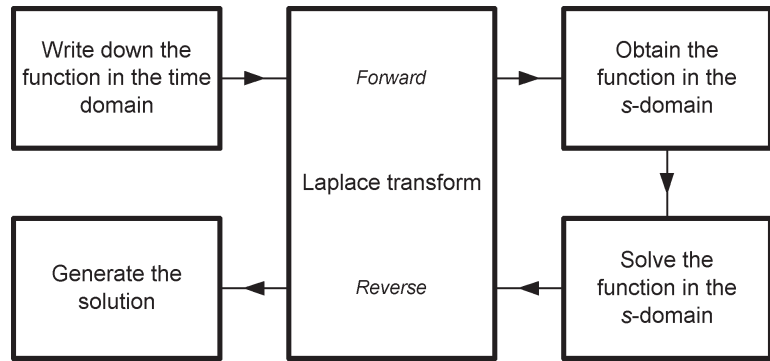


**Figure 4.4.4** Laplace transformation process

Frequently we need only to refer to a table of standard Laplace transforms (Figure 4.4.4.). Some of the most useful of these are summarised in the table below:

| $f(t)$                   | $F(s) = \mathcal{L}\{f(t)\}$                            | Comment               |
|--------------------------|---------------------------------------------------------|-----------------------|
| $a$                      | $\frac{1}{s}$                                           | A constant            |
| $t$                      | $\frac{1}{s^2}$                                         | Ramp                  |
| $e^{at}$                 | $\frac{1}{s-a}$                                         | Exponential growth    |
| $e^{-at}$                | $\frac{1}{s+a}$                                         | Exponential decay     |
| $e^{-at} \sin(\omega t)$ | $\frac{\omega}{(s+a)^2 + \omega^2}$                     | Decaying sine         |
| $e^{-at} \cos(\omega t)$ | $\frac{s+a}{(s+a)^2 + \omega^2}$                        | Decaying cosine       |
| $\sin(\omega t + \phi)$  | $\frac{s \sin \phi + \omega \cos \phi}{s^2 + \omega^2}$ | Sine plus phase angle |
| $\sin(\omega t)$         | $\frac{\omega}{s^2 + \omega^2}$                         | Sine                  |
| $\cos(\omega t)$         | $\frac{s}{s^2 + \omega^2}$                              | Cosine                |
| $\frac{df(t)}{dt}$       | $sF(s) - f(0)$                                          | First differential    |
| $\frac{d^2 f(t)}{dt^2}$  | $s^2 F(s) - sf(0) - \frac{df(0)}{dt}$                   | Second differential   |
| $\int f(t) dt$           | $\frac{1}{s} F(s) + \frac{1}{s} F(0)$                   | Integral              |



## Another view

You can think of the Laplace transform as a device that translates a given function in the time domain,  $f(t)$ , into an equivalent function in the  $s$ -domain,  $F(s)$ . We use the technique to translate a differential equation into an equation in the  $s$ -domain, solve the equation, and then use it in reverse to convert the  $s$ -domain solution to an equivalent solution in the time domain. Figure 4.4.4 illustrates this process.

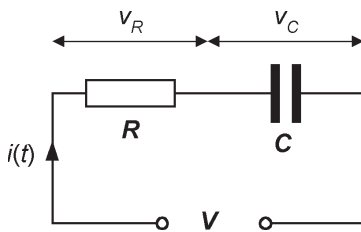


**Figure 4.4.5** See Example 4.4.1

### Example 4.4.1

Use Laplace transforms to derive an expression for the current flowing in the circuit shown in Figure 4.4.5 given that  $i = 0$  at  $t = 0$ .

The voltage,  $V$ , will be the sum of the two voltages,  $v_R$  and  $v_C$ , where  $v_R = Ri(t)$  and

$$v_C = \frac{1}{C} \int i(t) dt.$$

Note that to remind us that the current is a function of time we have used  $i(t)$  to denote current rather than just  $i$  on its own.

$$\text{Thus } V = v_R + v_C = Ri(t) + \frac{1}{C} \int i(t) dt.$$

Applying the Laplace transform gives:

$$\mathcal{L}\{V\} = \mathcal{L}\left\{Ri(t) + \frac{1}{C} \int i(t) dt\right\} = \mathcal{L}\{Ri(t)\} + \mathcal{L}\left\{\frac{1}{C} \int i(t) dt\right\}$$

$$\mathcal{L}\{V\} = R\mathcal{L}\{i(t)\} + \frac{1}{C} \mathcal{L}\left\{\int i(t) dt\right\}.$$

Using the table of standard Laplace transforms on page 357 gives:

$$\frac{V}{s} = Rl(s) + \frac{1}{sC}[l(s) + l(0)].$$

Note that we are now in the  $s$ -domain ( $s$  replaces  $t$ ) and the large  $l$  has replaced the small  $i$  (just as  $F$  replaces  $f$  in the general expression).

Now  $i = 0$  at  $t = 0$  thus  $l(0) = 0$ , hence:

$$\begin{aligned} \frac{V}{s} &= Rl(s) + \frac{1}{sC}[l(s) - (0)] \\ &= Rl(s) + \frac{1}{sC}l(s) = l(s)\left(R + \frac{1}{sC}\right) \end{aligned}$$

$$\text{or } l(s) = \frac{V}{s\left(R + \frac{1}{sC}\right)} = \frac{V}{Rs + \frac{1}{C}} = \frac{V}{R} \left( \frac{1}{s + \frac{1}{CR}} \right).$$

Now  $i(t) = \mathcal{L}^{-1}\{l(s)\}$  thus:

$$\begin{aligned} i(t) &= \mathcal{L}^{-1}\{l(s)\} = \mathcal{L}^{-1}\left\{\frac{V}{R} \left( \frac{1}{s + \frac{1}{CR}} \right)\right\} \\ &= \frac{V}{R} \mathcal{L}^{-1}\left\{\frac{1}{s + \frac{1}{CR}}\right\}. \end{aligned}$$

We now need to find the inverse transform of the foregoing equation. We can do this by examining the table of standard Laplace transforms on page 357 looking for  $F(s)$  of the form

$$\frac{1}{s+a} \quad \text{where } a = \frac{1}{CR}.$$

Notice, from the table, that the inverse transform of  $1/(s+a)$  is  $e^{-at}$ , that is:

$$\mathcal{L}^{-1}\left\{\frac{1}{s+a}\right\} = e^{-at}.$$

$$\text{Hence } i(t) = \mathcal{L}^{-1}\{I(s)\} = \frac{V}{R} \mathcal{L}^{-1}\left\{\frac{1}{s + \frac{1}{CR}}\right\} = \frac{V}{R} (e^{-t/CR}).$$

Thus the current in the circuit is given by:

$$i(t) = \frac{V}{R} (e^{-t/CR}) \text{ (exponential growth).}$$

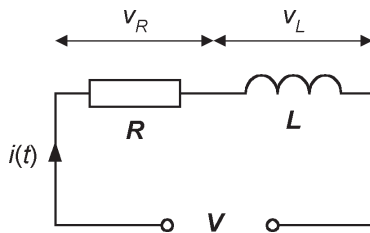


**Figure 4.4.6** See Example 4.4.2

### Example 4.4.2

Use Laplace transforms to derive an expression for the current flowing in the circuit shown in Figure 4.4.6 given that  $i = 0$  at  $t = 0$ . The voltage,  $V$ , will be the sum of the two voltages,  $v_R$  and  $v_L$ ,

where  $v_R = Ri(t)$  and

$v_L = L(di(t)/dt)$  and  $i(t)$  reminds us that the current,  $i$ , is a function of time.

$$\text{Thus } V = v_R + v_L = Ri(t) + L \frac{di(t)}{dt}.$$

Applying the Laplace transform gives:

$$\mathcal{L}\{V\} = \mathcal{L}\left\{Ri(t) + L \frac{di(t)}{dt}\right\} = \mathcal{L}\{Ri(t)\} + \mathcal{L}\left\{L \frac{di(t)}{dt}\right\}$$

$$\mathcal{L}\{V\} = R\mathcal{L}\{i(t)\} + L\mathcal{L}\left\{\frac{di(t)}{dt}\right\}.$$

Using the table of standard Laplace transforms on page 357 gives:

$$\frac{V}{s} = R I(s) + sL[I(s) - i(0)].$$

Note that we are now in the  $s$ -domain ( $s$  replaces  $t$ ) and the large  $I$  has replaced the small  $i$  (just as  $F$  replaces  $f$  in the general expression).

Now  $i = 0$  at  $t = 0$  thus  $i(0) = 0$ , hence:

$$\frac{V}{s} = RI(s) + sL[I(s) - 0] = RI(s) + sLI(s) = I(s) \times (R + sL)$$

$$\text{or } I(s) = \frac{V}{s(R + sL)}.$$

We can simplify this expression by using partial fractions, thus:

$$\frac{V}{s(R + sL)} = \frac{A}{s} + \frac{B}{R + sL} = \frac{A(R + sL) + Bs}{s(R + sL)}$$

$$\text{hence } V = A(R + sL) + Bs = AR + AsL + Bs.$$

We now need to find the values of  $A$ ,  $B$ , and  $C$

$$\text{when } s = 0, \quad V = AR + 0 \quad \text{thus} \quad A = \frac{V}{R},$$

$$\begin{aligned} \text{when } s = \frac{-R}{L}, \quad V &= AR + A\left(\frac{-R}{L}\right)L + B\left(\frac{-R}{L}\right) \\ &= AR - AR - B\frac{R}{L} \end{aligned}$$

$$\text{thus } V = \frac{-BR}{L} \quad \text{or} \quad B = \frac{-VL}{R}.$$

Replacing  $A$  and  $B$  with  $V/R$  and  $-VL/R$ , respectively, gives:

$$\begin{aligned} \frac{V}{s(R + sL)} &= \frac{A}{s} + \frac{B}{R + sL} = \frac{V}{Rs} + \frac{\left(\frac{-VL}{R}\right)}{R + sL} \\ &= \frac{V}{Rs} - \frac{VL}{R(R + sL)}. \end{aligned}$$

$$\text{Thus } I(s) = \frac{V}{Rs} - \frac{VL}{R(R + sL)}.$$

We need to find the inverse of the above equation, since:

$$i(t) = \mathcal{L}^{-1}\{I(s)\} = \mathcal{L}^{-1}\left\{\frac{V}{Rs} - \frac{VL}{R(R + sL)}\right\}$$

$$\begin{aligned} \text{thus } i(t) &= \mathcal{L}^{-1}\left\{\frac{V}{Rs} - \frac{VL}{R(R + sL)}\right\} \\ &= \frac{V}{R} \mathcal{L}^{-1}\left\{\frac{1}{s} - \frac{L}{R + sL}\right\} \end{aligned}$$

$$\text{hence } i(t) = \frac{V}{R} \mathcal{L}^{-1} \left\{ \frac{1}{s} - \frac{1}{s + \frac{R}{L}} \right\}.$$

Referring once again to the table on page 357 gives:

$$i(t) = \frac{V}{R} (1 - e^{-Rt/L}) \quad (\text{exponential decay}).$$

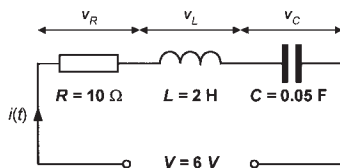


**Figure 4.4.7** See Example 4.4.3

### Example 4.4.3

Use Laplace transforms to derive an expression for the current flowing in the circuit shown in Figure 4.4.7 given that  $i = 0$  at  $t = 0$ . Here, the voltage,  $V$ , will be the sum of three voltages,  $v_R$ ,  $v_L$ , and  $v_C$ , where:

$$v_R = Ri(t), \quad v_L = L \frac{di(t)}{dt}, \quad \text{and} \quad v_C = \frac{1}{C} \int i(t) dt.$$

Yet again  $i(t)$  reminds us that the current,  $i$ , is a function of time.

$$\text{Thus } V = v_R + v_L + v_C = Ri(t) + L \frac{di(t)}{dt} + \frac{1}{C} \int i(t) dt.$$

Applying the Laplace transform gives:

$$\mathcal{L}\{V\} = \mathcal{L}\left\{ Ri(t) + L \frac{di(t)}{dt} + \frac{1}{C} \int i(t) dt \right\},$$

$$\mathcal{L}\{V\} = R\mathcal{L}\{i(t)\} + L\mathcal{L}\left\{ \frac{di(t)}{dt} \right\} + \frac{1}{C} \mathcal{L}\left\{ \int i(t) dt \right\}.$$

Using the table of standard Laplace transforms on page 357 gives:

$$\frac{V}{s} = RI(s) + sL[I(s) - i(0)] + \frac{1}{sC} [I(s) + I(0)].$$

Now  $i = 0$  at  $t = 0$  thus  $i(0) = 0$  and  $I(0) = 0$ , hence

$$\frac{V}{s} = RI(s) + sL[I(s) - 0] + \frac{1}{sC} [I(s) + I(0)]$$

$$\frac{V}{s} = RI(s) + sLI(s) + \frac{1}{sC} I(s) = I(s) \left( R + sL + \frac{1}{sC} \right)$$

$$\text{or } I(s) = \frac{V}{s \left( R + sL + \frac{1}{sC} \right)} = \frac{V}{sR + s^2L + \frac{1}{C}}.$$

Now  $V = 6 \text{ V}$ ,  $L = 2 \text{ H}$ ,  $R = 12 \Omega$  and  $C = 0.05 \text{ F}$ , thus:

$$I(s) = \frac{V}{s^2 L + sR + \frac{1}{C}} = \frac{6}{2s^2 + 12s + \frac{1}{0.05}}$$

$$= \frac{3}{s^2 + 6s + 10} \quad (\text{i})$$

In order to perform the inverse Laplace transformation we need to express Equation (i) in a form that resembles one of the standard forms listed in the table on page 357. The nearest form is as follows:

$$F(s) = \frac{\omega}{(s + a)^2 + \omega^2}, \quad \text{which has } f(t) = e^{-at} \sin \omega t \text{ as its reverse transform.}$$

Fortunately, it is not too difficult to rearrange Equation (i) into this form:

$$I(s) = 3 \times \frac{1}{(s + 3)^2 + 1}$$

from which  $a = 3$  and  $\omega = 1$ , thus the reverse transform is:  
 $i(t) = e^{-3t} \sin t$  (a decaying sine wave).

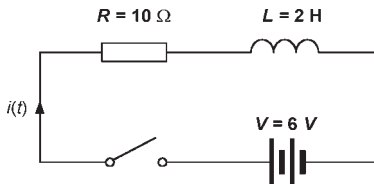


**Figure 4.4.8** See Questions 4.4.1

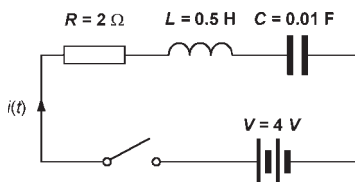


**Figure 4.4.9** See Questions 4.4.1

### Questions 4.4.1

- (1) A voltage given by  $V = 50 \sin(100t + \pi/4)$  is applied to a resistor of  $100 \Omega$ . Use the table of standard Laplace transforms to write down an expression for the current flowing in the resistor in the  $s$ -domain.
- (2) The current flowing in a circuit in the  $s$ -domain is given by:

$$I(s) = \frac{0.025}{0.025s + 1} \text{ mA.}$$

Use the inverse Laplace transform to write down an expression for the current in the time domain. Hence determine the current flowing in the circuit when  $t = 12.5 \text{ ms}$ .

- (3) Use Laplace transforms to derive an expression for the current flowing in the circuit shown in Figure 4.4.8 given that  $i = 0$  at  $t = 0$ .
- (4) Use Laplace transforms to derive an expression for the current flowing in the circuit shown in Figure 4.4.9 given that  $i = 0$  at  $t = 0$ .

# 5

## Mechanical principles

### Summary

In this chapter we will cover a range of mechanical principles which underpin the design and operation of mechanical engineering systems. The material is intended to extend the knowledge gained from Engineering Science (Chapter 3), which you should have studied first. It includes coverage of the strength of materials and mechanics of machines. The primary aim is to provide a firm foundation for work in engineering design and for the material to provide a basis for more advanced study.

### 5.1 COMPLEX LOADING SYSTEMS

You may from your previous work recall the definitions of Poisson's ratio and bulk modulus. In this section we are going to consider these subjects in a little more detail, by determining the behaviour of materials under complex loading conditions. We start by looking once again at Poisson's ratio, which is valid provided the material is only subject to loads within the elastic range:

$$\text{Poisson's ratio } (\nu) = -\frac{\text{Lateral strain}}{\text{Axial strain}}. \quad (5.1.1)$$

The minus sign results from the convention that compressive strains are considered negative and tensile strains positive. Since Poisson's ratio always produces a tensile strain accompanied by a compressive strain, then the laws of arithmetic always produce a *minus* sign.

For a three-dimensional solid, such as a bar in tension, the lateral strain represents a decrease in width (negative lateral strain) and the axial strain represents elongation (positive longitudinal strain). In tables it is normal to show only the magnitudes of the strains considered, so tabulated values of Poisson's ratio are positive. Table 5.1.1 shows typical values of Poisson's ratio, for a variety of materials. Note that for most metals, when there is no other available information, Poisson's ratio may be taken to be 0.3.

**Table 5.1.1** Typical values of Poisson's ratio for a variety of materials

|                           |          |
|---------------------------|----------|
| Aluminium                 | 0.33     |
| Manganese bronze          | 0.34     |
| Cast iron                 | 0.2–0.3  |
| Concrete (non-reinforced) | 0.2–0.3  |
| Marble                    | 0.2–0.3  |
| Nickel                    | 0.31     |
| Nylon                     | 0.4      |
| Rubber                    | 0.4–0.5  |
| Steel                     | 0.27–0.3 |
| Titanium                  | 0.33     |
| Wrought iron              | 0.3      |

### Poisson's ratio and two-dimensional loading

A *two-dimensional stress system* is one in which all the stresses lie within one plane, such as the plane of this paper the  $xy$  plane. Consider the flat plate (Figure 5.1.1) which is subjected to two-dimensional stress loading, as shown.

We know from our previous work that for a material within its elastic limit, we have from Hooke's law that:

$$\text{Elastic modulus } E = \frac{\text{Stress}}{\text{Strain}} = \frac{\sigma}{\varepsilon} \text{ and so strain} = \frac{\sigma}{E}.$$

Also lateral strain =  $\nu$  (axial or longitudinal strain) from *Poisson's ratio*. Then in the  $xx$  direction we have:

$$\text{Strain due to the stress acting along the } xx \text{ direction} = \frac{\sigma_x}{E} \text{ (tensile)}$$

and

$$\text{Strain due to stress acting along the } yy \text{ direction} = -\frac{\nu\sigma_y}{E} \text{ (compressive)}$$

(note minus sign for compressive strain in accordance with convention).

So combined strain due to the stresses acting perpendicular to each other in the  $x$ - and  $y$ -directions is:

$$\text{Combined strain in } xx \text{ direction } \varepsilon_x = \frac{\sigma_x}{E} - \frac{\nu\sigma_y}{E}. \quad (5.1.2)$$

Similarly:

$$\text{Combined strain in } yy \text{ direction } \varepsilon_y = \frac{\sigma_y}{E} - \frac{\nu\sigma_x}{E}. \quad (5.1.3)$$

Multiplying Equations (5.1.2) and (5.1.3) by  $E$  and solving simultaneously for the stresses  $\sigma_x$  and  $\sigma_y$  yields the following equations which relate stress, strain, and Poisson's ratio:

$$\sigma_x = \frac{E}{(1 - \nu^2)} (\varepsilon_x + \nu\varepsilon_y) \quad (5.1.4)$$

and

$$\sigma_y = \frac{E}{(1 - \nu^2)} (\varepsilon_y + \nu\varepsilon_x). \quad (5.1.5)$$

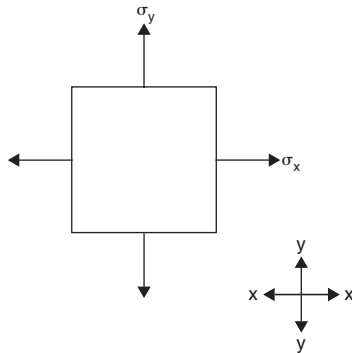


**Figure 5.1.1** Two-dimensional loading system

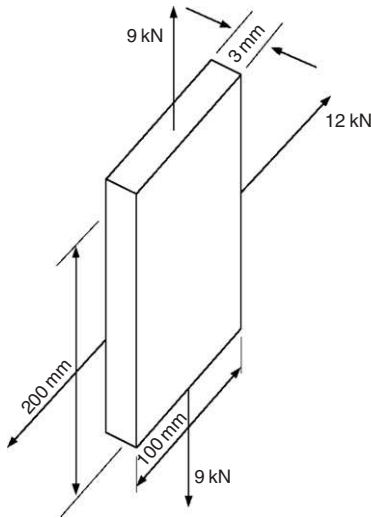


**Figure 5.1.2** Example 5.1.1 – plate loaded in two dimensions

### Example 5.1.1

If the plate loaded as shown in Figure 5.1.2 is made from a steel with  $\nu = 0.3$  and  $E = 205 \text{ GPa}$ .

Determine the changes in dimension in both the  $x$  and  $y$  directions.

In the  $x$  direction the plate is subject to a 12 kN load acting over an area of  $600 \text{ mm}^2$ .

Thus the stress in the  $x$  direction  $= L/A = 20 \text{ N/mm}^2$ .

Similarly, the stress in the  $y$  direction  $= 9000/300 \text{ N/mm}^2 = 30 \text{ N/mm}^2$ .

Then from Equation (5.1.2) the *total strain in the  $x$  direction* is:

$$\varepsilon_x = \frac{20 - 0.3(30) \text{ N/mm}^2}{205 \times 10^3 \text{ N/mm}^2}$$

(notice the careful manipulation of units)

$$\varepsilon_x = 5.37 \times 10^{-5}$$

and since strain = change in length/original length then:

Change in length = (strain)  $\times$  (original length)

$$= (5.37 \times 10^{-5}) \times (100) \text{ mm}$$

$$= 0.00537 \text{ mm.}$$

**So change in dimension in  $x$  direction is an extension = 0.00537 mm.**

Similarly from Equation (5.1.3) the *total strain in the  $y$  direction* is:

$$\varepsilon_y = \frac{30 - 0.3(20)}{205 \times 10^3}$$

$$\varepsilon_y = 1.17 \times 10^{-4}$$

and so change in length  $= (1.17 \times 10^{-4}) \times (200) \text{ mm} = 0.0234 \text{ mm.}$

**So change in dimension in  $y$  direction is an extension = 0.0234 mm.**

### Questions 5.1.1

- (1) A metal bar 250 mm long has a rectangular cross-section of  $60 \text{ mm} \times 25 \text{ mm}$ . It is subjected to an axial tensile force of 60 kN. Find the change in dimensions if the metal has an elastic modulus  $E = 200 \text{ GPa}$  and Poisson's ratio is 0.3.
- (2) A rectangular steel plate 200 mm long and 50 mm wide is subject to a tensile load along its length. If the width of the plate contracts by 0.005 mm find the change in length. Take Poisson's ratio for the steel as 0.28.



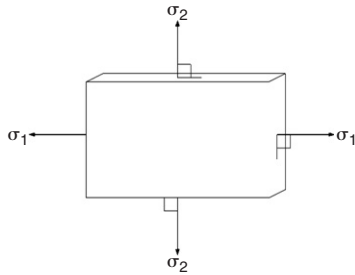


**Figure 5.1.3** Figure for Question 3 (Questions 5.1.1)

- (3) A flat aluminium plate is acted on by mutually perpendicular stresses  $\sigma_1$  and  $\sigma_2$  as shown in Figure 5.1.3. The corresponding strains resulting from these stresses are  $\epsilon_1 = 4.2 \times 10^{-4}$  and  $\epsilon_2 = 9.0 \times 10^{-5}$ . Find the values of the stress if  $E$  for aluminium = 70 GPa and the value of Poisson's ratio is that given for aluminium in Table 5.1.1.

### Poisson's ratio and three-dimensional loading

You know from our work on two-dimensional strain that if, for example, a flat plate is subject to an axial tensile stress (Figure 5.1.4a) then there is an axial extension and lateral contraction of the plate. This relationship is given by Equation 5.1.2, assuming that the axial strain is in the  $x$  direction.

If we extend the argument to three dimensions, you can see (Figure 5.1.4b) that an axial extension of the bar ( $x$  direction) results in a lateral contraction in both the mutually perpendicular  $y$  and  $z$  directions.

Then by analogy, three-dimensional strain resulting from a tensile stress in the  $x$  direction is given by:

$$\epsilon_x = \frac{\sigma_x}{E} - \frac{\nu\sigma_y}{E} - \frac{\nu\sigma_z}{E} \quad \text{or} \quad \epsilon_x = \frac{1}{E}(\sigma_x - \nu\sigma_y - \nu\sigma_z). \quad (5.1.6)$$

**Figure 5.1.4** One-dimensional stress leading to three-dimensional strain

**Figure 5.1.4a**

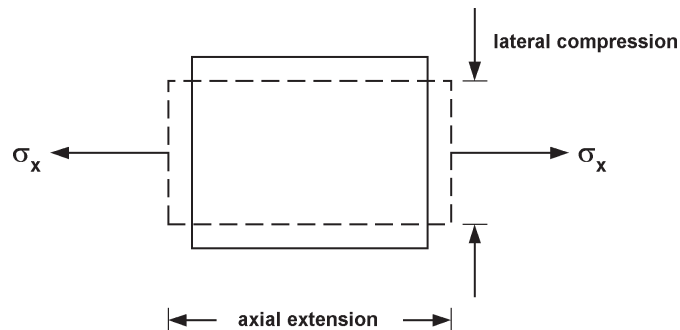
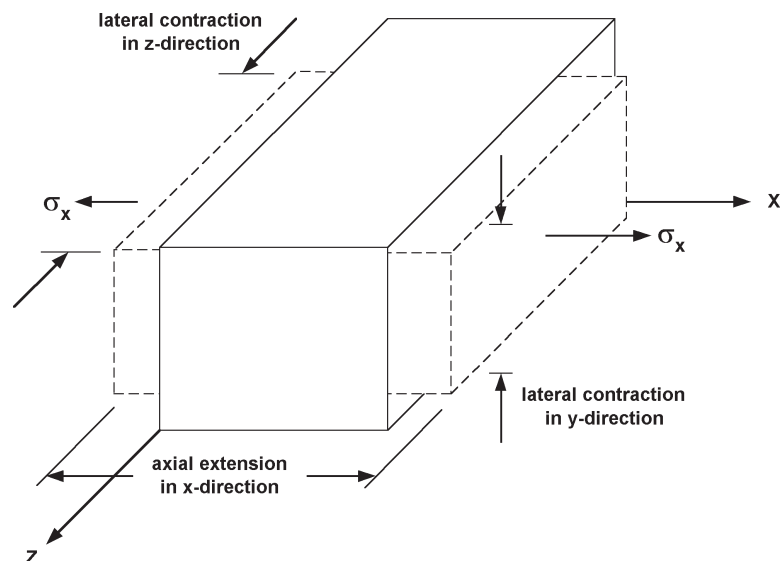


**Figure 5.1.4b**



Similarly the strains in the  $y$  and  $z$  directions in terms of the mutually perpendicular stresses are represented by:

$$\varepsilon_y = \frac{1}{E}(\sigma_y - \nu\sigma_x - \nu\sigma_z). \quad (5.1.7)$$

$$\varepsilon_z = \frac{1}{E}(\sigma_z - \nu\sigma_x - \nu\sigma_y). \quad (5.1.8)$$

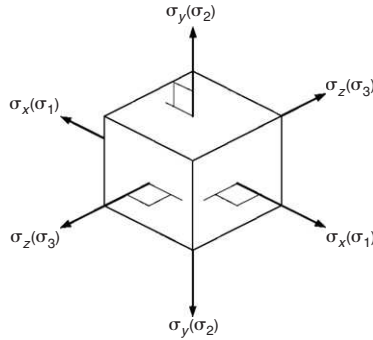


**Figure 5.1.5** *Principal stresses*

### Principal strains in terms of stresses

In the absence of shear stresses on the faces of the element shown in Figure 5.1.5, the stresses  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  are in fact principal stresses. So the principal strain in a given direction may be obtained from the principal stresses.

For principal stresses and strains the suffix,  $x$ ,  $y$ , and  $z$  are replaced by 1, 2, and 3, respectively. So we may rewrite the above equations in terms of principal stresses and strains as:

$$\varepsilon_1 = \frac{1}{E}(\sigma_1 - \nu\sigma_2 - \nu\sigma_3). \quad (5.1.6a)$$

$$\varepsilon_2 = \frac{1}{E}(\sigma_2 - \nu\sigma_1 - \nu\sigma_3). \quad (5.1.7a)$$

$$\varepsilon_3 = \frac{1}{E}(\sigma_3 - \nu\sigma_1 - \nu\sigma_2). \quad (5.1.8a)$$

### Bulk modulus

You will already be familiar with the elastic constants, where Young's modulus  $E$  and the shear modulus  $G$  are defined as the ratio of stress/strain for direct and shear loading, respectively.

There is a third member of the elastic constants known as bulk modulus  $K$ , where

$$K = \frac{p}{\delta V/V}, \text{ where } p \text{ is the uniform pressure or volumetric stress } (\sigma) \text{ and } \frac{\delta V}{V},$$

that is, change in volume over original volume, is the *volumetric strain*. This equation may therefore be written as:

$$\text{Bulk modulus } (K) = \frac{\text{Volumetric stress } (\sigma)}{\text{Volumetric strain } (\varepsilon_v)}. \quad (5.1.9)$$

### Volumetric strain

We know from the definition, that

$$\text{Volumetric strain} = \frac{\text{Change in volume } (\delta V)}{\text{Original volume } (V)}.$$

Now it can be shown that

*volumetric strain = the sum of the linear strains in the  $x$ ,  $y$ , and  $z$  directions*

$$\text{or } \varepsilon_v = \varepsilon_x + \varepsilon_y + \varepsilon_z. \quad (5.1.10)$$



**Figure 5.1.6** Bar subject to uniaxial stress

Circular bar subject to uniaxial tensile stress  $\sigma_x$ , causing lateral contraction and axial elongation with  $\sigma_y = \sigma_z = 0$

Now substituting Equations (5.1.6–5.1.8) into Equation (5.1.10) we produce an equation for the stresses in the  $x$ ,  $y$ , and  $z$  directions in terms of the volumetric strain, that is:

$$\begin{aligned}\epsilon_v = & \frac{1}{E}(\sigma_x - \nu\sigma_y - \nu\sigma_z) + \frac{1}{E}(\sigma_y - \nu\sigma_x - \nu\sigma_z) \\ & + \frac{1}{E}(\sigma_z - \nu\sigma_x - \nu\sigma_y)\end{aligned}$$

and on rearrangement we get:

$$\epsilon_v = \frac{1}{E}(\sigma_x + \sigma_y + \sigma_z)(1 - 2\nu). \quad (5.1.11)$$

Now let us consider the circular bar shown (Figure 5.1.6) which is subject to a uniaxial stress  $\sigma_x$ , then the stresses in the  $y$  and  $z$  planes are zero, therefore from Equation (5.1.11), we have:

$$\epsilon_v = \frac{\sigma_x}{E}(1 - 2\nu) \quad \text{since } \sigma_y = \sigma_z = 0.$$

We also know that the volumetric strain  $\epsilon_v = dV/V$  and letting  $\sigma_x = \sigma$ , then the general equation for calculating the volumetric strain or the change in volume of a circular bar is given as:

$$\epsilon_v = \frac{\delta V}{V} = \frac{\sigma}{E}(1 - 2\nu). \quad (5.1.12)$$

### Example 5.1.2

A bar of circular cross-section (Figure 5.1.7) is subject to a tensile load of 100 kN, which is within the elastic range. The bar is made from a steel with  $E = 210 \text{ GPa}$  and Poisson's ratio = 0.3, it has a diameter of 25 mm and is 1.5 m long. Determine the extension of the bar, the decrease in diameter and the increase in volume of the bar.

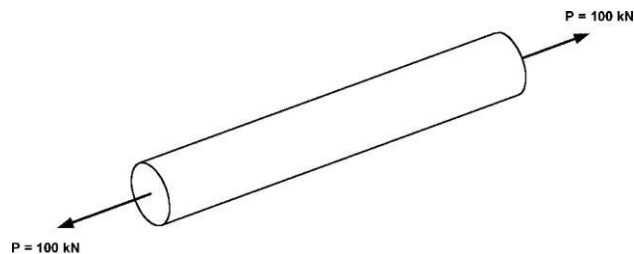


**Figure 5.1.7** Circular section bar subject to a tensile load of 100 kN

Since we are told that the loaded bar is in the elastic range then we may find the axial strain from:

$$\varepsilon = \frac{\sigma}{E} \text{ where } \sigma = 10^5/490.87 \text{ N/mm}^2 = 203.7 \text{ N/mm}^2$$

$$\text{then the axial strain } (\varepsilon) = \frac{203.7 \text{ MPa}}{210 \text{ GPa}} = 0.00097.$$

Now the extension (change in length) = strain  $\times$  original length.

$$\text{Extension} = (0.00097 \times 1.5) = \mathbf{1.455 \text{ mm}}.$$

To find the decrease in diameter we first find the lateral strain, from Poisson's ratio:

$$\varepsilon_{\text{lat}} = -\nu\varepsilon = (0.3 \times 0.00097) = -0.000291$$

Since decrease in diameter = lateral strain  $\times$  original diameter then:

$$\text{Decrease in diameter} = (-0.000291 \times 25) = \mathbf{0.007275 \text{ mm}}.$$

The change in volume is given by Equation (5.1.12), that is:

$$\delta V = V \frac{\sigma}{E} (1 - 2\nu) \quad \text{or} \quad \delta V = V \varepsilon (1 - 2\nu).$$

so

$$\delta V = (490.87 \times 10^{-6} \times 1.5)(0.00097)(1 - 0.6) = \mathbf{285.0 \text{ mm}^3}.$$

### ***Relationship between the elastic constants and Poisson's ratio***

You will remember from our review of shear stress and strain in Chapter 3 (page 133) that the modulus of rigidity, the *shear modulus*  $G$  is given by:

$$G = \frac{\tau}{\gamma}. \quad (5.1.13)$$

## **Mathematics in action**

We will use Equation 5.1.13 and Figure 5.1.8 to help us obtain an important relationship between the shear modulus  $G$ , Young's modulus  $E$ , and Poisson's ratio  $\nu$ .

The square section stress element ABCD (Figure 5.1.8a) is subject to pure shear by the application of stresses  $\tau$ , the faces are then distorted into a rhombus as shown in Figure 5.1.8b. The diagonal  $bd$  has lengthened, while the diagonal  $ac$  has shortened, as a result of the shear stresses (note that the corresponding strains would in practice be very small, they are exaggerated in Figure 5.1.8b for the purpose of illustration).

Now since the strains are small, angles  $abd$  and  $bde$  are each approximately  $45^\circ$  and so the triangle  $abd$  has sides with the ratios  $1:1:\sqrt{2}$ . Now by simple geometry:

The length of side  $bd = s\sqrt{2}(1 + \epsilon_{\max})$ ,

where  $\epsilon_{\max}$  is the normal strain in the  $45^\circ$  direction along  $bd$ . In Figure 5.1.8b the shear strain is given by the angle of deformation in radians. You should now see that the angles  $bda$  and  $dba$  are both equal to  $\pi/4 - \gamma/2$ , and also due to the extension of diagonal  $bd$  the angle  $dab$  is equal to  $\pi/2 + \gamma$ , as shown.

Now using the *cosine rule* for triangle  $dab$  we have:

$$[s\sqrt{2}(1 + \epsilon_{\max})]^2 = s^2 + s^2 - 2s^2 \cos\left(\frac{\pi}{2} + \gamma\right).$$

$$\text{So } 2s^2(1 + \epsilon_{\max})^2 = 2s^2 - 2s^2 \cos\left(\frac{\pi}{2} + \gamma\right)$$

and on division by  $2s^2$  we have:

$$(1 + \epsilon_{\max})^2 = 1 - \cos\left(\frac{\pi}{2} + \gamma\right).$$

Now using the trigonometric identity  $\cos(\pi/2 + x) = -\sin x$  and expanding LHS gives:

$$1 + 2\epsilon_{\max} + (\epsilon_{\max})^2 = 1 + \sin \gamma.$$

Also since  $\epsilon_{\max}$  and  $\gamma$  are very small strains we may use the approximation that  $\sin \gamma = \gamma$  and also ignore  $(\epsilon_{\max})^2$  when compared to  $2\epsilon_{\max}$ .

So now we have:

$$1 + 2\epsilon_{\max} = 1 + \gamma$$

$$\text{or } \epsilon_{\max} = \frac{\gamma}{2}. \quad (5.1.14)$$

This condition shows the relationship for pure shear between the normal strain acting in the  $45^\circ$  direction.

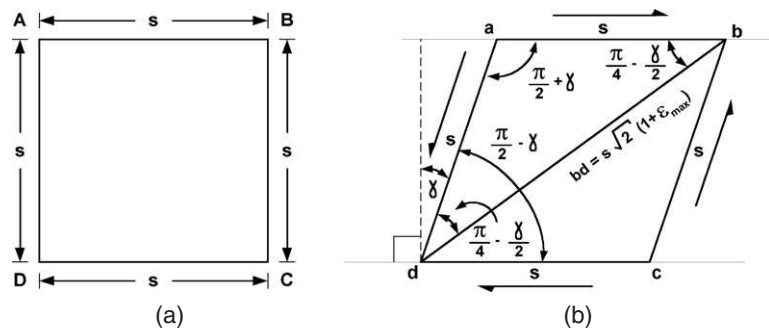


**Figure 5.1.8** Square section stress element subject to pure shear

Equation (5.1.14) shows the relationship between the normal strain in the  $45^\circ$  direction  $\varepsilon_{\max}$  and the shear strain  $\gamma$ .

Now consider the strains that occur in the stress element orientated at zero degrees shown in Figure 5.1.9a. This can be replaced (from Hooke's law) by a system of *direct stresses* orientated at  $45^\circ$  as shown in Figure 5.1.9b.

Then using Equation (5.1.2) for the direct stress system along the diagonals and, noting the sign convention for the stresses we have:

The strain in the direction of positive normal diagonal:

$$\begin{aligned}\varepsilon_{\max} &= \frac{\sigma_1}{E} - \frac{\nu\sigma_2}{E} \\ &= \frac{\tau}{E} - \frac{\nu}{E}(-\tau) \\ \varepsilon_{\max} &= \frac{\tau}{E}(1 + \nu).\end{aligned}\quad (5.1.15)$$

Also from Equation (5.1.13) we have  $\gamma = \frac{\tau}{G}$ .

Now substituting this expression for  $\gamma$  and the expression for  $\varepsilon_{\max}$  Equation (5.1.15) into Equation (5.1.14) gives:

$$\frac{\tau}{E}(1 + \nu) = \frac{\tau/G}{2} \quad \text{and so} \quad G = \frac{E}{2(1 + \nu)}.\quad (5.1.16)$$

Do not worry too much if you were unable to follow the mathematical argument, much more important is that you understand, and are able to use, the relationships we have found. In determining the relationship given in Equation (5.1.16), we have only considered an element of material under *pure shear* conditions. The stresses imposed under these conditions are shown in Figure 5.1.9.

We will return to the relationships between shear stress and shear strain in the next section when we look at the loading of thin-walled cylinders, such as pressure vessels. Before we leave our study of Poisson's ratio and the elastic constants, let us look at one or two examples.

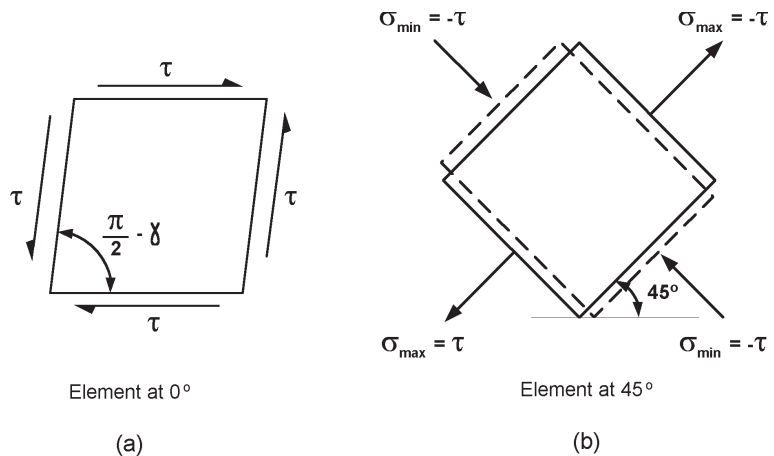


**Figure 5.1.9** Stresses resulting from pure shear

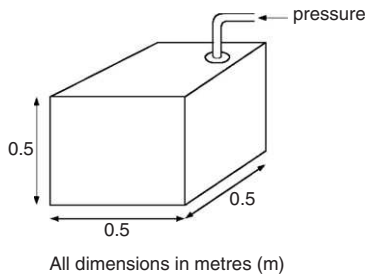


**Figure 5.1.10** Figure for Example 5.1.3

### Example 5.1.3

Assume that the pressure vessel (Figure 5.1.10) contains a pressurised liquid, which subjects the vessel to volumetric (hydrostatic) stress. Show, stating all assumptions, that the elastic constants  $E$  and  $K$ , and Poisson's ratio are related by the expression:

$$K = \frac{E}{3(1 - 2\nu)}.$$

In order to find this relationship we need to remember one very important fact which is:

*Volumetric strain = the sum of the strains in the x, y, and z directions.*

Now for hydrostatic stress the stresses in the x, y, and z directions are equal so:

Volumetric strain =  $3 \times$  the linear strain.

*Assumption:* that all surfaces of the pressure vessel are subject to equal pressure and have the same material properties. Then, sidewall stresses are equal and volumetric strain is equal to three times the linear strain.

So from Equation (5.1.6) where:

$$\varepsilon_x = \frac{1}{E}(\sigma_x - \nu\sigma_y - \nu\sigma_z) \text{ the volumetric strain,}$$

on rearrangement, is given by:

$$\varepsilon_v = \frac{3\sigma}{E}(1 - 2\nu) \quad (5.1.17)$$

where  $\sigma_x = \sigma_y = \sigma_z = \sigma$ .

Now we need an expression involving the bulk modulus  $K$ , if we are to relate the elastic constants  $E$  and  $K$  with Poisson's ratio.

We know that the bulk modulus  $K$  = volumetric stress/ volumetric strain therefore:

$$\text{Volumetric strain } \varepsilon_v = \frac{\sigma}{K} \quad (5.1.18)$$

and from Equations (5.1.17) and (5.1.18) we have:

$$\frac{\sigma}{K} = \frac{3\sigma}{E}(1 - 2\nu) \text{ and on rearrangement:}$$

$$K = \frac{E}{3(1 - 2\nu)}. \quad (5.1.19)$$

Equation (5.1.19) is the relationship we require.

**Example 5.1.4**

A cylindrical compressed air cylinder 1.5 m long and 0.75 m internal diameter has a wall thickness of 8 mm. Find the increase in volume when the cylinder is subject to an internal pressure of 3.5 MPa, also find the values for the constants  $G$  and  $K$ . Given that  $E = 207$  GPa and Poisson's ratio  $\nu = 0.3$ .

For this problem, the increase in volume can be found directly using Equation (5.2.6), which you will meet in the next section. Then:

$$\text{Change in internal volume} = \frac{pd}{4tE}(5 - 4\nu)V$$

where  $p$  = internal pressure,  $\nu$  = Poisson's ratio,  $d$  = internal diameter,  $V$  = internal volume, and  $t$  = wall thickness.

So change in volume:

$$= \frac{(3.5 \times 10^6)(0.75)(5 - 4 \times 0.3)(0.6627)}{(4)(0.008)(207 \times 10^9)}$$

$$\text{change in volume} = 0.000998 \text{ m}^3.$$

Also using Equations (5.1.16) and (5.1.19) then:

$$G = \frac{207 \times 10^9}{2(1 + 0.3)} = 79.6 \text{ GPa}$$

$$\text{and } K = \frac{207 \times 10^9}{3(1 - 0.6)} = \mathbf{172.5 \text{ GPa.}}$$

**Questions 5.1.2**

- (1) The principal strains ( $\varepsilon_1$  and  $\varepsilon_2$ ) at a point on a loaded steel sheet are found to be  $350 \times 10^{-6}$  and  $210 \times 10^{-6}$ , respectively. If the modulus of elasticity  $E = 210$  GPa and Poisson's ratio for the steel is 0.3. Determine the corresponding principal stresses.
- (2) An aluminium alloy has a modulus of elasticity of 80 GPa and a modulus of rigidity of 33 GPa. Determine the value of Poisson's ratio and the bulk modulus.
- (3) A rod 30 mm in diameter and 0.5 m in length is subjected to an axial tensile load of 90 kN. If the rod extends in length by 1.1 mm and there is a decrease in diameter of 0.02 mm. Determine the values of Poisson's ratio and the values of  $E$ ,  $G$ , and  $K$ .

## 5.2 LOADED BEAMS AND CYLINDERS

We continue our theme of complex loading by first considering the way beams bend or deflect under load, where we call upon the knowledge you have already gained when you studied the static outcome in Engineering Science. The loads imposed on pressure vessels as a result of internal and



external loads are then considered. Both thin- and thick-walled pressure vessels will be considered.

## Slope and deflection of beams

You have already been introduced to the deflection curve, in your earlier study of beams in Unit 3. In order that engineers may estimate the loading characteristics of beams in situations where, for example, they are used as building support lintels or supports for engineering machinery. It is important to establish mathematical relationships which enable the appropriately dimensioned beam of the correct material to be chosen.

We have already established one very important relation, engineers theory of bending, from which other important equations may be found.

From Equation (3.1.8) we know that:

$$\frac{M}{I} = \frac{E}{R}$$

and on rearrangement we may write:

$$\frac{1}{R} = \frac{M}{EI} \quad (5.2.1)$$

Now consider Figure 5.2.1 which shows a cantilever beam being deflected downwards (compare with Figure 3.1.3). At A the beam has been deflected by an amount  $y$  and at B the deflection is  $y + dy$ .

The deflection of beams in practice, is very small, therefore from simple geometry for very small angles we may assume that the arc distance  $ds$  is approximately equal to  $dx$ . So as  $d\theta$  approaches zero then:

$$dx = R d\theta \text{ from which } \frac{1}{R} = \frac{d\theta}{dx}.$$

Also, since the arc distance AB is very small, we may approximate it to a straight line which has a gradient (*slope*) equal to the tangent of the angle  $\theta$ , therefore:

$$\theta = \frac{dy}{dx} \text{ and so on differentiation } \frac{d\theta}{dx} = \frac{d^2y}{dx^2}.$$

Then from the above two equations

$$\frac{1}{R} = \frac{d^2y}{dx^2} \quad (5.2.2)$$

Finally combining Equations 5.2.1 and 5.2.2 yields:

$$\frac{d^2y}{dx^2} = \frac{M}{EI} \quad (5.2.3)$$

## Deflections by integration

The above bending moment differential equation provides the fundamental relationship for the deflection curve of a loaded beam. Further manipulation of this equation using integration produces relationships in terms of the *slope* and *deflection* of the beam, as can be seen below.

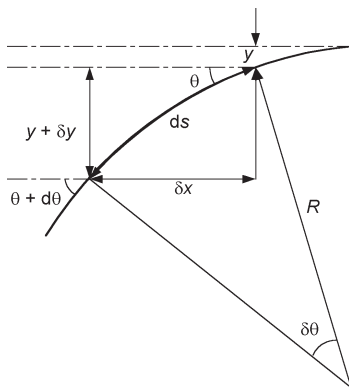


**Figure 5.2.1** Element of deflection curve of beam

## Mathematics in action

### *Deflection by integration*

When the beam is of uniform cross-section and  $M$  is able to be expressed mathematically as a function of  $x$ , we may integrate the equation  $d^2y/dx^2 = M/EI$  with respect to  $x$  and get:

$$\frac{dy}{dx} = \frac{1}{EI} \int M dx + A \quad (5.2.4)$$

where  $A$  is the constant of integration and  $dy/dx$  is the 'slope' of the beam.

Integrating a second time yields:

$$Y = \frac{1}{EI} \iint M dx dx + Ax + B \quad (5.2.5)$$

which produces the expression for the 'deflection' of the beam  $y$ .

Also from Equations (3.1.1) and (3.1.2) we have:

$$\frac{dF}{dx} = -\omega \quad \text{and} \quad \frac{dM}{dx} = F$$

where  $\omega$  is the distributed load and  $F$  the shear force due to point load.

Then differentiating:  $\frac{d^2y}{dx^2} = \frac{M}{EI}$  once more gives:

$$\frac{d^3y}{dx^3} = \frac{dM}{dx} \times \frac{1}{EI}$$

and substituting  $dM/dx = F$  gives:

$$\frac{d^3y}{dx^3} = \frac{F}{EI}. \quad (5.2.6)$$

Finally, using  $\frac{dF}{dx} = -\omega$  and differentiating  $\frac{d^3y}{dx^3} = \frac{F}{EI}$  again, we get:

$$\frac{d^4y}{dx^4} = \frac{dF}{dx} \times \frac{1}{EI} \quad \text{and so}$$

$$\frac{d^4y}{dx^4} = -\frac{\omega}{EI}. \quad (5.2.7)$$

Note that the slope and deflection equations, we have just found all contain  $1/EI$ , where the modulus ( $E$ ) multiplied by the second moment of area ( $I$ ) provides a measure of the *flexural rigidity* of the beam. Since the elastic modulus is a measure of the stiffness of the beam material and the second moment of area (sometimes known as the moment of inertia) is a measure of the beam's resistance to bending, about a particular axis. The *signs* associated with the equations we have just integrated, adopt the convention that the *deflection* ( $y$ ) is *positive upwards* and *bending moments* ( $M$ ) *causing sagging are positive*.

The integration method can be applied to standard cases, which include cantilevers and simply supported beams having a variety of

combinations of concentrated loads, distributed loads, or both. Space does not permit us to consider all possible cases, so just one example of these applications is given in full below. This is followed by a table of some of the more important cases that may be solved by use of the *integrating method*.

When applying the integrating method to the slope and deflection of beams we will use the following symbols, which match with our previous theory:  $x$  as the distance measured from left-hand support of beam or from the fixed end of a cantilever;  $l$  for length;  $y$  for deflection  $w$  per unit length for distributed loads, and  $F$  for shear force due to point load  $W$ .

*Note:* In some textbooks  $v$  is used for  $y$  and  $P$  for  $W$ .

### Example 5.2.1

To determine the deflection and slope of a cantilever subject to a point load consider the cantilever shown in Figure 3.1.3, subject to a single point load at its free end. Now the load causes a 'hogging' bending moment in the beam, which according to our convention, is negative so

$$-M = W(l - x) = Wl - Wx.$$

Using our fundamental Equation 5.2.3, where:

$$\frac{d^2y}{dx^2} = -\frac{M}{EI} \text{ then on substitution } EI \frac{d^2y}{dx^2} = -W(l - x).$$

so on integration we have:

$$EI \frac{dy}{dx} = -W \left( lx - \frac{x^2}{2} \right) + A.$$

The constant of integration can be found from the 'boundary conditions'. The slope of the beam at the fixed end is zero, since the beam is deemed to be level at this point, in other words when  $x = 0$ , the slope  $dy/dx = 0$ . This implies that  $A = 0$ , so our equation becomes:

$$EI \frac{dy}{dx} = -W \left( lx - \frac{x^2}{2} \right)$$

Or the 'slope' of the beam is given by:

$$\frac{dy}{dx} = -\frac{W}{EI} \left( lx - \frac{x^2}{2} \right). \quad (5.2.8)$$

If we integrate the above equation again then we get:

$$y = -\frac{W}{EI} \left( \frac{lx^2}{2} - \frac{x^3}{6} \right) + B.$$

Again the constant of integration  $B$  can be found from the boundary conditions. These are that  $y = 0$  at  $x = 0$ , this infers that  $B = 0$ .

So our equation for deflection  $y$  may be written as:

$$y = -\frac{W}{EI} \left( \frac{x^2}{2} - \frac{x^3}{6} \right). \quad (5.2.9)$$

We can see, quite clearly, from Figure 3.1.3 that the maximum deflection occurs at the free end, where  $x = l$ , then

$$y_{\max} = -\frac{W}{EI} \left( \frac{l^3}{2} - \frac{l^3}{6} \right)$$

$$\text{and so } y_{\max} = -\frac{Wl^3}{3EI}. \quad (5.2.10)$$

Also the maximum slope occurs at the free end and from Equation 5.2.8 when  $x = l$  we have:

$$\left( \frac{dy}{dx} \right)_{\max} = -\frac{Wl^2}{2EI}. \quad (5.2.11)$$

Note the use of boundary conditions to establish values for the constants. You should ensure that you follow the logic of the argument when establishing such conditions, which vary from case to case.

Now let us assume that the cantilever discussed above is made from a steel with an elastic modulus  $E = 210 \text{ GPa}$ , and the point load  $W = 40 \text{ kN}$ . If the maximum deflection  $y_{\max} = 1.5 \text{ mm}$ , determine:

- (i) the second moment of area for the beam;
- (ii) the maximum slope.

(i) The second moment of area is easily calculated using Equation (5.2.10), where:

$$I = \frac{Wl^3}{3Ey_{\max}} = \frac{(40 \times 10^3)(2.5)^3}{(3)(210 \times 10^9)(1.5 \times 10^{-3})} = 661.3 \times 10^{-6} \text{ m}^4$$

$$\text{or } I = 661.3 \times 10^{-6} \times 10^{12} = \mathbf{661.3 \times 10^6 \text{ mm}^4}.$$

Note the careful use of units!

(ii) The maximum slope may now be determined using Equation (5.2.11)

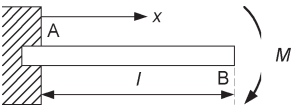
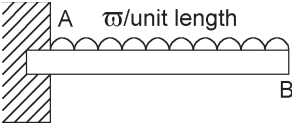
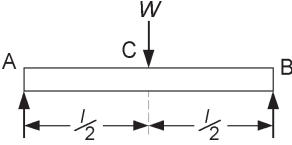
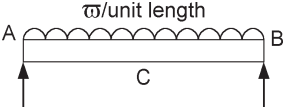
$$\begin{aligned} \left( \frac{dy}{dx} \right)_{\max} &= \frac{Wl^2}{2EI} \\ &= \frac{(40 \times 10^3)(2.5)^2}{(2)(210 \times 10^9)(661.3 \times 10^{-6})} = 0.0022 \text{ m/m}. \end{aligned}$$

**So maximum slope is 0.0022 m/m or 2.2 mm/m.**

The process shown in the above example, where successive integration is applied to the fundamental equations to establish expressions for the slope and deflection, may be used for many standard cases involving plane bending. Remembering that the constants of integration may be found by applying the unique boundary conditions, which exist, for each individual case.

Laid out in Table 5.2.1 are some standard cases for slopes and deflection associated with simply supported beams and cantilevers.

**Table 5.2.1** Some standard cases for slope and deflection of beams

| Situation                                                                                                                          | Slope $\left(\frac{dy}{dx}\right)$                         | Deflection ( $y$ )                                                               | $\left(\frac{dy}{dx}\right)$   | $y_{\max}$                  |
|------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|----------------------------------------------------------------------------------|--------------------------------|-----------------------------|
| Cantilever with end couple<br>                    | $-\frac{Mx}{EI}$                                           | $-\frac{Mx^2}{2EI}$                                                              | $-\frac{Ml}{EI}$ at B          | $-\frac{Mx^2}{EI}$ at B     |
| Cantilever with concentrated end load<br>        | $-\frac{W}{EI}\left(lx - \frac{x^2}{2}\right)$             | $-\frac{W}{EI}\left(\frac{x^2}{2} - \frac{x^3}{6}\right)$                        | $-\frac{Wl^2}{2EI}$ at B       | $-\frac{Wl^3}{3EI}$ at B    |
| Cantilever with distributed load<br>            | $-\frac{w}{2EI}\left(l^2x - lx^2 + \frac{x^3}{3}\right)$   | $-\frac{w}{2EI}\left(\frac{l^2x^2}{2} - \frac{lx^3}{3} + \frac{x^4}{12}\right)$  | $-\frac{wl^3}{6EI}$ at B       | $-\frac{wl^4}{8EI}$ at B    |
| Simply supported beam with point load<br>       | $\frac{W}{EI}\left(\frac{lx}{2} - \frac{x^2}{2}\right)$    | $\frac{W}{2EI}\left(\frac{lx^2}{4} - \frac{x^3}{6} - \frac{l^3}{24}\right)$      | $\frac{Wl^2}{16EI}$ at A and B | $-\frac{Wl^3}{48EI}$ at C   |
| Simply supported beam with distributed load<br> | $\frac{w}{2EI}\left(\frac{l^2x}{4} - \frac{x^3}{3}\right)$ | $\frac{w}{2EI}\left(\frac{l^2x^2}{8} - \frac{x^4}{12} - \frac{5l^4}{192}\right)$ | $\frac{wl^3}{24EI}$ at A and B | $-\frac{5wl^4}{384EI}$ at C |

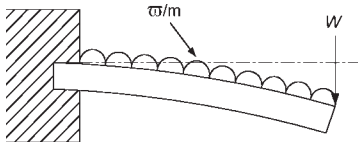


**Figure 5.2.2** Cantilever beam subject to loads causing deflection

### Principle of superposition

This principle states that if, one relationship connecting the bending moment, slope and deflection for a particular loading has been determined, for example a cantilever subject to a concentrated load at its free end (Figure 5.2.2). Then we may add algebraically, any other known relationship to it for a given loading situation.

The cantilever shown in Figure 5.2.2 is subject to both a concentrated load at its free end and a distributed load over its whole length so, for example, using the expressions for maximum deflections in Table 5.2.1 and, the principle of superposition. We may write:

$$y_{\max} = -\frac{Wl^3}{3EI} + \frac{wl^4}{8EI}.$$

### Macauly's method

No doubt if you have attempted Question 2 you will realise that the method of successive integration to determine the slope and deflection equations for each separate loading situation, is tedious and time consuming. We are required to find values for the constants of integration by considering boundary conditions on each separate occasion and, apart from the most simple of cases, this can involve quite complex algebraic processes. All is not lost however, since a much simpler method exists which enables us to formulate one equation which takes into account the bending moment expression for the whole beam and also enables us to find the necessary boundary conditions much more easily. This simplified process is known as *Macauly's method*.

To assist us in formulating the Macauly expression for a given loading situation of a beam, we need to use Macauly functions. The graph of the Macauly unit ramp function (Figure 5.2.3a) will help us to understand its nature.

Macauly functions are used to represent quantities that begin at some particular point on the  $x$  axis. In our case for the unit ramp function, this is at  $x = a$ . To the left of the point  $a$  the function has the value zero, to the right of  $a$ , the function denoted by  $F_1$  has the value  $(x - a)$ . This relationship may be written mathematically as:

$$F_1(x) = \langle x - a \rangle^1 = \begin{cases} 0 & \text{when } x \leq a \\ x - a & \text{when } x \geq a \end{cases}$$

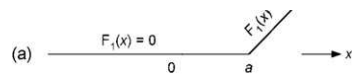
Now the pointed brackets indicate that the function is a *discontinuity function*. At zero our function can take *two* values, zero or one. This may be expressed mathematically as:

$$F_0(x) = \langle x - a \rangle^0 = \begin{cases} 0 & \text{when } x \leq a \\ 1 & \text{when } x \geq a \end{cases}$$

So what does all this mean in practice? If we apply our function to a beam where a point load  $W$  is applied some way along the beam

Graph of Macauly function  $F_1$  (unit ramp function)

$$F_1(x) = \langle x - a \rangle^1 \text{ where } F = \text{function.}$$

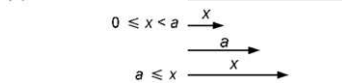


At point 1,  $M = 0$

At point 2,  $M = W(x - a)$

So Macauly expression for all values of  $x$  is

$$M = W\langle x - a \rangle$$



At point 1,  $M = 0$

At point 2,  $M = \frac{1}{2}w(x - a)^2$

So again Macauly expression for all values of  $x$  is

$$M = \frac{1}{2}w\langle x - a \rangle^2$$

**Figure 5.2.3** Macauly function application to beams

(Figure 5.2.3b) at a distance  $a$ . We can determine the bending moments in the normal way, by first considering forces to the left of  $x$ , which gives:

$$M = 0 \quad \text{for } x \geq 0 < a.$$

Also for forces to the right of  $a$ , we may write the bending moment as:

$$M = W(x - a) \quad \text{for } a \leq x.$$

Then the Macaulay expression for the bending moment for all values of  $x$  is given as:

$$M = W\langle x - a \rangle.$$

Now in using the Macaulay expression we must remember that we are applying the unit ramp function and step function and different rules apply. In particular, all terms which make the value inside the bracket negative, are given the value zero.

Figure 5.2.3c shows our beam subject to a distributed load commencing at distance  $a$ . Again the bending moment at a distance  $x$  along the beam may be determined in the normal manner. Considering forces to the left of  $x$ , we have:

$$M = 0 \quad \text{when } 0 \leq x \leq a$$

$$\text{and } M = \frac{1}{2} w \langle x - a \rangle^2.$$

Thus the Macaulay expression for the bending moment for all values of  $x$  is:

$$M = \frac{1}{2} w \langle x - a \rangle^2.$$

Do remember the constraints that apply when considering the expression inside the pointed brackets.

The method you should adopt in determining the Macaulay expression for the whole beam is illustrated by the next example.

We will now consider an example, where the total procedure for obtaining the Macaulay expression for the entire beam is illustrated.

### Example 5.2.2

Determine the Macaulay expressions for the slope and deflection of the simply supported beam shown in Figure 5.2.4.

Now we take the origin from the extreme left-hand end, as shown in Figure 5.2.4 and consider a section at the extreme right-hand end of the beam from which we write down the Macaulay expression. Once written the ramp function rules apply.

So taking moments in the normal way gives:

$$M = R_A x - W_1(x - a) - W_2(x - b)$$

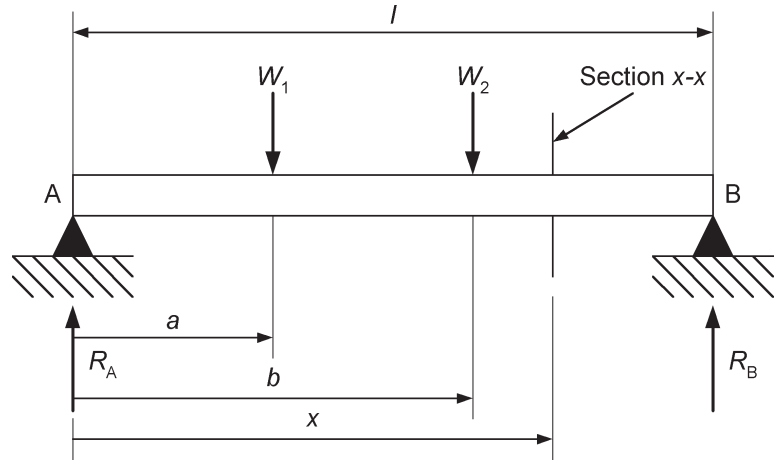


**Figure 5.2.4** Figure for Example 5.2.2

therefore the Macaulay expression for the entire beam is:

$$M = R_A x - W_1 \langle x - a \rangle - W_2 \langle x - b \rangle$$

and using Equation 5.2.2 where  $\frac{d^2 y}{dx^2} = \frac{M}{EI}$  gives:

$$EI \frac{d^2 y}{dx^2} = (R_A x - W_1 \langle x - a \rangle - W_2 \langle x - b \rangle)$$

and integrating in the normal way gives:

$$\frac{dy}{dx} = \frac{1}{EI} \left( \frac{R_A x^2}{2} - W_1 \frac{\langle x - a \rangle^2}{2} - W_2 \frac{\langle x - b \rangle^2}{2} + A \right)$$

and integrating a second time gives the expression for the deflection

$$y = \frac{1}{EI} \left( \frac{R_A x^3}{6} - W_1 \frac{\langle x - a \rangle^3}{6} - W_2 \frac{\langle x - b \rangle^3}{6} + Ax + B \right).$$

The constraints of integration can be found by applying the boundary conditions. At the supports the deflection is zero so when  $x = 0$ ,  $y = 0$  and when  $x = l$ ,  $y = 0$ .

From the above equation it can be seen that when  $x = 0$ ,  $y = 0$  and so  $B = 0$ .

Also, when  $x = l$ ,  $y = 0$  then from the slope equation, the value of  $A$  can be found by evaluating the expression

$$A = \frac{R_1 l^2}{6} - \frac{F_1(l-a)^3}{6l} - \frac{F_2(l-b)^3}{6l}.$$



Therefore, we have all that is necessary to determine the slope and deflection for any point along the beam.

Do remember that all negative Macaulay brackets should be treated as zero.

### Using the Macaulay method for uniformly distributed loads

In using the Macaulay ramp function for a uniformly distributed load (UDL) we started from a point  $a$ , an arbitrary distance from the zero datum, this was in order to comply with the requirements of the function. We then made the assumption that the UDL continued to the right-hand extremity of the beam. This has two consequences for use of the Macaulay method.

- (i) For a UDL that covers the complete span of the beam, the Macaulay method is not appropriate. If we wish to find the slope and deflection for UDLs, which span the entire beam, then we need to use the successive integration methods, discussed earlier.
- (ii) For a UDL that starts at  $a$ , but does not continue for the entire length of the beam (Figure 5.2.5a) we cannot directly apply the Macaulay terms to our bending moment equation, for the reason given above. We need to modify our method in some way to accommodate this type of loading. In effect, what we do is allow our UDL to continue to the extremity of the beam (Figure 5.2.5b) and counter this additional loading by introducing an equal and opposite loading into the bending moment expression.

Figure 5.2.5a shows a simply supported beam with a UDL positioned between point  $a$  and  $b$ . If we imagine that this load is extended to the end of the beam (Figure 5.2.5b) then our bending moment equation, as previously explained, would be:

$$BM = R_A x - \frac{\omega}{2} \langle (x - a)^2 \rangle$$

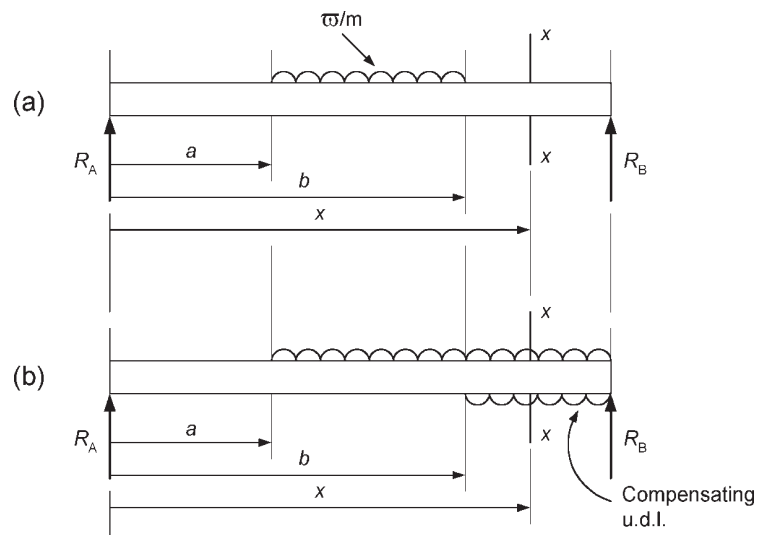


**Figure 5.2.5** Macaulay compensation method for UDL that does not span entire length

This of course is incorrect, but if we now counter-balance the increase in the UDL by a corresponding decrease over the distance  $(x - b)$ , and introduce this as an additional Macaulay term, we obtain the correct Macaulay equation for a beam subject to a partial length UDL. The corrected equation is:

$$BM = R_A x - \frac{w}{2} \langle x - a \rangle^2 + \frac{w}{2} \langle x - b \rangle^2.$$

### Example 5.2.3

For the beam shown in Figure 5.2.6. Determine the deflection of the beam at its centre, if  $EI = 100 \text{ MNm}^2$ .

We first find the reactions at the supports by taking moments about  $R_B$ , then

$$10R_A = (30 \times 7) + (80 \times 5) + (50 \times 3)$$

therefore  $R_A = 76 \text{ kN}$  and so  $R_B = 84 \text{ kN}$ .

Taking the origin as  $R_A$  we apply the Macaulay method shown previously. So by considering section  $x-x$  to the extreme right of the beam and taking moments of the forces to the left of  $x-x$  we get:

$$BM = 10^3 \left[ 76x - 30\langle x - 3 \rangle - 50\langle x - 7 \rangle - \frac{20}{2}\langle x - 3 \rangle^2 + \frac{20}{2}\langle x - 7 \rangle^2 \right]$$

$$\text{and since } \frac{d^2 y}{dx^2} = \frac{BM}{EI}$$

then

$$\frac{d^2 y}{dx^2} = \frac{10^3}{EI} \left[ 76x - 30\langle x - 3 \rangle - 50\langle x - 7 \rangle - \frac{20}{2}\langle x - 3 \rangle^2 + \frac{20}{2}\langle x - 7 \rangle^2 \right]$$

therefore

$$\frac{dy}{dx} = \frac{10^3}{EI} \left[ \frac{76x^2}{2} - \frac{30}{2}\langle x - 3 \rangle^2 - \frac{50}{2}\langle x - 7 \rangle^2 - \frac{20}{6}\langle x - 3 \rangle^3 + \frac{20}{6}\langle x - 7 \rangle^3 + A \right]$$

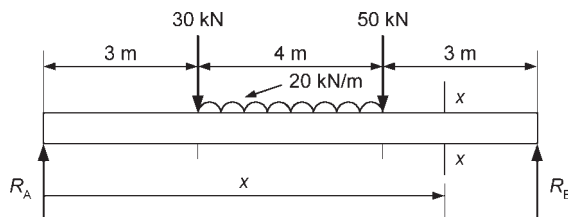


**Figure 5.2.6** Figure for Example 5.2.3

and

$$y = \frac{10^3}{EI} \left[ \frac{76x^3}{6} - \frac{30}{6}\langle x-3 \rangle^3 - \frac{50}{6}\langle x-7 \rangle^3 - \frac{20}{24}\langle x-3 \rangle^4 + \frac{20}{24}\langle x-7 \rangle^4 + Ax + B \right]$$

Now the boundary conditions yield the following values. When  $x = 0$ ,  $y = 0$ , so that  $B = 0$ , since all the Macaulay terms are negative and so equate to zero.

When  $x = 10$ ,  $y = 0$ , now all the Macaulay terms are positive and so, we find  $A = -879.3$ .

When  $x = 5$ , then:

$$y = \frac{10^3}{EI} \left[ \frac{(76)(5)^3}{6} - \frac{30}{6}(2)^3 - \frac{20}{24}(2)^4 - 4396.7 \right]$$

$$y = \frac{(10^3)(-2866.7)}{100 \times 10^6}, \text{ giving}$$

$$y = -0.0287 \text{ m or } -28.7 \text{ mm.}$$

The exercises set out below should help you to consolidate, this rather difficult subject. You will need to refer to Chapter 6, if you are unclear on the mathematics associated with establishing ( $I$ ) the second moment of area or the slope and deflection equations.

### Questions 5.2.1

- (1) Use the free-body approach to draw the shear force and bending moment diagrams for the simply supported beams (Figure 5.2.7) shown below, giving values for the maximum bending moment in each case.
- (2) A circular section cantilever beam 2 m long and 80 mm in diameter, supports a concentrated load of 1.2 kN acting 1.5 m from its fixed end. If the elastic modulus for the beam is  $E = 210 \text{ GPa}$ , calculate the deflection at the load.
- (3) A beam simply supported at its ends is 4 m long and carries point loads of 60 and 30 kN at distances 1 and 2.5 m from the left-hand end of the beam. Determine the position and magnitude of the maximum deflection of the beam, given that  $EI = 16 \times 10^6 \text{ Nm}^2$ .
- (4) In a laboratory experiment, a steel beam 1.8 m long with rectangular cross-section 25 mm wide and 5 mm deep is simply supported on two knife edges 0.8 m apart spanning the centre section of the beam. A load of 30 N is hung at either end of the beam. Neglecting the mass of the beam and given that the steel has

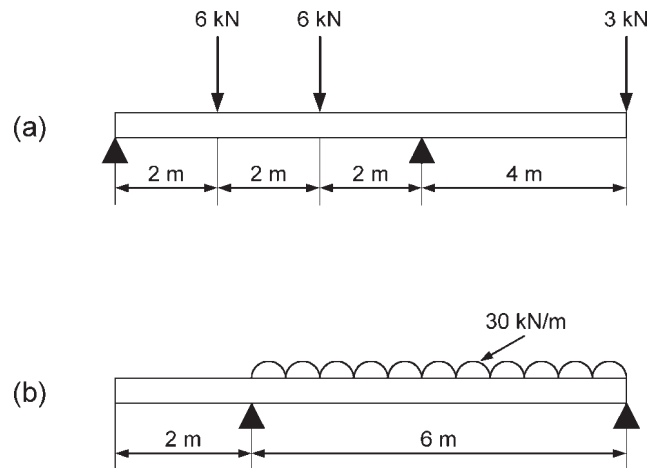


**Figure 5.2.7** Figure for Question 5.2.1 (1)

an elastic modulus  $E = 200 \text{ GN/m}^2$ :

- (i) show that the bending moment is constant between the supports and find the value of this bending moment;
  - (ii) find the maximum stress due to bending;
  - (iii) calculate the radius of curvature between the supports.
- (5) A beam of length 6 m is simply supported at its ends and carries a distributed load ( $\omega = 10 \text{ kN/m}$ ) which commences 1 m from the left-hand end of the beam and continues to the right-hand end of the beam. Two point loads of 30 and 40 kN act 1 and 5 m, respectively, from the left-hand end. If the beam material is made of a steel with  $EI = 20 \times 10^6 \text{ Nm}^2$  determine the deflection at the centre.

## 5.3 PRESSURE VESSELS

Pressure vessels are divided into two major categories, thin- and thick-walled. We start by considering thin-walled pressure vessels that are used, for example, as compressed air containers, boilers, submarine hulls, aircraft fuselages, condenser casings, and hydraulic reservoirs. We then take a look at thick-walled vessels subject to pressure, these for example, function as; hydraulic linear actuators, extrusion dies, gun barrels, and high-pressure gas bottles.

### Thin-walled pressure vessels

When a thin-walled pressure vessel is subject to internal pressure, three mutually perpendicular principal stresses are set up in the vessel material. These are the *hoop stress* (often referred to as the circumferential stress), the *radial stress*, and the *longitudinal stress*. We may define a *thin-walled pressure vessel* as one in which the *ratio* of the wall thickness to that of the inside diameter of the vessel is *greater* than 1:20, that is *less* than the fraction 1/20. Under these circumstances it is reasonable to assume that the hoop and longitudinal stresses are uniform across the wall thickness and that the radial stress is so small in comparison with the hoop and longitudinal stresses that for the purpose of our calculations it may be ignored.

**Figure 5.3.1** Vessel subject to internal pressure

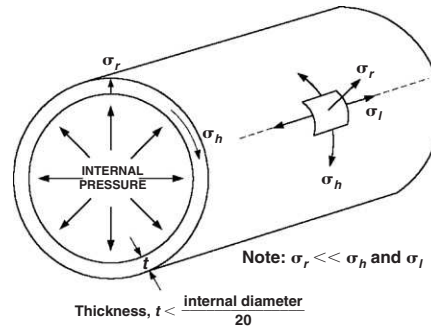


Figure 5.3.1 shows a vessel subject to internal pressure and the nature of the resulting, hoop, longitudinal and radial stresses.

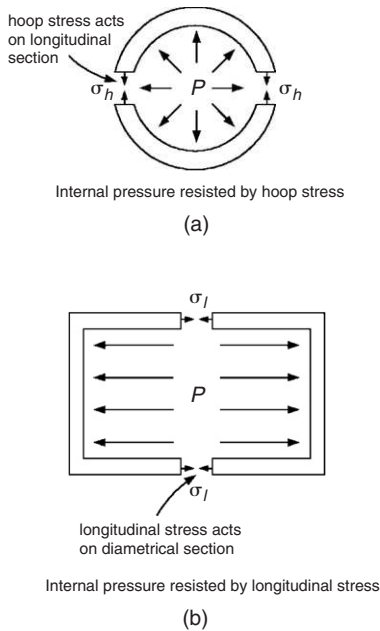


**Figure 5.3.2** Hoop and longitudinal stress resulting from internal pressure

### Hoop and longitudinal stress

*Hoop or circumferential stress* is the stress which resists the bursting effect of the applied internal pressure. If the internal diameter is  $d$ , the applied pressure is  $p$ , the length is  $L$ , and the wall thickness is  $t$ , then the force tending to burst or separate the vessel (Figure 5.3.2a) is given by pressure  $\times$  internal diameter  $\times$  length or  $force = pdL$ . This is resisted by the hoop stress ( $\sigma_h$ ) acting on an area  $2tL$ , so for equilibrium these values may be equated, that is:

$$\text{Hoop stress} = \frac{F}{A} \quad \text{so} \quad F = \sigma_h \times A \quad \text{or} \quad F = \sigma_h \times 2tL$$

also from above  $F = pdL$  then:

$$pdL = 2\sigma_h tL \quad \text{so} \quad \sigma_h = \frac{pd}{2t}. \quad (5.3.1)$$

In a similar manner the force trying to separate the pressure vessel across its diameter, that is longitudinally (Figure 5.3.2b) is given by:

$$\text{Pressure} \times \text{Area} = \frac{p\pi d^2}{4}.$$

This force is resisted by the longitudinal stress ( $\sigma_l$ ) acting on the area  $\pi dt$  so for equilibrium we have:

$$\frac{p\pi d^2}{4} = \sigma_l \pi dt \quad \text{so} \quad \sigma_l = \frac{pd}{4t}. \quad (5.3.2)$$

In developing the formulae for hoop and longitudinal stress we have ignored the efficiency of the joints that go to make up a pressure vessel. You will note that the hoop stress acts on a longitudinal section and that the longitudinal stress acts on a diametrical section (Figure 5.3.2). It therefore follows that the efficiency of the longitudinal joints which go to make up the pressure vessel directly affects the hoop stress at the joint. Similarly the longitudinal stress is affected by the efficiency of the diametrical joints.

The joint efficiency ( $\eta$ ) is often quoted as the strength of the joint/strength of the undrilled plate. So taking this into account for hoop stress we may write:

$$\sigma_h = \frac{pd}{2t\eta_l}. \quad (5.3.3)$$

Similarly for longitudinal stress we may write:

$$\sigma_l = \frac{pd}{4t\eta_h}. \quad (5.3.4)$$

### Example 5.3.1

A cylindrical compressed air vessel has an internal diameter of 50 cm and is manufactured from steel plate 4 mm thick. The longitudinal plate joints have an efficiency of 90% and the circumferential joints have an efficiency of 55%. Determine the maximum air pressure which may be accommodated by the vessel if the maximum permissible tensile stress of the steel is 100 MPa.

Using Equations (5.3.3) and (5.3.4) we have:

$$(i) \quad \sigma_h = \frac{pd}{2\eta_l t} = \frac{0.5p}{2 \times 0.9 \times 0.004}$$

and since maximum stress = 100 MPa then the maximum air pressure accommodated by the longitudinal joints is 1440 kPa.

$$(ii) \quad \text{similarly: } \sigma_l = \frac{pd}{4\eta_h t} = \frac{0.5p}{4 \times 0.55 \times 0.004}$$

then the maximum air pressure accommodated by circumferential joints is 1760 kPa.

So the maximum air pressure is determined by the longitudinal joints and is 1440 kPa.

### Dimensional change in thin cylinders

The change in length of a thin cylinder subject to internal pressure may be determined from the longitudinal strain, provided we ignore the radial stress. The longitudinal strain may be found from Equation (5.1.2), that is:

$$\varepsilon_l = \frac{\sigma_l}{E} - \frac{\nu\sigma_h}{E}$$

and since change in length = longitudinal strain  $\times$  original length, then:

$$\text{Change in length} = \frac{1}{E}(\sigma_l - \nu\sigma_h)l$$

so on substituting Equations (5.3.1) and (5.3.2) for hoop and longitudinal stress into the above equation for change in length, then:

$$\text{Change in length} = \frac{pd}{4tE} (1 - 2\nu)l. \quad (5.3.5)$$

We can also find a relationship for the change in internal volume of a thin-walled cylinder by considering some of our earlier work on volumetric strain. You will remember that volumetric strain is equal to the sum of the mutually perpendicular linear strains (Equation 5.1.10). For a

cylinder we have longitudinal strain ( $\varepsilon_l$ ) and two perpendicular diametrical strains, so from:

$$\varepsilon_v = \varepsilon_x + \varepsilon_y + \varepsilon_z$$

on substituting  $2\varepsilon_D$  (diametrical strain for  $\varepsilon_y + \varepsilon_z$ ) then:

the volumetric strain for a cylinder is  $= \varepsilon_l + 2\varepsilon_D$

and knowing that longitudinal strain  $= 1/E (\sigma_l - \nu\sigma_h)$  and substituting into above equation for volumetric strain, then:

$$\varepsilon_v = \frac{1}{E}(\sigma_l - \nu\sigma_h) + \frac{2}{E}(\sigma_h - \nu\sigma_l)$$

and finally, substituting Equation (5.3.5) into above equation after simplification, gives:

$$\text{Change in internal volume} = \frac{pd}{4tE}(5 - 4\nu)V. \quad (5.3.6)$$

The full derivation of Equation (5.3.6) together with the determination of a relationship for the change in diameter is left as an exercise at the end of this section.

### Example 5.3.2

A thin cylinder has an internal diameter of 60 mm, and is 300 mm in length. If the cylinder has a wall thickness of 4 mm and is subject to an internal pressure of 80 kPa. Then assuming Poisson's ratio is 0.3 and  $E = 210$  GPa, then determine the:

- (a) change in length of the cylinder;
- (b) hoop stress;
- (c) longitudinal stress.

(a) Using Equation (5.3.5), then change in length

$$\begin{aligned} &= \frac{pd}{4tE}(1 - 2\nu)l \\ &= \frac{80 \times 10^3 \times 60 \times 10^{-3} \times 300 \times 10^{-3}}{4 \times 4 \times 10^{-3} \times 210 \times 10^9}(1 - 0.6) \\ &= 0.1714 \times 10^{-6} \text{ m or } \mathbf{0.1714 \mu\text{m}.} \end{aligned}$$

$$\begin{aligned} \text{(b) Hoop stress is given by } \sigma_h &= \frac{pd}{2t} \\ &= \frac{80 \times 10^3 \times 60 \times 10^{-3}}{2 \times 4 \times 10^{-3}} \\ &= \mathbf{0.6 \text{ MPa}.} \end{aligned}$$

(c) Longitudinal stress is given by

$$\sigma_L = \frac{pd}{4t} = \frac{80 \times 10^3 \times 60 \times 10^{-3}}{4 \times 4 \times 10^{-3}} = \mathbf{0.3 \text{ MPa}.}$$

### Pressure vessel applications

An application of the theory of thin cylinders and other thin- and thick-walled shells may be applied to the design and manufacture of pressure vessels. These vessels could be used for food or chemical processing, the storage of gases and liquids under pressure or even the fuselage of a pressurised aircraft. Whatever the industrial use, formulae can be simply derived from some of the relationships we have already found.

Let us first consider some useful relationships which aid the design of cylinders and spherical shells for the plant and process industry.

For a thin-walled cylindrical shell the minimum thickness required to resist internal pressure can be determined from Equation (5.3.1). We use this equation because the hoop stress is greater than the longitudinal stress and so must always be considered first.

Let  $d$  be the internal diameter and  $e$  be the minimum thickness required, then the mean diameter will be given by:

$$d + (2 \times e/2) \text{ or simply } (d + e)$$

and substituting this into Equation (5.3.1) gives:

$$e = \frac{p(d + e)}{2\sigma_d}$$

where  $\sigma_d$  is the design stress and  $p$  the internal pressure. Rearranging the above formula gives:

$$e = \frac{pd}{2\sigma_d - p} \quad (5.3.7)$$

which is the form the equation takes in British Standards (BS) 5500, the standards manual used by pressure vessel designers.

The relationship for the principal stresses of a spherical shell is similar to those we have found for the hoop and longitudinal stress (principal stresses) of a thin-walled cylinder.

That is, the principal stresses for a spherical shell are given by:

$$\sigma_1 = \sigma_2 = \frac{pd}{4t} \quad (5.3.8)$$

Then in a similar manner to above, an equation for minimum thickness of a sphere can be obtained from Equation (5.3.8), that is:

$$e = \frac{pd}{4\sigma_d - p} \quad (5.3.9)$$

Equation (5.3.9) differs slightly from that given in BS 5500, because it is derived from thick-walled shell theory, which we will be looking at next. The formula for determining the minimum thickness of material required to resist internal pressure within a sphere is given in BS 5500 as:

$$e = \frac{pd}{4\sigma_d - 1.2p} \quad (5.3.10)$$

If our pressure vessel has welded joints then we need to take into account the integrity of these joints, by using joint efficiency factors ( $\eta$ ). The efficiency of welded joints are dependent on the heat treatment



received after welding, and on the amount of non-destructive examination used to ascertain their integrity. If the joint integrity can be totally relied upon as a result of heat treatment and non-destructive examination, then the joint may be considered to be 100% efficient. In practice welded joint efficiency factors vary from about 0.65 up to 1.0. In other words the efficiency of welded joints normally ranges from around 65% up to 100%.

If we wish to consider cylindrical and spherical welded vessels then Equations 5.3.7 and 5.3.10 may be written as follows:

$$\text{For a cylinder } e = \frac{pd}{2\sigma_d \eta - p} \quad (5.3.11)$$

$$\text{For a sphere } e = \frac{pd}{4\sigma_d \eta - p} \quad (5.3.12)$$

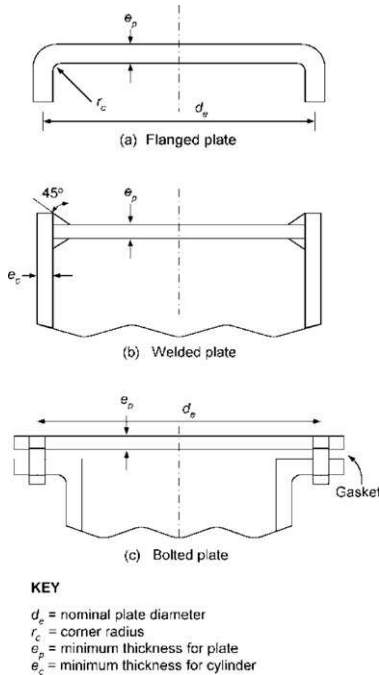


**Figure 5.3.3** Flat plate closures for cylindrical pressure vessel

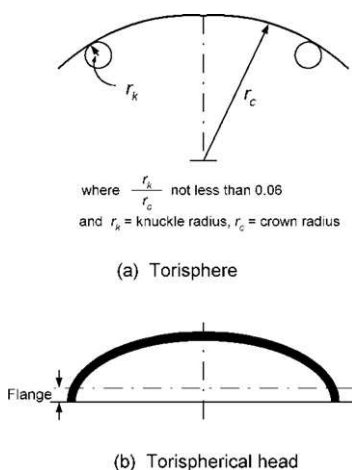


**Figure 5.3.4** Torispherical domed head for cylindrical pressure vessel

When designing cylindrical pressure vessels due consideration must be given to the way in which the vessels are closed at their ends. For example, flat plates or domes of various shape could be used. There are in fact four principal methods used to close cylindrical pressure vessels, these are: flat plates and formed flat heads, hemispherical heads, and torispherical and ellipsoidal heads.

Flat plates may be plain or flanged, and then bolted or welded in position (Figure 5.3.3). Torispherical domed heads, which are often used to close cylindrical pressure vessels are formed from part of a torus and part of a sphere (Figure 5.3.4). Torispherical heads are generally preferred to their elliptical counterparts (Figure 5.3.5), because they are easier and cheaper to fabricate. Hemispherical heads (Figure 5.3.6) are the strongest form of closure that can be used with cylindrical pressure vessels and are generally used for high-pressure applications. They are however costly, so their use is often restricted to applications where pressure containment and safety take priority over all other considerations, including costs.

Minimum thickness formulae for the methods of closure can be determined in a similar manner to those we found earlier for cylinders and spheres, by analysing their geometry, stresses, and method of constrain at their periphery. Time does not permit us to derive these relationships but three typical formulae for flat, ellipsoidal, and torispherical heads are given below.

The equation for determining the thickness of flat heads is:

$$e = c_p d_e \sqrt{\frac{p}{\sigma_d}} \quad (5.3.13)$$

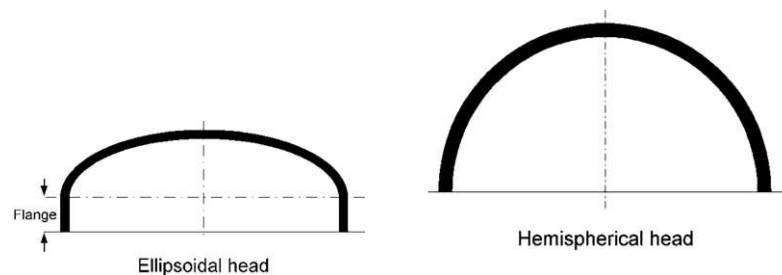


**Figure 5.3.5** Ellipsoidal head for cylindrical pressure vessel

**Figure 5.3.6** Hemispherical head for cylindrical pressure vessel

where  $c_p$  = a design constant, dependent on edge constraint,  
 $\sigma_d$  = design stress,  $d_e$  = nominal diameter.

The minimum thickness required for ellipsoidal heads is given by:

$$e = \frac{pd}{2\eta\sigma_d - 0.2p} \quad (5.3.14)$$

where the symbols have their usual meaning.

Finally the minimum thickness required for torispherical heads is given by:

$$e = \frac{pR_c C_s}{2\eta\sigma_d + p(C_s - 0.2)} \quad (5.3.15)$$

where  $C_s$  = stress concentration factor for torispherical heads and is found from:

$$C_s = \frac{1}{4} \left( 3 + \sqrt{\frac{R_c}{R_k}} \right)$$

where  $R_c$  = crown radius and  $R_k$  = knuckle radius.

We have spent some time looking at formulae which can be used in determining minimum thickness of materials for pressure vessels. For a full account of all aspects of pressure vessel design you should refer to BS 5500.

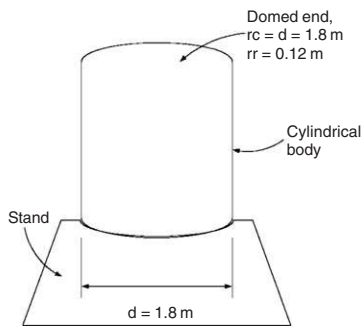


**Figure 5.3.7** Figure for Example 5.3.3

### Example 5.3.3

Consider the pressure vessel (Figure 5.3.7). Estimate the thickness of material required for the cylindrical section and domed ends. Assume that the pressure vessel is to be used in the food processing industry and so the material for construction is to be a stainless steel (SS). The vessel is to operate at a design temperature of 250°C and a design pressure of 1.6 MPa, so we will use a typical design stress of 115 MPa.

(i) *The cylindrical section*

Using Equation (5.3.7) we have:

$$e = \frac{1.6 \times 10^6 \times 1.8}{(2 \times 115 \times 10^6) - (1.6 \times 10^6)} = 0.0126 \text{ m.}$$

Now we need to consider any relevant safety factors. The design pressure will have a safety factor built-in, so it will already be set at a higher value than normal operating pressure. Therefore we need only consider an allowance for corrosion and fabrication problems. The SS is highly resistant to corrosion, but as a result of the fabrication process, may have 'locked-in' stresses which we are not aware of, therefore it would be prudent to make some allowance for the unknown without incurring too much unnecessary expense or material wastage. In practice, further detail design will determine the necessary allowances.

Let us assume a further 2 mm is the appropriate allowance, then:

$$e = 12.6 + 2 = 14.6 \text{ mm.}$$

Now we are unlikely to obtain 14.6 mm SS plate as stock, so we use 15 mm plate.

(ii) *The domed head*

We shall first consider a standard dished torispherical head and assume that the head is formed by pressing, therefore it has no welded joints and so  $\eta$  may be taken as 1.0. Let us assume that the torisphere has a crown radius equal to the diameter of the cylindrical section (its maximum value) and a knuckle radius of 0.12 m.

Then using Equation (5.3.15) we have for a torispherical head:

$$e = \frac{pR_c C_s}{2\sigma_d \eta + p(C_s - 0.2)}$$

$$\text{where } C_s = \frac{1}{4} \left( 3 + \sqrt{\frac{R_c}{R_k}} \right)$$

$$C_s = \frac{1}{4} \left( 3 + \sqrt{\frac{1.8}{0.12}} \right)$$

$$C_s = 1.72 \quad \text{and} \quad \eta = 1.0$$

$$\text{so } e = \frac{1.6 \times 10^6 \times 1.8 \times 1.72}{(2 \times 115 \times 10^6) + 1.6 \times 10^6(1.72 - 0.2)}$$

therefore,  $e = 0.0215 \text{ m}$  or 21.5 mm.

Let us now try an ellipsoidal head.

From Equation 5.3.14 we get:

$$e = \frac{1.6 \times 10^6 \times 1.8}{(2 \times 115 \times 10^6) - (0.2 \times 1.6 \times 10^6)}$$

$$e = 0.0125 \text{ m or } 12.5 \text{ mm.}$$

Since the fabrication of both the torispherical and ellipsoidal head incur similar effort, then the most economical head would be ellipsoidal with a minimum thickness of 15.0 mm, after the normal allowances are added.

## Thick-walled pressure vessels

When dealing with the theoretical analysis of thin cylinders we made the assumption that the hoop stress was constant across the thickness of the cylinder (Figure 5.3.2a) and that there was no difference in pressure across the cylinder wall. When dealing with thick pressure vessels where the wall thickness is substantial we can no longer accept either of the above two assumptions. So with thick-walled vessels, where the wall

thickness is normally greater than one tenth of the diameter, both the hoop ( $\sigma_h$ ) stress and the radial ( $\sigma_r$ ) stress vary across the wall.

If thick-walled vessels are subject to pressure, then the value of these stresses may be determined by use of the *Lamé equations* given below:

$$\sigma_r = a - \frac{b}{r^2} \quad (5.3.16)$$

$$\text{and } \sigma_h = a + \frac{b}{r^2}. \quad (5.3.17)$$

These equations may be used to determine the radial and hoop stress at any radius  $r$  in terms of constants  $a$  and  $b$ . For any pressure condition there will always be two known stress conditions that enable us to find values for the constants. These known stress conditions are often referred to as ‘boundary conditions’.

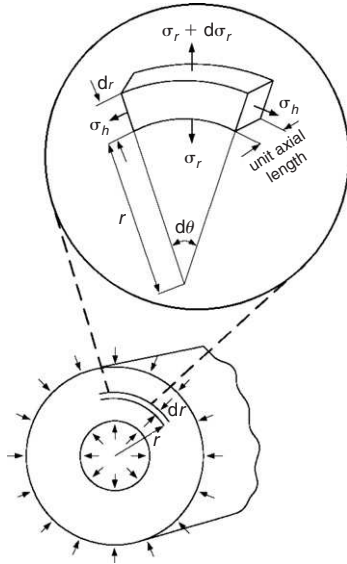


**Figure 5.3.8** Element of cylinder wall subject to internal and external pressure

## Mathematics in action

Consider a cylinder which is long in comparison with its diameter, where the longitudinal stress and strain are assumed to be constant across the thickness of the cylinder wall. If we take an element of the cylinder wall (Figure 5.3.8) when the cylinder is subject to an internal and external pressure, and where the axial length is considered to be unity.

Then for radial equilibrium, applying elementary trigonometry and remembering that for small angles  $\sin \theta$  approximately equals  $\theta$  (in radians), we may write:

$$(\sigma_r + d\sigma_r)(r + dr)d\theta = \sigma_r r d\theta + 2\sigma_h dr \frac{d\theta}{2}.$$

From which, after simplification, neglecting second-order small quantities we may write:

$$\sigma_r dr + r d\sigma_r = \sigma_h dr$$

which on division by  $(dr)$  gives:

$$\sigma_r + r \frac{d\sigma_r}{dr} = \sigma_h. \quad (5.3.18)$$

Now if we let the longitudinal stress be  $\sigma_L$  and the longitudinal strain be  $\epsilon_L$ , then following our assumption that the longitudinal stress and strain are constant across the cylinder walls we may write:

$$\epsilon_L = \frac{1}{E}(\sigma_L - \nu(\sigma_r + \sigma_h)). \quad (5.3.19)$$

Compare this equation with Equation (5.1.16).

Also, with  $\sigma_L$  and  $\epsilon_L$  constant then  $\sigma_r + \sigma_h = \text{constant}$  (say  $2a$ )

$$\text{or } \sigma_h = 2a - \sigma_r. \quad (5.3.20)$$

Then on substituting Equation (5.3.20) for  $\sigma_h$  into Equation 5.3.18 and simplifying, we get:

$$2ar - r^2 \frac{d\sigma_r}{dr} - 2\sigma_r r = 0$$

which on integration gives:

$$ar^2 - r^2\sigma_r = \text{constant (say } b)$$

and on rearrangement and division by  $r$  we get:

$$\sigma_r = a - \frac{b}{r^2} \quad (5.3.21)$$

and substituting Equation (5.3.21) for  $\sigma_r$  into Equation (5.3.20) gives:

$$\sigma_h = a + \frac{b}{r^2}$$

which are Equations (5.3.16) and (5.3.17), the Lamé equations for internal components.

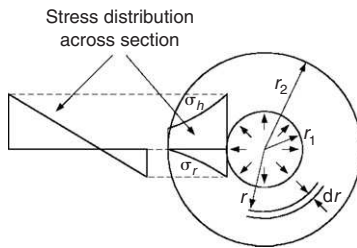


**Figure 5.3.9** Thick cylinder subject to internal pressure

### Thick cylinders subject to internal pressure

Consider a thick cylinder which is only subject to internal pressure ( $p$ ), the external pressure being zero. For the situation shown in Figure 5.3.9, the boundary conditions are,  $\sigma_r = -p$  when  $r = r_1$  and  $\sigma_r = 0$  when  $r = r_2$ . The internal pressure is considered as a negative radial stress, because it tends to produce thinning (compression) of the cylinder wall.

By substituting the boundary conditions into Equations (5.3.16) and (5.3.17), and determining expressions for the constants  $a$  and  $b$ , it can be shown that for thick cylinders subject to internal pressure only the:

$$\text{Radial stress } \sigma_r = \frac{pr_1^2}{r_2^2 - r_1^2} \left( \frac{r^2 - r_2^2}{r^2} \right) \quad (5.3.22)$$

and the

$$\text{Hoop stress } \sigma_h = \frac{pr_1^2}{r_2^2 - r_1^2} \left( \frac{r^2 + r_2^2}{r^2} \right). \quad (5.3.23)$$

The maximum values for the radial and hoop stress across the section can be seen in Figure 5.3.9 and may be determined using:

$$\sigma_{r_{\max}} = -p \quad (5.3.24)$$

$$\sigma_{h_{\max}} = p \left( \frac{r_1^2 + r_2^2}{r_2^2 - r_1^2} \right). \quad (5.3.25)$$

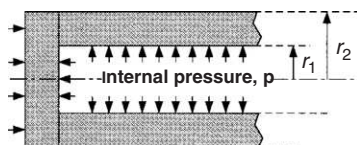


**Figure 5.3.10** Longitudinal equilibrium at end of cylinder

### Longitudinal and shear stress

So far we have established that the longitudinal stress in thick-walled pressure vessels, may be assumed to be uniform. This constant longitudinal stress ( $\sigma_l$ ) is determined by considering the longitudinal equilibrium at the end of the cylinder. Figure 5.3.10 shows the arrangement for the simple case where the cylinder is subject to internal pressure.

So assuming equilibrium conditions and resolving forces horizontally we have:

$$p \times \pi r_1^2 = \sigma_1 \times \pi(r_2^2 - r_1^2)$$

and

$$\sigma_1 = \frac{pr_1^2}{r_2^2 - r_1^2}. \quad (5.3.26)$$

In a similar manner it can be shown that for a thick-walled cylinder subject to both *internal and external pressure* the constant longitudinal stress is given by:

$$\sigma_1 = \frac{p_1 r_1^2 - p_2 r_2^2}{r_2^2 - r_1^2}. \quad (5.3.27)$$

The radial stress ( $\sigma_r$ ), hoop stress ( $\sigma_h$ ) and longitudinal stress ( $\sigma_1$ ) are all principal stresses. The maximum shear stress ( $\tau$ ) is equal to half the difference between the maximum and minimum principal stresses. So if we consider the case for *internal pressure only*, we have already seen that radial stress is compressive, the hoop and longitudinal stress are tensile. Therefore, remembering the sign convention, and noting the fact that the hoop stress is larger than the longitudinal stress. Then the maximum difference is given by ( $\sigma_h - \sigma_r$ ) and so the maximum shear stress is:

$$\tau_{\max} = \frac{\sigma_h - \sigma_r}{2}$$

and substituting for  $\sigma_h$  and  $\sigma_r$  from the Lamé equations we have:

$$\tau_{\max} = \frac{\left(\frac{a+b}{r^2}\right) - \left(\frac{a-b}{r^2}\right)}{2} = \frac{b}{r^2}.$$

Now applying the boundary conditions to the Lamé equation for radial stress where the cylinder is subject to internal pressure only gives:

$$-p = a - \frac{b}{r_1^2} \quad \text{and} \quad 0 = a - \frac{b}{r_2^2}$$

$$\text{from which } b = \frac{pr_1^2 r_2^2}{r_2^2 - r_1^2}$$

$$\text{and so } \tau_{\max} = \frac{pr_1^2 r_2^2}{(r_2^2 - r_1^2)r^2} = \frac{b}{r^2}.$$

So the maximum shear stress ( $\tau_{\max}$ ) occurs at the inside wall of the cylinder where  $r = r_1$

$$\text{or } \tau_{\max} = \frac{pr_2^2}{r_2^2 - r_1^2}. \quad (5.3.28)$$

**Example 5.3.4**

The cylinder of a hydraulic actuator has a bore of 100 mm and is required to operate up to a pressure of 12 MPa. Determine the required wall thickness for a limiting tensile stress of 36 MPa.

In this example the boundary conditions are that at  $r = 50$  mm,  $\sigma_r = -12$  MPa and since the maximum tensile hoop stress occurs at the inner surface then, at  $r = 50$  mm,  $\sigma_h = 36$  MPa. In order to simplify the arithmetic when calculating the constants in the Lamé equations we will use the relationship that  $1 \text{ MPa} = 1 \text{ N/mm}^2$  and work in  $\text{N/mm}^2$ .

Therefore using  $\sigma_h = a + b/r^2$ ,  $\sigma_r = a - b/r^2$  (inner surface) where  $\sigma_r = -12 \text{ N/mm}^2$ ,  $r = 50 \text{ mm}$ ,  $\sigma_h = 36 \text{ N/mm}^2$

$$\text{we have } -12 = a - \frac{b}{2500}$$

$$\text{and } 36 = a + \frac{b}{2500}.$$

Adding the two equations:

$$24 = 2a, \text{ so } a = 12$$

from which on substitution

$$-12 = 12 - \frac{b}{2500} \quad \text{and } b = 60\,000.$$

At the outer surface,  $\sigma_r = 0$ ; therefore:

$$0 = a - \frac{b}{r^2} = 12 - \frac{60\,000}{r^2}$$

$$\text{and } r = 70.7 \text{ mm}.$$

**Questions 5.3.1**

- (1) A thin cylinder has an internal diameter of 50 mm, a wall thickness of 3 mm and is 200 mm long. If the cylinder is subject to an internal pressure of 10 MPa, determine the change in internal diameter and the change in length. Take  $E = 200 \text{ GPa}$  and Poisson's ratio  $\nu = 0.3$ .
- (2) A cylindrical reservoir used to store compressed gas is 1.8 m in diameter and assembled from plates 12.5 mm thick. The longitudinal joints have an efficiency of 90% and the circumferential joints have an efficiency of 50%. If the reservoir plating is to be limited to a tensile stress of 110 MPa, determine the maximum safe pressure of the gas.
- (3) A thin cylinder has an internal diameter of 240 mm, a wall thickness of 1.5 mm and is 1.2 m long. The cylinder material has  $E = 210 \text{ GPa}$  and Poisson's ratio  $\nu = 0.3$ . If

the internal volume change is found to be  $14 \times 10^{-6} \text{ m}^3$  when pressurised.

Find

- (i) the value of this pressure;
  - (ii) the value of the hoop and longitudinal stresses.
- (4) A steel pipe of internal diameter 50 mm and external diameter 90 mm is subject to an internal pressure of 15 MPa. Determine the radial and hoop stress at the inner surface.
- (5) A thick cylinder has an internal diameter 200 mm and external diameter of 300 mm. If the cylinder is subject to an internal pressure of 70 MPa and an external pressure of 40 MPa.
- Determine:
- (i) the hoop and radial stress at the inner and outer surfaces of the cylinder;
  - (ii) the longitudinal stress, if the cylinder is assumed to have closed ends.

## 5.4 DYNAMICS OF ROTATING SYSTEMS

In this final section we start with an investigation into power transmission between machines using belt drives, clutches, and gear trains. We then turn our attention to the dynamics of rotating systems, looking particularly at static and dynamic balancing, flywheel dynamics and finally at the coupling of these components into rotational systems.

So, for example, the torque created by an electric motor may be converted into mechanical work by coupling the motor output to a gearbox, chain, or belt drive. Motor vehicle transmission systems convert the torque from the engine into the torque at the wheels via a clutch, gearbox, and drive shafts. The clutch is used to engage and disengage the drive from the engine, the gearbox assembly transmits the power to the wheel drive shafts, as dictated by the driving conditions.

Gears, shafts, and bearings provide relatively rigid, but positive power transmission and may be used in many diverse applications ranging, for example, from power transmission for the aircraft landing flap system on a jumbo jet to the drive mechanism for a child's toy.

Belts, ropes, chains, and other similar elastic or flexible machine elements are used in conveying systems and in the transmission of power over relatively long distances. In addition, since these elements are elastic and usually quite long, they play an important part in absorbing shock loads and in damping out and isolating the effects of unwanted vibration.

A classic example of the effects of a component being out of balance may be felt through the steering wheel of a motor vehicle. Where a minor knock to a road wheel may be sufficient to put it out of balance and so cause rotational vibration which is felt back through the steering geometry. Any rotational machinery, which is incorrectly balanced or goes out of balance as a result of damage or a fault, will produce vibration, which is at best unacceptable and at worse may cause dangerous failure. In this section, you will be briefly introduced to the concept of balance and the techniques for determining out of balance moments.

We finish this section with a brief study of flywheels. These components provide a valuable source of rotational energy, for use in many



engineering situations, such as machine presses, mills, cutting machines, and prime movers for transport and agricultural machinery. We look, finally, at the inertia and energies involved when transmission and rotational components are coupled together, to form rotational systems.

We begin our study of power transmission by looking first at belt drives.

## Belt drives

Power transmission from one drive shaft to another can be achieved by using drive pulleys attached to the shafts of machines being connected by belts of varying cross-section and materials.

A wide variety of belts are available, including flat, round, V, toothed, and notched, the simplest of these and one of the most efficient is the flat belt. Belts may be made from leather or a variety of reinforced elastomer materials. Flat belts are used with crowned pulleys, round and V-belts are used with grooved pulleys, while timing belts require toothed wheels or sprockets.

To illustrate the set-up between a drive belt and pulleys a V-belt system is shown in Figure 5.4.1. Note that the pulley size is given by its pitch diameter and that in the case of a V-belt, it rides slightly higher than the external diameter of the pulley. The meaning of the symbols used in Figure 5.4.1 (if they are not obvious) are explained in the following paragraphs.

The speed ratio between the driving and driven pulleys is inversely proportional to the ratio of the pulley pitch diameters. This follows from

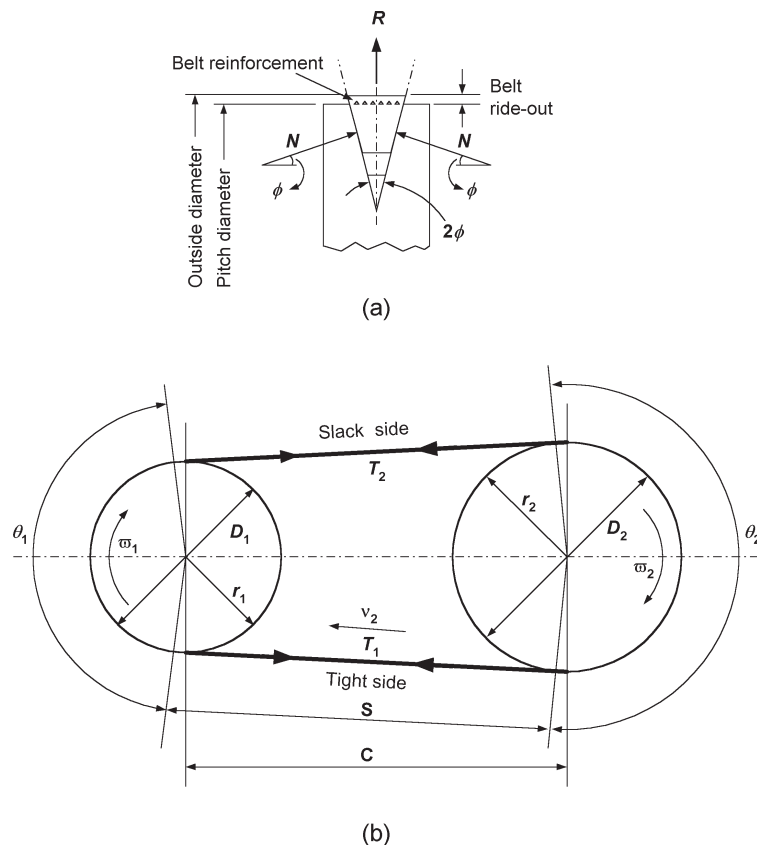


**Figure 5.4.1** Basic belt drive geometry

the observation that there is no slipping (under normal loads). Thus the linear velocity of the pitch line of both pulleys is the same and equal to the belt velocity ( $\nu$ ), then:

$$\nu = r_1\omega_1 = r_2\omega_2 \quad (5.4.1)$$

$$\text{or } \nu = \frac{D_1\omega_1}{2} = \frac{D_2\omega_2}{2}$$

$$\text{so the angular velocity ratio (VR) is } \frac{\omega_1}{\omega_2} = \frac{D_2}{D_1} \quad (5.4.2)$$

The angles  $\theta_1$  and  $\theta_2$  are known as the *angle of lap* or *angle of contact* (Figure 5.4.1), this is measured between the contact length of the belt with the pulley and the angle subtended at the centre of the pulley.

The *torque* in belt drives and subsequent *transmitted power* is influenced by the differences in tension that exist within the belt during operation. The differences in tension result in a tight side and slack side of the belt, as shown. These differences in tension influence the torque on the driving and driven pulley wheels, in the following way:

$$\text{Torque on driving pulley} = T_1r_1 - T_2r_2 = (T_1 - T_2)r_1$$

$$\text{Torque on driven pulley} = T_1r_2 - T_2r_2 = (T_1 - T_2)r_2.$$

Now we know that power = torque  $\times$  angular velocity, and since angular velocity is given by the linear speed divided by the radius of action, then from Equation (5.4.1) we may write that:

$$\text{Transmitted power (P)} = (T_1 - T_2)V. \quad (5.4.3)$$

The *stresses* set up in the drive belt are due to:

- (i) the tensile force in the belt, which is a maximum on the tight side of the belt;
- (ii) the bending of the belt around the pulleys, which is a maximum as the tight side of the belt bends around the smaller diameter pulley;
- (iii) the centrifugal forces created as the belt moves around the pulleys.

The maximum total stress occurs where the belt enters the smaller pulley and the bending stress is a major part. Thus there are recommended minimum pulley diameters for standard belts. Using smaller than recommended drastically reduces belt life.

### **Belt tension equations**

Belts when assembled on pulleys are given an initial tension ( $T_0$ ), you will be only too well aware of this fact if you have ever been involved in setting up the fan belt on a motor vehicle engine. If this initial tension is too slack the belt may slip below the maximum design power or, if too tight, undue stresses may be imposed on the belt causing premature failure. The initial pre-tension may be determined by using the formula:

$$T_0 = \frac{T_1 + T_2}{2}. \quad (5.4.4)$$

This follows from the argument that since the belt is made from an elastic material, when travelling around the pulleys, the belt length remains

constant. Thus the increase in length on the tight side is equal to the reduction in length on the slack side, and so the corresponding tensions are given by,  $T_1 - T_0 = T_0 - T_2$ , from which Equation (5.4.4) results.

As already mentioned, when the belt passes over a pulley, there is an increase in the stresses imposed within the belt. These are due to an increase in centrifugal force, as the belt passes over the pulley, creating an increase in the tension of the belt.

Therefore, we have an increase in belt tension, additional to that of normal transmission power, which is due to the centrifugal forces set up in the belt as it passes over the pulleys. Our standard power Equation (5.4.3) needs to be modified to take the effects of centrifugal tension into account.

We can formulate this relationship, leading to the required power equation if we are prepared to undertake a little mathematical gymnastics! This exercise has been covered in Chapter 6, under Trigonometric Methods. The important results of this exercise *and* the relationship between the coefficient of friction ( $\mu$ ) the angle of lap ( $\theta$ ), and the belt tensions ( $T_1, T_2$ ) are given below:

$$T_c = m^1 \nu^2 \quad (5.4.5)$$

where  $T_c$  is the additional centrifugal tension,  $m^1$ , the mass per unit length of the belt, and  $\nu$ , the belt velocity as indicated in Figure 5.4.1.

$$\frac{T_1 - T_c}{T_2 - T_c} = e^{\mu\theta}. \quad (5.4.6)$$

Under low-speed condition the centrifugal tension term in Equation (5.4.6) is relatively small and may be neglected to give the following relationship:

$$\frac{T_1}{T_2} = e^{\mu\theta}. \quad (5.4.7)$$

Also, as suggested we may modify our *power* equation to take into account the centripetal tension term, under high- or low-speed conditions, then substituting these terms into Equation (5.4.3) gives:

$$P = (T_1 - T_c)(1 - e^{-\mu\theta})\nu \quad (5.4.8)$$

$$\text{or, for low-speed conditions } P = T_1(1 - e^{-\mu\theta})\nu. \quad (5.4.9)$$

### Modified equation for V-belt drives

The above equations have been formulated based on geometry related to flat belts travelling over flat drum pulleys (see Chapter 6). V-belt drives require us to replace  $\mu$  by  $\mu/\sin\phi$ , where it appears in Equations 5.4.6 to 5.4.9.

Consider the V-pulley drive shown in Figure 5.4.1, where the groove angle is  $(2\phi)$  and the reactions ( $N$ ) act at angle  $\phi$  as shown. The normal reaction ( $R$ ) between the belt and pulley is given by simple trigonometry as:

$$R = 2N \sin \phi$$

and so the *friction force* between the belt and pulley is:

$$2\mu N = \frac{\mu R}{\sin \phi}. \quad (5.4.10)$$

So in effect, going from flat to V-grooved belt drives, transforms  $\mu$  to  $\mu/\sin \phi$ .

### ***Belt drive design requirements***

In order to select a drive belt and its associated pulley assemblies for a specified use, and to ensure proper installation of the drive, several important factors need to be considered. The basic data required for drive selection is listed below. They are the:

- rated power of the driving motor or other prime mover;
- the service factor, which is dependent on the time in use, operating environment and the type of duty performed by the drive system;
- belt power rating;
- belt length;
- size of driving and driven pulleys and centre distance between pulleys;
- correction factors for belt length and lap angle on smaller pulley;
- number of belts;
- initial tension on belt.

Many design decisions depend on the required usage and space constraints. Provision must be made for belt adjustment and pre-tensioning, so the distance between pulley centres must be adjustable. The lap angle on the smaller pulley should be greater than  $120^\circ$ . Shaft centres should be parallel and carefully aligned so that the belts track smoothly, particularly if grooved belts are used.

The operating environment should be taken into consideration, where hostile chemicals, pollution, or elevated temperatures may have an effect on the belt material. Consider other forms of drive, such as chains, if the operating environment presents difficulties.

From the above list, it can be seen that relationships between belt length, distance between centres, pulley diameters, and lap angles are required if the correct belt drive is to be selected. Equation (5.4.11) relates the belt pitch length,  $L$ , centre distance,  $C$ , and the pulley pitch diameters,  $D$ , by:

$$L = 2C + 1.57(D_2 - D_1) + \frac{(D_2 - D_1)^2}{4C}. \quad (5.4.11)$$

$$\text{Also } C = B + \frac{\sqrt{B^2 - 32(D_2 - D_1)^2}}{16} \quad (5.4.12)$$

where  $B = 4L - 6.28(D_2 + D_1)$ .

The contact angle (*lap angle*) of the belt on each pulley is:

$$\begin{aligned} \theta_1 &= 180^\circ - 2 \sin^{-1} \left[ \frac{D_2 - D_1}{2C} \right] \\ \theta_2 &= 180^\circ + 2 \sin^{-1} \left[ \frac{D_2 - D_1}{2C} \right] \end{aligned} \quad (5.4.13)$$

One other useful formula is given below which enables us to find the span length  $S$  (see Figure 5.4.1), given the distance  $C$  between pulley centres and the pulley diameters.

$$S = \sqrt{C^2 - \left[ \frac{D_2 - D_1}{2} \right]^2}. \quad (5.4.14)$$

This relationship is important since it enables us to check for correct belt tension, by measuring the force required to deflect the belt at mid-span. This length also has a direct bearing on the amount of vibration set up in the belt during operation therefore, we need to ensure that span length falls within required design limits.

### Example 5.4.1

A flat belt drive system consists of two parallel pulleys of diameter 200 and 300 mm, which have a distance between centres of 500 mm. Given that the maximum belt tension is not to exceed 1.2 kN, the coefficient of friction between the belt and pulley is 0.4 and the larger pulley rotates at 30 rads/sec. Find:

- (i) the belt lap angle for the pulleys;
- (ii) the power transmitted by the system;
- (iii) the belt pitch length  $L$ ;
- (iv) the pulley system span length between centres.

(i) Lap angles are given by Equation (5.4.13).

so

$$\begin{aligned} \theta_1 &= 180^\circ - 2 \sin^{-1} \left( \frac{300 - 200}{2 \times 500} \right) \\ &= 180^\circ - 2 \sin^{-1} 0.1 \\ &= 180^\circ - 11.48^\circ = 168.52^\circ \end{aligned}$$

$$\text{and } \theta_2 = 180^\circ + 11.48^\circ = 191.48^\circ.$$

(ii) The maximum power transmitted will depend on the angle of lap of the smaller pulley. So, using Equation (5.4.9) and assuming low-speed conditions, we get:

$$P = T_1(1 - e^{-\mu\theta})\nu \quad \text{where } \nu = r\omega = (0.15)(30) = 4.5.$$

$$P = 1.2 \times 10^3(1 - e^{-(0.4)(2.94)})4.5 \quad \text{note } \theta \text{ in radians.}$$

$$P = 1.2 \times 10^3(1 - 0.308)4.5.$$

$$\mathbf{P = 3.74 \text{ kW.}}$$

(iii) The belt pitch length is found using Equation (5.4.11)

$$L = 2C + 1.57(D_2 - D_1) + \frac{(D_2 - D_1)^2}{4C}.$$

The distance between pulley centres  $C$  is given as 500 mm. Then:

$$L = (2)(0.5) + 1.57(0.3 - 0.2) + \frac{(0.3 - 0.2)^2}{(4)(0.5)}$$

$$\text{giving } \mathbf{L = 1.162 \text{ m.}}$$

(iv) The span length  $S$  is given by Equation (5.4.14). Then:

$$S = \sqrt{(0.5)^2 - \left(\frac{0.3 - 0.2}{2}\right)^2}$$

giving  $S = 0.495 \text{ m}$ .

## Friction clutches

With the sliding surfaces encountered in most machine components, such as bearings, gears, levers, and cams it is desirable to minimise friction, so reducing energy loss, heat generation, and wear. When considering friction in clutches and brakes, we try to maximise the friction between the driving and driven plates. This is achieved in dry clutches by choosing clutch plate materials with high coefficients of friction. Much research has been undertaken in materials designed to maximise the friction coefficient and minimise clutch disc wear, throughout a wide range of operating conditions during its service life.

The primary function of a clutch is to permit smooth transition between connection and disconnection of the driving and driven components and, when engaged, to ensure that the maximum possible amount of available torque is transmitted. Clutches are used where there is a constant rotational torque produced at the drive shaft, by a prime mover such as an engine. The drive shaft can then be engaged through the clutch to drive any load at the output, without the need to alter the power produced by the prime mover. The classic application of this transmission principle is readily seen in automobiles, where the engine provides the power at the drive shaft which is converted to torque at the wheels. When we wish to stop the vehicle, without switching off the engine, we disengage the drive using the clutch.

### Disc clutches

A simple disc clutch, with one driving and one driven disc is illustrated in Figure 5.4.2. The principle of operation is dependent on the driving friction created between the two clutch discs when they are forced together. The means of applying pressure to the clutch discs may be mechanical or hydraulic.

As with other components that rely on friction, they are designed to operate either dry or wet using oil. Many types of clutch used in automobile

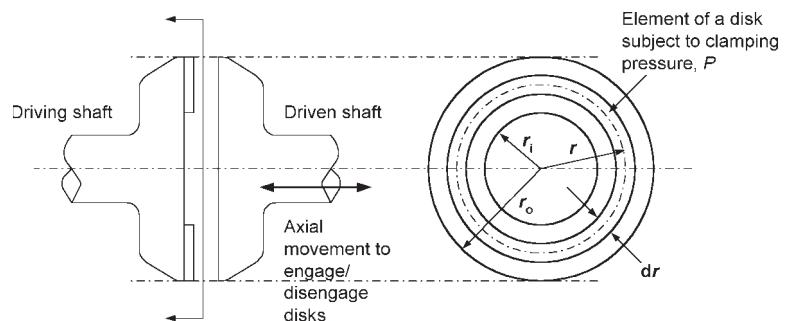


**Figure 5.4.2** Simple disc clutch

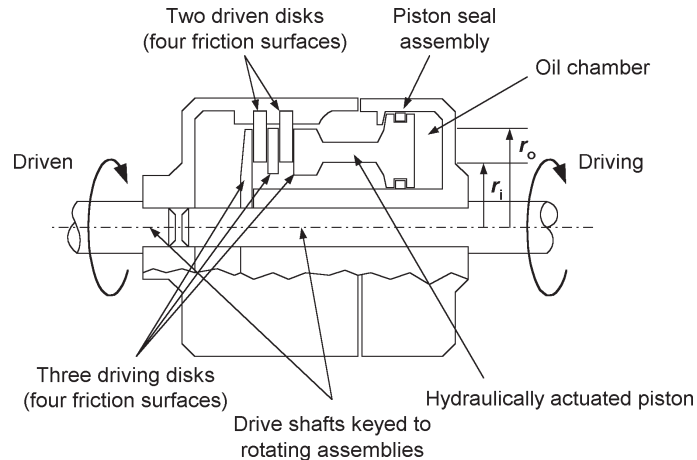


**Figure 5.4.3** *Hydraulically operated clutch assembly*

transmissions are designed to operate wet. The oil provides an effective cooling medium for the heat generated during clutch engagement and instead of using a single disc (Figure 5.4.2.), multiple discs may be used to compensate for the inevitable drop in the coefficient of friction.

In a multi-disc clutch (Figure 5.4.3) the driving discs rotate with the input shaft, the clutch is engaged using hydraulic pressure which is generated from a master cylinder at the clutch pedal, or by some other means. The driving discs are constrained, normally by a key and keyway, to rotate with the input shaft. On engagement, the rotating drive discs are forced together with the driven disc to provide torque at the output. When the clutch is disengaged by releasing the hydraulic pressure, or through spring action, the discs are free to slide axially and so separate themselves.

From our diagram you can see that the driving discs include the backing plates, which only provide one friction surface, whereas the middle driving disc and the two driven discs each provide two contact surfaces. So in our example the five discs provide a total of four driving and four driven contact surfaces.

### **Friction clutch equations**

In order to estimate the torque and power from clutch driven shafts we need to use equations which relate factors such as: clutch size, friction coefficients, torque capacity, axial clamping force, and pressure at the discs. When formulating such equations, we make two very important basic assumptions:

- (i) we assume uniform distribution of pressure at the disc interface;
- (ii) we assume a uniform rate of wear at the disc interface.

## **Mathematics in action**

The *first assumption* is valid for an unworn (new) accurately produced and installed clutch, with rigid outer discs. Consider a very small element of the disc surface (Figure 5.4.2), which is subject to an axial force normal to the friction surface. Then if *constant* pressure  $p$  acts

on the element at radius  $r$  and the element has width  $dr$ , then the force on the element:

$$= (2\pi r dr)p.$$

So the total normal force acting on the area of contact is:

$$W = \int_{r_i}^{r_o} 2\pi p r dr = \pi p (r_o^2 - r_i^2) \quad (5.4.15)$$

where  $W$  is the axial force damping the driving and driven plates together.

Now we know that the friction torque that can be developed on the disc 'element' is the product of the friction force and radius. The friction force is the product of the force on the element and the coefficient of friction. So we may write:

$$\text{Torque on disc element} = (2\pi p r dr)\mu r$$

and the total torque that can be transmitted is:

$$T = \int_{r_i}^{r_o} 2\pi p \mu r^2 dr = \frac{2}{3} \pi p \mu (r_o^3 - r_i^3). \quad (5.4.16)$$

The above equation represents the torque transmitted by one driving disc (plate) mating with one driven plate.

For clutches that have more than one set of discs, that is,  $N$  discs, where  $N$  is an even number (see Figure 5.4.3), then:

$$T = \frac{2}{3} \pi p \mu (r_o^3 - r_i^3) N. \quad (5.4.17)$$

If we transpose Equation (5.4.15) for  $p$  and substitute it into Equation (5.4.17) we get:

$$T = \frac{2W\mu(r_o^3 - r_i^3)}{3(r_o^2 - r_i^2)} N. \quad (5.4.18)$$

This equation relates the torque capacity of the clutch with the axial clamping force.

If we now assume a uniform rate of wear. We know that the rate at which wear takes place between two rubbing surfaces is proportional to the rate at which it is rubbed (the rubbing velocity) multiplied by the pressure applied. So as a pair of clutch plates slide over each other, their relative velocities are linear, our clutch plates rotate with angular velocity, so the rubbing velocity is dependent on the radius of rotation. From the above argument we may write:

$$\text{wear} \propto \text{pressure} \times \text{velocity (where velocity is dependent on radius)}$$

$$\text{so } \text{wear} \propto \text{pressure} \times \text{radius}$$

and introducing the constant of proportionality gives:

$$pr = c.$$

Now substituting the above expression into Equations 5.4.15 and 5.4.16 give, respectively:

$$W = \int_{r_i}^{r_o} 2\pi c dr = 2\pi c (r_o - r_i) \quad (5.4.19)$$



and

$$T = \int_{r_i}^{r_o} 2\pi c \mu r dr = \pi c \mu (r_o^2 - r_i^2) N \quad (5.4.20)$$

and on transposition of Equation (5.4.15) for  $p$  and then substitution into Equation (5.4.20) we get:

$$T = W \mu \left( \frac{r_o + r_i}{2} \right) N. \quad (5.4.21)$$

We have said that  $pr = c$ , now it is known that the greatest pressure on the clutch plate friction surface occurs at the inside radius  $r_i$ . So for a maximum allowable pressure ( $p_{\max}$ ) for the friction surface, we may write:

$$pr = c = p_{\max} r_i. \quad (5.4.22)$$

Clutch plates are designed with a specified ratio between their inside and outside radius, to maximise the transmitted torque between surfaces. This is given below, without proof:

$$r_i = 0.58r_o.$$

In practice clutches are designed within the approximate range  $0.5r_o \leq r_i \leq 0.8r_o$ .

### Cone clutches

A cone clutch has a *single pair* of conical friction surfaces (Figure 5.4.4) and is similar to a disc clutch, except that in the case of the disc clutch the coning angle take a specific value of  $90^\circ$ . So the cone clutch may be thought of as the generalised case, of which the disc clutch is a special case. One of the advantages of a cone clutch, apart from their simplicity, is that due to the wedging action the clamping force required is drastically reduced when compared to their disc clutch counterpart.

The formulae for cone clutches are derived in a very similar way to those you met earlier concerning the plate clutch and V-belts. We therefore adopt a very similar mathematical approach.

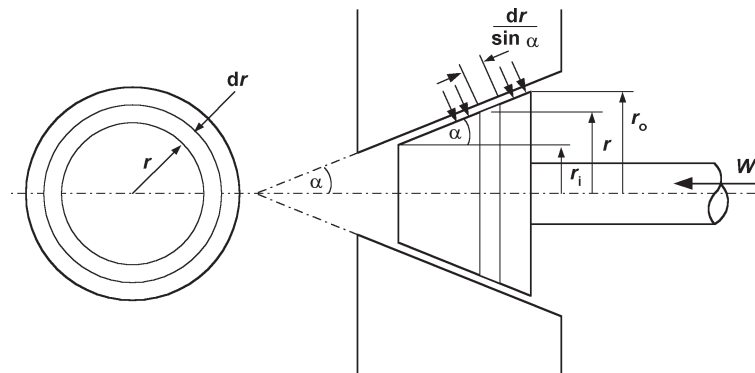


Figure 5.4.4 Cone clutch

## Mathematics in action

### Conical clutches

Consider the element of width  $dr$  in Figure 5.4.4, from the geometry the area of this element is:

$$= \frac{2\pi r dr}{\sin \alpha}.$$

If  $p$  is again the normal pressure, then the normal force acting on the element is:

$$= \frac{(2\pi r dr)p}{\sin \alpha}$$

and the corresponding axial clamping force is:

$$= (2\pi r dr)p$$

the above expression is exactly the same as for the disc clutch element, shown earlier. So the total axial force acting on the mating cone surfaces is:

$$W = \int_{r_i}^{r_o} 2\pi p r dr \text{ (same as Equation (5.4.15))}$$

and assuming 'uniform pressure'

$$W = \pi p (r_o^2 - r_i^2) \text{ as before.}$$

We also know that the friction force is equal to the product of the coefficient of friction ( $\mu$ ) and the normal reaction force. If we again consider the normal reaction force in terms of pressure, we may write:

$$\text{Friction force} = \frac{(2\pi r \mu dr)p}{\sin \alpha}$$

and the torque transmitted by the element is:

$$= \frac{(2\pi r \mu dr)pr}{\sin \alpha}.$$

Then the total torque transmitted is:

$$T = \int_{r_i}^{r_o} 2\pi \mu p r^2 dr$$

and assuming uniform pressure

$$T = \frac{2\pi \mu p}{3 \sin \alpha} (r_o^3 - r_i^3) \quad (5.4.23)$$

and as before eliminating  $p$ , using the force ( $W$ ) equation, we have:

$$T = \frac{2\mu W}{3 \sin \alpha} \left( \frac{r_o^3 - r_i^3}{r_o^2 - r_i^2} \right). \quad (5.4.24)$$

Finally assuming *uniform wear* we have:

$$W = 2\pi p_{\max} r_i (r_o - r_i) \quad (5.4.25)$$

$$T = \frac{\pi \mu p_{\max} r_i}{\sin \alpha} (r_o^2 - r_i^2) \quad (5.4.26)$$

and in terms of  $W$

$$T = \frac{\mu W}{2 \sin \alpha} (r_o + r_i). \quad (5.4.27)$$

Note that in Equations (5.4.25) and (5.4.26), the constant 'c' is replaced by  $p_{\max} r_i$ , since  $pr = c = p_{\max} r_i$  (compare this with Equations (5.3.18) and (5.3.19)).

### Example 5.4.2

A multiple plate clutch, similar to that shown in Figure 5.4.3, needs to be able to transmit a torque of 160 Nm. The external and internal diameters of the friction plates are 100 and 60 mm, respectively. If the friction coefficient between rubbing surfaces is 0.25 and  $p_{\max} = 1.2$  MPa determine the total number of discs required, *and* the axial clamping force. State all assumptions made.

Since this is a design problem, we make the justifiable assumption that the wear rate will be uniform at the clutch plate rubbing surfaces. We must also assume that the torque is shared equally by all the clutch discs and, that the friction coefficient remains constant during operation.

This last assumption is unlikely, and contingencies in the design would need to be made to ensure the clutch met its service requirements. In practice, designers ensure that a new clutch is slightly over engineered to allow for *bedding in*. After the initial bedding-in period wear tends to be uniform and friction coefficients are able to meet the clutch design specification under varying conditions, so the clutch assembly performs satisfactorily.

Using Equation 5.4.20, where  $c = r_i p_{\max}$ , then:

$$N = \frac{T}{\pi r_i p_{\max} \mu (r_o^2 - r_i^2)}.$$

$$\text{so } N = \frac{160}{(\pi)(30 \times 10^{-3})(1.2 \times 10^6)(0.25)(0.05^2 - 0.03^2)}$$

$$N = \frac{160}{45.29} = 3.53.$$

Now since  $N$  must be an even integer, then the nearest  $N$  above design requirements is  $N = 4$ . This is because

the friction interfaces must transmit torque in parallel pairs. In Figure 5.4.3 there are four friction interfaces, which require 'five' discs, the two outer discs having only one friction surface.

We also need to find the axial clamping force ( $W$ ).

This is given by Equation (5.4.21), where:

$$T = W\mu\left(\frac{r_o + r_i}{2}\right)N.$$

$$\text{so } W = \frac{T}{\mu N} \left( \frac{2}{r_o - r_i} \right) = \frac{160}{(0.25)(4)} \left( \frac{2}{0.05 - 0.03} \right)$$

$$W = 16 \text{ kN.}$$

So we require a total of five discs and an axial clamping force of 16 kN.

## Gear trains

We continue our study of power transmission by looking at gears. The gearbox is designed to transfer and modify rotational motion and torque. For example, electric motor output shafts may rotate at relatively high rates, producing relatively low torque. A reduction gearbox interposed between the motor output and load, apart from reducing the speed to that required, would also increase the torque at the gearbox output by an amount equivalent to the gearbox ratio.

The gear wheel itself is a toothed device designed also, to transmit rotary motion from one shaft to another. As mentioned in the introduction to transmission systems, gears are rugged, durable, and efficient transmission devices, providing rigid drives. The simplest and most commonly used gear is the *spur gear*, they are designed to transmit motion between parallel shafts (Figure 5.4.5). Spur gears can be helical, providing smooth and efficient meshing, which provides a quieter action and tends to reduce wear. The gears used in automobile transmission systems are often helical, for the above reasons (synchronised meshing or synchromesh). If the gear teeth are bevelled, they enable the gears to transmit motion at right angles, the common hand drill is a good example of the use of bevelled gears.

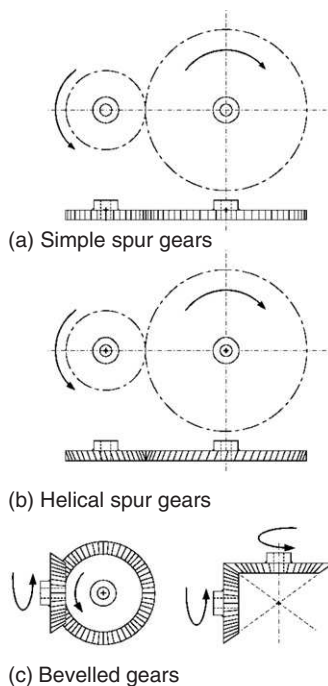


**Figure 5.4.5** Gear wheels

### Spur gears

Figure 5.4.6 shows the basic geometry of two meshing spur gears, there are several important points to note about this geometry, which will help you when we consider gear trains. The gear teeth profile is *involute*, this geometric profile is the curve generated by any point on a taut thread as it unwinds from a *base circle*. Once the gear teeth have cut to this involute shape they will, if correctly spaced, mesh without jamming. It is interesting to note that this is the only shape where this can be achieved. The gear teeth involute profile is extended outwards beyond the *pitch circle* (the point of contact of inter-meshing teeth) by a distance called the addendum. Similarly the tooth profiles are extended inwards from the pitch circle by an identical distance called the dedendum. When we refer to the diameter of gear wheels, we are always referring to the *pitch diameter*.

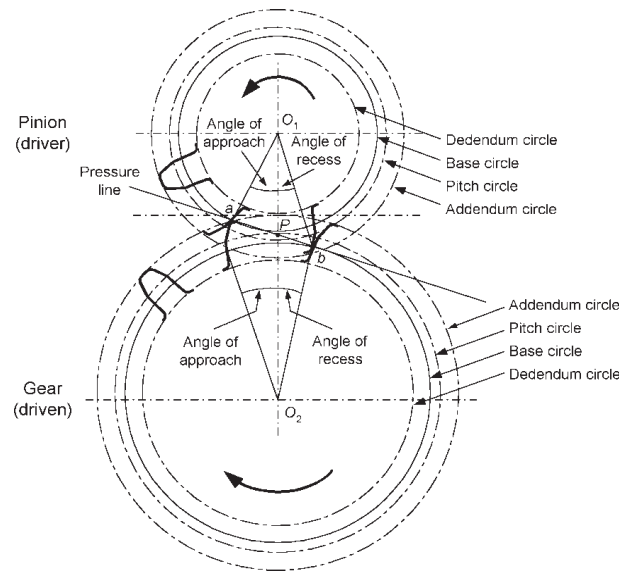


**Figure 5.4.6** Geometry for two gears in mesh

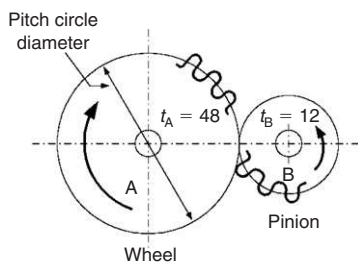


**Figure 5.4.7** Simple gear train

### Simple gear train

Meshing spur gears which transmit rotational motion from an input shaft to an output shaft are referred to as *gear trains*. Figure 5.4.7 shows a simple gear train consisting of just two spur gears of differing size, with their corresponding direction of rotation being given. If the larger gear *wheel* has 48 teeth (gear A) and the smaller *pinion* has 12 teeth (gear B). Then the smaller gear will have to complete 4 revolutions for each revolution of the larger gear. Now we know that the VR is defined as:

$$VR = \frac{\text{Distance moved by effort}}{\text{Distance moved by load}}$$

The VR when related to our gear train, will depend on which of the gears is the *driver* (effort) and which the *driven* (load). The distance moved is dependent on the number of teeth on the gear wheel. In our example above, if the larger wheel is the driver then we have:

$$VR = \frac{\text{Distance moved by driver}}{\text{Distance moved by driven}} = \frac{1}{4}$$

What this shows is the relationship between the number of gear teeth and the VR or, when considering gears, the *gear ratio*.

Now when the gear train is in motion, the ratio of the angular velocities of the gears is directly proportional to the angular distances travelled and as can be seen from above is the inverse ratio of the number of teeth on the gears. Thus the *gear ratio*  $G$  may be written formally as:

$$\text{Gear ratio } (G) \frac{t_B}{t_A} = \frac{\omega_A}{\omega_B} \text{ which for our gear pair gives } \frac{12}{48} = \frac{1}{4}$$

So the gear ratio in this case is 1:4.

The gear teeth on any meshing pair must be the same size, therefore the number of teeth on a gear is directly proportional to its *pitch circle*

diameter. So, taking into account this fact the *gear ratio formula*, may be written as:

$$G = \frac{\omega_A}{\omega_B} = \frac{t_B}{t_A} = \frac{d_B}{d_A}. \quad (5.4.28)$$

*So to summarise; the gear ratio is inversely proportional to the teeth ratio, and so inversely proportional to the pitch circle diameter ratio.*

### Example 5.4.3

A gearwheel (A) rotates with an angular velocity of 32 rad/s, it has 24 teeth, it drives a gear B. If the gear ratio  $G$  is 8:1, determine the velocity of B and the number of teeth on B.

From Equation (5.4.28)

$$G = \frac{\omega_A}{\omega_B} = \frac{t_B}{t_A} \text{ or } \frac{8}{1} = \frac{32}{4} = \frac{t_B}{24} \text{ and } t_B = 192.$$

Then the velocity of gear B = 4 rad/s and it has 192 teeth.

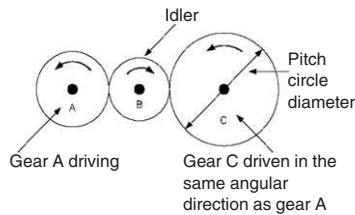


**Figure 5.4.8** Gear train with idler

A simple gear train can, of course, carry more than two gears. When more than two gears are involved we still consider the gear train in terms of a combination of gear pairs. Figure 5.4.8 shows a simple gear train consisting of three gears.

The overall gear ratio is the ratio of the input velocity over the output velocity, in our example this is,  $\omega_A/\omega_C$ . By considering the VRs of each gear pair sequentially from the input, we can determine the VR for the whole train. So for the first pair of meshing gears, AB, the VR is given in the normal manner as  $\omega_A/\omega_B$ . Now the second pair of meshing gears is  $\omega_B/\omega_C$ . So the overall gear ratio is:

$$G = \frac{\omega_A}{\omega_B} \times \frac{\omega_B}{\omega_C} = \frac{\omega_A}{\omega_C}. \quad (5.4.29)$$

You will note that the middle gear, plays no part in the final gear ratio. This suggests that it is not really required. It is true that the size of the middle gear has no bearing on the final gear ratio, but you will note from Figure 5.4.8, that it does have the effect of reversing the direction of angular motion between the driving and driven gear, for this reason it is known as an *idler gear*.

### Compound gear train

In a compound train at least one shaft carries a compound gear; that is, two wheels which rotate at the same angular velocity. The advantage of this type of arrangement is that compound gear trains can produce higher gear ratios, without using the larger gear sizes that would be needed in a simple gear train. Examples of compound gear trains may be found in automobile gearboxes and in a variety of industrial machinery such as lathes and drilling machines.

The following example illustrates how the gear ratio may be determined by considering the angular velocities of the gears and their relationship with the number of teeth.

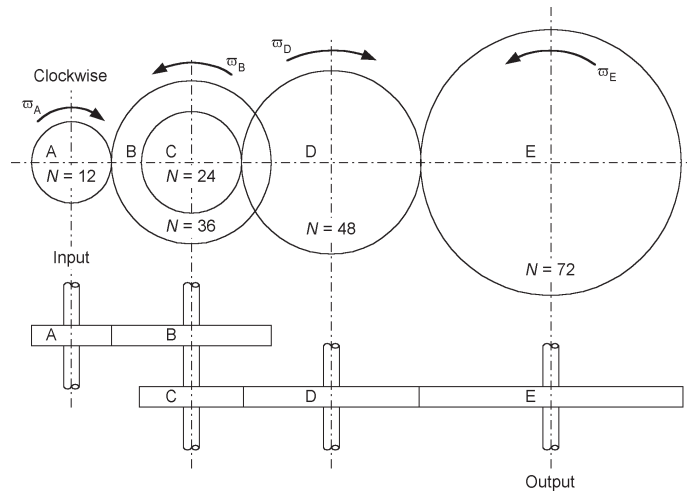


**Figure 5.4.9** Compound gear train

### Example 5.4.4

The compound gear train illustrated in Figure 5.4.9 consists of five gears, of which two, gears B and C, are on the same shaft. If gear A drives the system at 240 rpm (clockwise) and the number of teeth on each gear ( $N$ ) are as given. Determine the angular velocity of gear E at the output and the gear ratio for the system.

Again the system may be treated as gear pairs. Noting that gears B and C, being on the same shaft, rotate with the same angular velocity, then we proceed as follows.

From Equation (5.4.29), for a simple gear train, then the overall gear ratio ( $G$ ) is given by:

$$G = \frac{\omega_A}{\omega_E} = \frac{\omega_A}{\omega_B} \times \frac{\omega_B}{\omega_C} \times \frac{\omega_C}{\omega_D} \times \frac{\omega_D}{\omega_E} \quad \text{where } \omega_B = \omega_C,$$

so that the equation can be rewritten as:

$$G = \frac{\omega_A}{\omega_E} = \frac{\omega_A}{\omega_B} \times 1 \times \frac{\omega_C}{\omega_D} \times \frac{\omega_D}{\omega_E}.$$

The angular velocity of gear E can then be found as follows:

$$\omega_E = \left( \frac{N_A}{N_B} \times 1 \times \frac{N_C}{N_D} \times \frac{N_D}{N_E} \right) \omega_A$$

$$\omega_E = \left( \frac{12}{36} \times 1 \times \frac{24}{48} \times \frac{48}{72} \right) \omega_A = \frac{1}{9} \omega_A$$

$$\text{then } \omega_E = \frac{240}{9} = 26.66 \text{ rpm (counterclockwise).}$$

The direction can be established as counterclockwise from the velocity arrows shown in Figure 5.4.9.

Now from above  $G = \frac{\omega_A}{\omega_E} \neq \frac{N_A}{N_E}$  and considerable care is needed in compound gear examples.

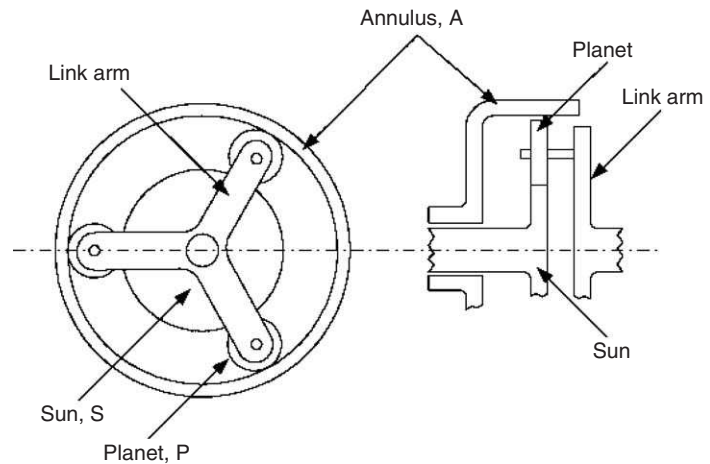


**Figure 5.4.10** Epicyclic gear train

### Epicyclic gear trains

Unique gear ratios and movements can be obtained in a gear train by permitting some of the gear axes to rotate about others. Such gear combinations are called *planetary* or *epicyclic* gear trains. Epicyclic gear trains (Figure 5.4.10) always include a *sun* gear, a *planet* carrier or *link arm*, one or more *planet* gears. Often a ring is included which has internal gear teeth that mesh with the planet gear, this is known as the *annulus*. Epicyclic gear trains are unusual mechanisms in that they have two degrees of freedom. This means, for constrained motion, that any two of the elements in the train provide an input.

Epicyclic gear trains are used for most automatic gearboxes and some complex steering mechanisms. The major advantages of epicyclic mechanisms are the attainment of high gear ratios and the fact that they act as bearings as well as transmissions.

We will now attempt to analyse the motion of the gear train illustrated in Figure 5.4.10, the following procedure may be used.

- (1) First imagine that the whole assembly is locked and rotated through 1 revolution, this has the effect of adding +1 revolution to all elements.
- (2) Next apply a rotation of  $-1$  to the gear which is to be fixed, this has the effect of cancelling the rotation given initially. Now the rotations of the other elements can be compared with the rotation of the fixed element, by considering the number of teeth on each of the gears.
- (3) Construct a table for each of the elements and add the results obtained from above actions. (When we sum the columns, using above procedure, we are able to ensure that the fixed link has zero revolution.)

#### Example 5.4.5

An epicyclic gear train (Figure 5.4.10) has a fixed annulus (A) with 180 teeth, and a sun wheel (S) with 80 teeth.

- (i) Determine the gear ratio between the sun and the link arm.
- (ii) Find the number of teeth on the planet gears and so determine the gear ratio between the planet and link arm.



So following our procedure, we first lock the whole assembly and give it +1 revolution (so the annulus is zero when we sum). Next we apply a -1 revolution to the gear specified as being at rest (so this element is zero when we sum) in this case the annulus. Then, to determine the arm in relation to the sun wheel, we fix the arm. Now the number of revolutions of the sun wheel to that of the annulus equals the inverse ratio of their teeth. The results of these various actions are given below.

| Action                                            | Effect |    |                   |                   |
|---------------------------------------------------|--------|----|-------------------|-------------------|
|                                                   | Arm    | A  | S                 | P                 |
| (1) Lock assembly give all elements +1 revolution | +1     | +1 | +1                | +1                |
| (2) Fix the link arm and give -1 to A             | 0      | -1 | $\frac{+180}{80}$ | $\frac{-180}{50}$ |
| (3) Sum 1 and 2                                   | (+1)   | 0  | 3.25              | (-2.6)            |

Then a rotation of +1 for the link arm results in +3.25 for the sun wheel. The gear ratio is 3.25:1.

Note that the number of teeth on the planet (P) was not required for part (1), because it is in effect an idler gear between the sun and the annulus.

(iii) Referring to Figure 5.4.10, we note that:

$r_s = 2r_p = r_A$  and we know from Figure that

$$\frac{d_S}{d_P} = \frac{t_S}{t_P} \text{ and } \frac{d_P}{d_A} = \frac{t_P}{t_A} \quad (a)$$

so  $\frac{r_S}{r_P} = \frac{2r_P}{r_P} = \frac{r_A}{r_P}$  then since  $2r = d$  we may write:

$$\frac{d_S}{d_P} = 2 = \frac{d_A}{d_P} \text{ and so from (a) } \frac{t_S}{t_P} = 2 = \frac{t_A}{t_P}$$

which on multiplication by  $t_P$  gives  $t_S = 2t_P = t_A$ .

$$\text{Then } t_P = \left( \frac{t_A - t_S}{2} \right) = \left( \frac{180 - 80}{2} \right) = 50.$$

The required results for the gear ratio between the planet and the link arm is -2.6:1. The negative sign indicates that the angular velocity of the planet is in the opposite direction to that found for the sun. Since angular velocities determine the gear ratio, this would be expected. These results are shown in brackets in the table.

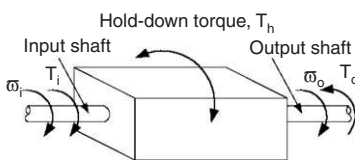


Figure 5.4.11 Gearbox torques and angular velocities

### Torque in gear trains

The torques and angular velocities at the input and output of a gear train are shown in Figure 5.4.11. The gear train could take the form of a gearbox. The input torque will be in the same direction as the rotation of

the input shaft, the torque at the output will act in the opposite direction to the rotation of the output shaft. This is because, the torque created by the shaft driving the load at the output, acts in opposition to the rotation of the drive shaft.

Under normal circumstances there will be a difference between the torque created at the input ( $T_i$ ) and the torque created at the output ( $T_o$ ). This difference produces a third torque which creates a positive or negative turning moment, which acts on the gear train casing. To counter this effect we need to ensure that the gear train assembly is held down securely, hence we refer to this third torque as the *hold-down torque* ( $T_h$ ). Under these circumstances the gear train assembly is in equilibrium and so the sum of the torques is zero, or

$$T_i + T_h + T_o = 0. \quad (5.4.31)$$

Also if there are no losses through the gear train then, the input power equals the output power. We already know that power ( $P$ ) =  $T\omega$ , and for our system the output torque acts in opposition to the input torque (in a negative sense), so we may write:

$$T_i\omega_i = -T_o\omega_o.$$

The above assumes that no losses occur through the system, in practice there are always losses, due to bearing friction, viscous friction and gear drag. Now since the efficiency of a machine is defined as:

$$\text{Efficiency } (\zeta) = \frac{\text{Power at output}}{\text{Power at input}}$$

$$\text{then } \zeta = \frac{T_o\omega_o}{T_i\omega_i}. \quad (5.4.32)$$

### Example 5.4.6

A gearbox powered by a directly coupled electric motor, is required to provide a minimum of 12 kW of power at the output, to drive a load at 125 rpm. The gearbox chosen consists of a compound train (Figure 5.4.12), which has an efficiency of 95%. Determine the:

- (i) gear ratio;
- (ii) electric motor output shaft velocity;

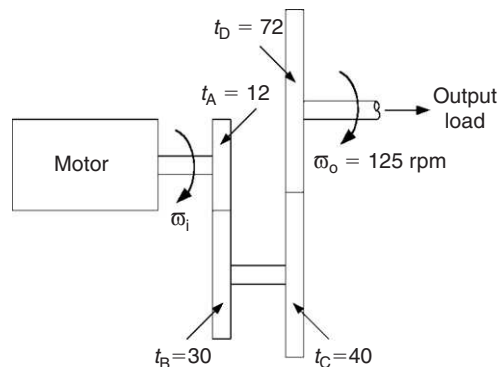


**Figure 5.4.12** Figure for Example 5.4.6

- (iii) power required from the electric motor;
- (iv) hold-down torque.

$$(i) \quad G = \frac{t_B}{t_A} \times \frac{t_C}{t_B} \times \frac{t_D}{t_C} = \frac{t_D}{t_A} \quad (\text{inverse of teeth ratios}).$$

From the diagram

$$G = \frac{30}{12} \times \frac{40}{30} \times \frac{72}{40} = \frac{72}{12} = \frac{6}{1} = \mathbf{6:1}.$$

- (ii)  $\omega_i = 6\omega_o = (6)(125) = \mathbf{750 \text{ rpm.}}$
- (iii) Power required at input is given by:

$$\text{Efficiency} = \frac{\text{Power at output}}{\text{Power at input}}$$

$$\text{so } P_i = \frac{12 \text{ kW}}{0.95} = \mathbf{12.63 \text{ kW.}}$$

- (iv) The hold-down torque may be found from Equation (5.4.31).

$$\text{Then since } P = T\omega \text{ so } T_i = \frac{P_i}{\omega_i} \quad \text{and} \quad \omega_i = 78.54 \text{ rad/s}$$

$$\text{then } T_i = \frac{12.63 \times 10^3}{78.54} = 160.8 \text{ Nm.}$$

$$\text{Similarly, at output } T_o = \frac{12 \times 10^3}{13.09} = -916.7 \text{ Nm} \\ (T_o \text{ negative}).$$

Then from Equation (5.4.31), where  $T_o + T_h + T_i = 0$ .

$$T_h = -T_o - T_i = 916.7 - 160.8 = \mathbf{755.9 \text{ Nm.}}$$

From the results you will note that the torque required from the motor is six times smaller than that required at the output. This is one of the reasons for placing a gearbox between the motor and the load, we are then able to substantially reduce the size of the motor!

## Balancing

As already mentioned in the introduction to this section, the effects of components being out of balance can be severe. You have already met the concept of *inertia* applied to rotational systems, when you studied angular kinetic energy in Chapter 3. Inertia forces and accelerations occur in all rotational machinery and such forces must be taken into account by the engineering designer. Here, we are concerned with the alleviation of inertia effects on rotating masses.

If a mass  $m$  is rigidly attached to a shaft which rotates with an angular velocity  $\omega$  about a fixed axis and the centre of the mass is at a radius  $r$

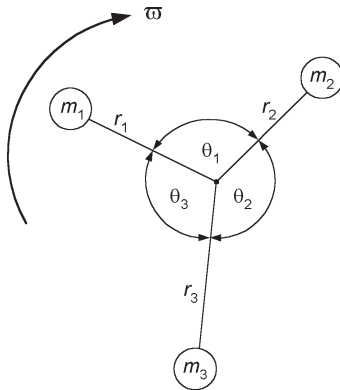


**Figure 5.4.13** Fixed masses about centre of rotating shaft

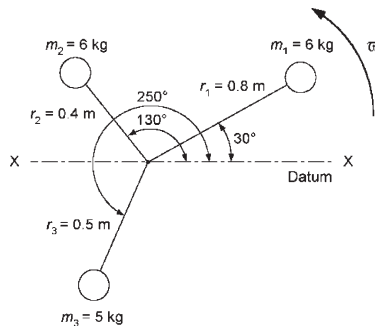


**Figure 5.4.14** Rotating mass system

from the axis. Then we know from our previous work that the centrifugal force which acts on the shaft as a result of the rotating mass is given by:

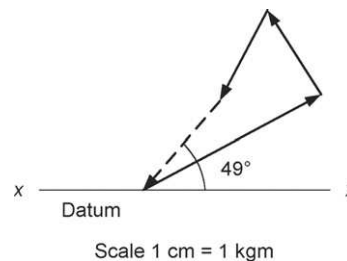
$$F_c = mr\omega^2. \quad (5.4.33)$$

When additional masses are added, which *act in the same plane* (Figure 5.4.13), they also produce centrifugal forces on the shaft. Then for equilibrium *to balance* the shaft, we require that the sum of all the centrifugal forces,  $\Sigma m\omega^2 r = 0$ . Since all the masses are rigidly fixed, they all rotate with the same angular velocity, so  $\omega$  may be omitted from the above expression to give  $\Sigma mr = 0$ . Now to determine whether or not the shaft is in balance we can simply produce a vector polygon, where the vector product  $mr$ , is drawn to represent each mass at its radius, in the system.

### Example 5.4.7

A rotating system consists of three masses rigidly fixed to a shaft, which all act in the same plane, as shown in Figure 5.4.14. Determine the magnitude and direction of the mass required for balance, if it is to be set at a radius of 0.5 m from the centre of rotation of the shaft.

Then from the diagram, we draw our vector polygon, where each component is  $mr$ . Then  $m_1 r_1 = 4.8 \text{ kg m}$ ,  $m_2 r_2 = 2.4 \text{ kg m}$ , and  $m_3 r_3 = 2.5 \text{ kg m}$ . Now constructing the vector polygon, we get which gives a value for the balancing couple



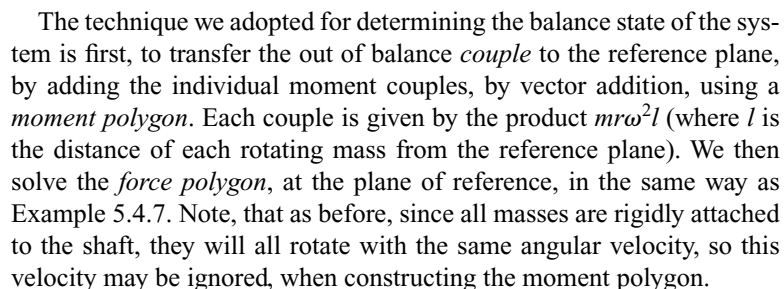
of 2.5 kg m acting at the angle shown above. So required mass at radius 0.5 m is:

$$\frac{2.5 \text{ kg m}}{0.5 \text{ m}} = 5 \text{ kg}.$$

Note that the common angular velocity,  $\omega$ , was not needed to determine the out of balance force.

If the masses rotate in *different planes*, then not only must the centrifugal forces be balanced but the moments of these forces about any plane of revolution must also be balanced. We are now considering rotating masses, in three dimensions.

This system is illustrated in Figure 5.4.15, where masses now rotate about our shaft, some distance ( $l$ ) from a reference plane. The result of these masses rotating in *parallel* planes to the reference plane is that they create moment effects. Quite clearly the magnitude of these moments will depend on the position chosen for the reference plane.

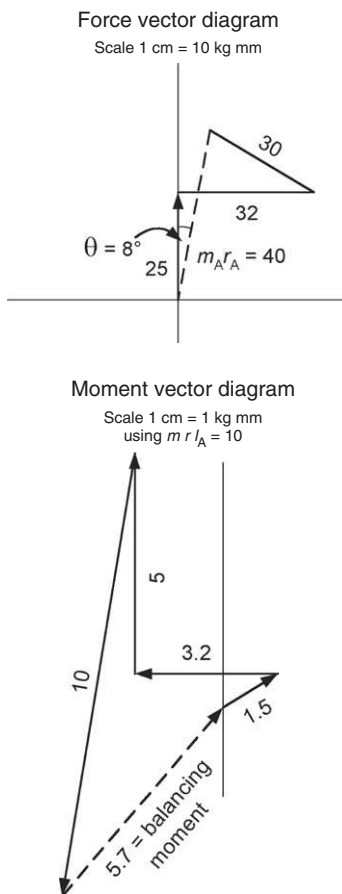
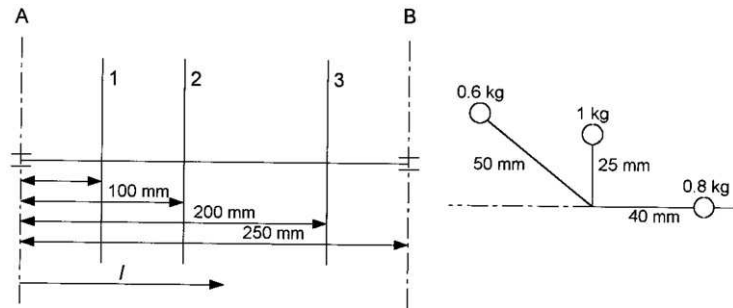


So we have a *two stage* process which first involves the production of a *moment polygon* for the couples, from which we determine whether or not there is an out of balance moment. If there is, we will know the nature of this moment; that is: its mass, radius of action, angular position, and distance from the reference plane (*mrl*). The *force polygon* may then be constructed, which will include the *mr* term from the moment polygon. We are then able to balance the system by placing the *balance* mass, derived from our polygons, at a position (*l*), angle (*t*), and radius (*r*) to achieve *dynamic balance* of the rotating system.

If the shaft is supported by bearings, radial balance across the bearing will be achieved by adopting the above method. However, axial forces may still exist at the bearing if this is not chosen as the reference plane for the unknown mass. This would result in longitudinal in-balance of the shaft (Figure 5.4.15b). There would then be a requirement for the addition of a counter-balance mass, to be placed in an appropriate position to one side of the shaft bearing. Thus there may be the need for two balance masses, to ensure total dynamic balance throughout the system.

The above procedure should be readily understood by considering the following example.

**Figure 5.4.16** Figure for Example 5.4.8



### Example 5.4.8

A rotating shaft is supported by bearings at A and B. It carries three disc cams, which are represented by the equivalent concentrated masses shown in Figure 5.4.16. Their relative angular positions are also shown.

For the given situation, calculate the magnitude of the reaction at bearing A resulting from dynamic imbalance when the shaft rotates at 3000 rpm.

Using  $F = m r \omega^2$  we first need to establish the  $m_A r_A$  value, then find  $m_A r_A l_A$  and so determine the reaction at bearing A, as required.

To assist us we will set up a table, and by taking moments about 'B', we eliminate it from the calculation, in the normal manner. Then from the diagram we find values as:

| Plane | $m$ (kg) | $r$ (mm) | $mr$      | $l$ (m) | $mrl$          |
|-------|----------|----------|-----------|---------|----------------|
| A     | $m_A$    | $r_A$    | $m_A r_A$ | 0.25    | $0.25 m_A r_A$ |
| 1     | 1        | 25       | 25        | 0.2     | 5              |
| 2     | 0.8      | 40       | 32        | 0.1     | 3.2            |
| 3     | 0.6      | 50       | 30        | 0.05    | 1.5            |
| B     | $m_B$    | $r_B$    | $m_B r_B$ | 0       | 0              |

Check that you arrive at the values in the table. Next we construct our force polygon as in the previous example.

Now the required force may be found since  $m_A r_A =$

$$\frac{5.7}{0.25} \text{ kg mm} = 22.8 \text{ kg mm, then}$$

$$F_A = m_A r_A \omega^2 = (22.8 \times 10^{-3}) \left( \frac{3000 \times 2\pi}{60} \right)^2 = 2250 \text{ N.}$$

so force on bearing A = 2250 N.

Note that reaction B may be found in a similar manner by taking moments about A.

### Flywheels

A flywheel is an energy storage device. It absorbs mechanical energy by increasing its angular velocity and gives out energy by decreasing its angular velocity. The mechanical energy is given to the flywheel by a torque, which is often provided by a motor. It is accelerated to operating speed, thus increasing its inertial energy, which can be used to drive a load after the accelerating torque has been removed.

## Mathematics in action

If a flywheel is given an input torque  $T_i$  at angular velocity  $\omega_i$ , this will cause the flywheel to increase speed. Then if a load or some other output acts on the flywheel, it will create an opposing torque  $T_o$  at angular velocity  $\omega_o$ , which will cause the flywheel to reduce speed. The difference between these torques will cause the flywheel to accelerate or retard, according to the relationship.

$$T_i(\omega_i) - T_o(\omega_o) = I\omega'$$

where  $\omega' = d\omega/dt$  which is the rate of change of velocity with respect to time, which of course is the angular acceleration  $\alpha$  and  $I$ , you should recognise as the second moment of inertia (see Equation 3.2.14). We could have applied the torque from rest through some angle  $t$ , then both angular displacement and angular velocity become variables. Then the above relationship connecting input, output torque, and acceleration would be written:

$$T_i(\theta_i, \omega_i) - T_o(\theta_o, \omega_o) = I\theta''$$

where  $\theta'' = d^2\theta/dt^2$  which is the second derivative of angular displacement with respect to time or again, acceleration  $\alpha$ .

Now assuming that the flywheel is mounted on a rigid shaft, then  $\theta_i = \theta_o = \theta$  and so  $\omega_i = \omega_o = \omega$ .

Then the above equation may now be written as:

$$T_i(\theta, \omega) - T_o(\theta, \omega) = I\theta'' = I\alpha. \quad (5.4.34)$$

The above equation can be solved by direct integration, when the start values for angular displacement and angular velocity are known.

Now the *work input* to a flywheel  $= T_i(\theta_1 - \theta_2)$ , where the torque acts on the flywheel shaft through the angular displacement  $(\theta_1 - \theta_2)$ . Similarly the work output to a flywheel  $= T_o(\theta_3 - \theta_4)$  where again the torque acts on the shaft through some arbitrary angular displacement.

We can write these relationships in terms of kinetic energy (see Equation (3.2.21)). At  $\theta = \theta_1$  the flywheel has a velocity  $\omega_1$  rad/s, and so its kinetic energy is:

$$E_1 = \frac{1}{2}I\omega_1^2$$

and similarly at,  $\theta_2$  with velocity  $\omega_2$

$$E_2 = \frac{1}{2}I\omega_2^2.$$

So the change in the kinetic energy of the flywheel is given by:

$$E_2 - E_1 = \frac{1}{2}I(\omega_2^2 - \omega_1^2). \quad (5.4.35)$$

### Example 5.4.9

A steel cylindrical flywheel is 500 mm in diameter, has a width of 100 mm and is free to rotate about its polar axis. It is uniformly accelerated from rest and takes 15 s to reach an angular velocity of 800 rpm. Acting on the flywheel is a constant friction torque of 1.5 Nm. Taking the density of the steel to be

7800 kg/m<sup>3</sup>, determine the torque which must be applied to the flywheel to produce the motion.

To solve this problem we will need to use not only Equation (5.4.34), but also to refer back to Chapter 3 for the work we did on angular motion.

To find the angular acceleration, we may use  $\omega_f = \omega_i + \alpha t$ , which gives:

$$83.8 = 0 + 15\alpha \text{ (where 800 rpm} = 83.8 \text{ rad/s)}$$

$$\text{then } \alpha = 5.58 \text{ rad/s}^2.$$

Now the mass moment of inertia  $I$ , for a cylindrical disc, is given by:

$$I = \frac{1}{2}Mr^2 \text{ (see Figure 3.2.5).}$$

In order to find  $I$  we first need to find the mass of the flywheel, that is:

$$M = (0.25\pi)(0.5)^2(0.1)(7800)$$

$$M = 153.15 \text{ kg.}$$

$$\text{Then } I = (0.5)(153.15)(0.25)^2$$

$$I = 4.79 \text{ kg m}^2.$$

Now the net accelerating torque is given from Equation (5.4.34):

$$T_N = T_i - T_o = I\alpha$$

$$T_N = (T_i - 1.5) = (4.79)(5.58).$$

Thus the accelerating torque  $T_i = 28.23 \text{ Nm}$ .

A further example of an integrated system involving the use of a flywheel is given in the next section, when we deal with coupled systems.

### Coupled systems

In the final part of this section we are going to look at *coupled systems*. For example, an electric motor may be used as the initial power source for a lathe. The rotary and linear motion required for machining operations being provided by a gearbox and selector mechanism, positioned between the motor and the workpiece. The motor vehicle provides another example where the engine is the prime mover, coupled through a clutch to a gearbox and drive shafts.

We are already armed with the necessary under pinning principles, from our previous work on belt drives, clutches, gear trains, rotating masses, and flywheels. We may, however, also need to draw on our mechanical knowledge from Chapter 3, particularly with respect to angular motion.

We will study this mainly through examples, which are intended to draw together and consolidate our knowledge of power transmission systems.

Let us first consider the inertia characteristics of a rotational system, in this case a gearbox. Remembering our work on free-body diagrams from our study of shear force and bending moments!



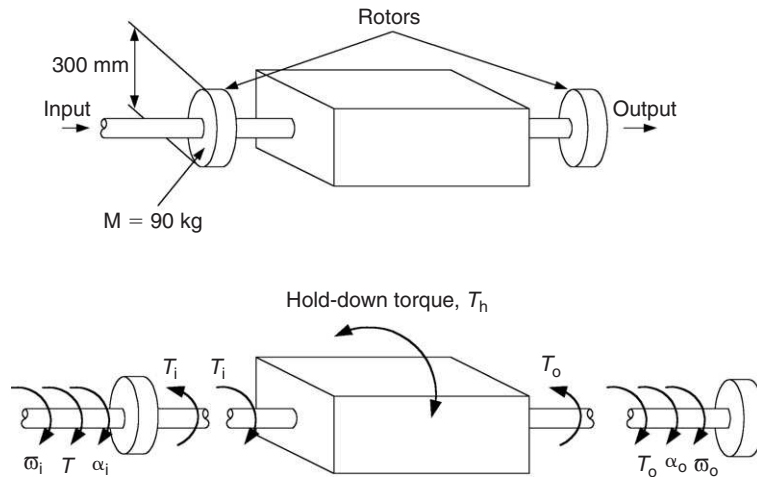


**Figure 5.4.17** Gearbox assembly with rotors

Figure 5.4.17 shows a gearbox assembly with rotors attached to its input and output shaft. Let us assume that the gearbox ratio is a step-down; that is, the input velocity is higher than the output velocity and these are related to the overall gear ratio  $G$ , by Equation (5.4.29).

Now each of the rotors will have a mass moment of inertia which depends on their radius of gyration, for the purpose of our analysis we will assume the rotors are cylindrical. Then if we draw the free-body diagram for the system, *with respect to the input* (Figure 5.4.17b), we are able to consider with respect to any component in the system, the resulting torques, velocities, and accelerations. We will use these ideas in the following example.

### Example 5.4.10

A gearbox has rotors attached to its input and output as shown in Figure 5.4.17a, where for the output rotor  $I_o = 25 \text{ kg m}^2$ , the input rotor is cylindrical having a mass of 90 kg and diameter of 300 mm. The angular velocity of the rotor on the input shaft is higher than that of the rotor at the output. If a torque of 1.0 kN is applied to the high-speed input shaft. Calculate the value of the angular acceleration of the input shaft, when  $G = 5$  and the gearbox has an efficiency of 95%.

In order to solve this problem we first need to formulate the relationships, which enable us to determine the appropriate torques and accelerations for the system, from the point of view of the input.

The free-body diagram in Figure 5.4.17b shows the isolated system elements (when the torque is applied to the input rotor shaft), as the input rotor and drive shaft, the gearbox assembly, and the output rotor and drive shaft. For the input rotor the torque,  $T$ , is opposed by the torque produced at the gearbox side of the shaft, so the equation of motion for  $I_i$  is given as:

$$T - T_i = I_i \alpha_i \quad (\text{a})$$

Compare this with Equation (5.4.34).

For the gearbox the input and output torques are directly related to the gearbox ratio and the efficiency of the gearbox by:

$$T_o = T_i G \zeta. \quad (b)$$

The above relationship is true for a reduction gearbox, if you remember that one of the primary functions of a gearbox is to provide an increase in power at the output and if the velocity at the output is reduced, as in this case, the torque is raised (see Equation (5.4.32)).

Now at the output rotor, from the free-body diagram we have the relationship:

$$T_o = I_o \alpha_o \quad (\text{as before}) \quad (c)$$

$$\text{and } G \alpha_o = \alpha_i. \quad (d)$$

Now combining equations (b) and (c) gives:

$$T_i G \zeta = I_o \alpha_o$$

and substituting the above expression into equation (a) and at the same time from (d) substituting for  $\alpha_o$  we get:

$$T - I_o \alpha_o = I_i \alpha_i.$$

$$\text{and so finally } T = \alpha_i [I_i + I_o / G^2 \zeta]. \quad (e)$$

So this equation tells us that the torque for the system (when for a reduction gearbox the torque is assumed to act at the high velocity input side) is equal to the acceleration caused by this torque multiplied by the equivalent inertia ( $I_e$ ) of the system, which is the term in the square brackets.

As you can see, most of this example has been concerned with formulating the required relationship, the above technique can be used when developing system equations for many similar situations.

The required acceleration at the input can be found by direct substitution of the variables into equation e, once we have determined  $I_i$ . Then mass moment of inertia for a cylinder gives

$$I_i = \frac{1}{2} M r^2 = (0.5)(90)(0.15)^2 = 1.01 \text{ kg m}^2.$$

Now from equation (e), acceleration of input shaft is:

$$\alpha_i = \frac{1.0}{(1.01 + 25/23.75)} = \frac{1.0}{(1.01 + 1.053)} = 0.48 \text{ rad/s}^2.$$

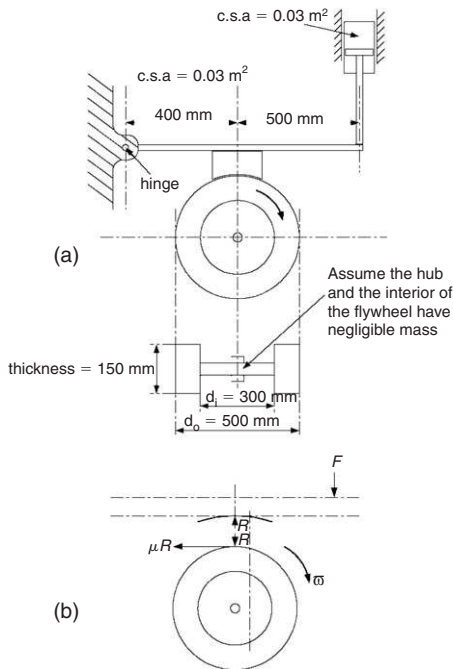


**Figure 5.4.18** Flywheel and shoe brake system

We will now consider another fairly simple system which consists of a shoe brake, being used to retard a flywheel. The flywheel may be thought of as an annulus, where the internal thickness is considered, as a first approximation, to be negligible (Figure 5.4.18).

### Example 5.4.11

The steel flywheel shown in Figure 5.4.18 has an external diameter of 500 mm, an internal diameter of 300 mm, and the

annulus formed by these diameters is 150 mm thick. The flywheel is retarded by means of a lever brake. Just prior to the application of the brake the flywheel is rotating with an angular velocity of 3500 rpm and then the brake is applied for 50 s. A piston assembly is used to activate the brake assembly acting on the end of the lever. If the pressure applied by the piston is 3.5 kPa and the piston area is  $0.03 \text{ m}^2$ . Find:

- (i) the angular velocity of the flywheel at the end of the braking period;
- (ii) the power released from the flywheel during the braking period.

Take the flywheel density as  $7800 \text{ kg/m}^3$  and the coefficient of friction between the brake and the flywheel as 0.3.

Figure 5.4.18a shows the situation diagram with all relevant dimensions, Figure 5.4.18b shows the free-body diagram, illustrating the braking reaction.

(i) The solution to this part of the problem relates to the equation  $T = I\alpha$ , which is a simplified form of Equation (5.4.34), where, in our case,  $T = T_B = \text{torque due to braking}$ . We need to find this and the mass moment of inertia of the flywheel.

To find the braking torque we first determine the braking force, this is:

$$F = P \times A = (3500)(0.03) = 105 \text{ N.}$$

Now considering the equilibrium of the brake lever, the reaction force  $R$  is found by taking moments about the hinge (see free-body diagram).

$$\text{So } R \times 400 = (105)(500) = 131.25 \text{ N.}$$

So, assuming the friction is limiting friction and using the fact that braking force  $F = \mu R$ , then braking torque  $T_B = (\mu R)r$  ( $r = \text{radius of flywheel}$ ).

$$\text{So } T_B = (0.3)(131.25)(0.25) = 9.84 \text{ Nm.}$$

Now the mass moment of inertia of a disc is given by:

$$I = \frac{1}{2}Mr^2$$

Logic suggests that  $I$  for an annulus will have the form  $\frac{1}{2}M(R^2 - r^2)$ , which it does. However, it is normally expressed as:

$$I = \pi \frac{(d_o^4 - d_i^4)t\rho}{32}$$

where  $t = \text{thickness of the flywheel}$  and  $\rho = \text{density of flywheel material}$ . The above form is often used as a first approximation where the hub and inner disc is considered insignificant, when compared with the bulk of the rim.

$$\text{So } I_F = \frac{\pi(0.5^4 - 0.3^4)(0.15)(7800)}{32}$$

giving  $I_F = 6.25 \text{ kg m}^2$ .

To find the angular retardation  $\alpha$ , then

$$T_B = I\alpha \quad \text{and} \quad \alpha = \frac{9.84}{6.25} = 1.574 \text{ rad/s}^2.$$

So final angular velocity is given by Equation (3.2.9).

$$\omega_f = \omega_i + \alpha t \quad \text{and} \quad \omega_f = (366.5 - 78.5)$$

and  $\omega_f = \mathbf{288 \text{ rad/s}}$  or 2750 rpm.

(ii) To find the power given out by the flywheel during braking, we first find the kinetic energy change. Then using Equation (5.4.35) where  $E_2 - E_1 = \frac{1}{2}I(\omega_2^2 - \omega_1^2)$  then energy change =  $(0.5)(6.25)(366.5^2 - 288^2)$

$$= 160.588 \text{ kJ}$$

$$\text{and so power } P = \frac{J}{S} = \frac{160.558}{50} = \mathbf{3.21 \text{ kW}}.$$

We now discuss a system for the *operation of a crank pressing machine*, which consists of a prime mover, gearbox, and flywheel, to drive the press crank.

Let us first consider the selection of a power unit for a punch pressing machine. What information do we require to help with our selection? Well, at the very least, we would need information about the service load, the frequency of the pressing operation, the torque requirements, and rotational velocities of the system.

Figure 5.4.19 shows the torque requirements for a typical pressing operation. Note that the torque requirement fluctuates, very high torque being required for the actual pressing operation, which is of course what we would expect.

We first determine the torque versus angular displacement graph (Figure 5.4.19).

Let us assume that the press operates at 150 rpm and the pressing operation is completed in 0.3 s and takes place 30 times a minute. This is not an unreasonable assumption, many punch presses operate at much higher rates than this, obviously dependent on the intricacy and force requirements of the product. Assume that the resisting torque during the pressing operation is 1800 Nm.

Based on the above information we can deduce the angular displacement of the crank for one complete cycle of events (Figure 5.4.19).

The punching operation takes place at 30 times a minute; therefore, the punch press crank will rotate 5 revolutions, for each operation. So the angular displacement for one complete operation is  $5 \times 2\pi = 10\pi$  rad.

The angular displacement of the crankshaft during the actual pressing operation, is easily determined by comparing the time for one complete cycle of events (2 s), with the time for the actual pressing operating (0.3 s).

So angular displacement for pressing operation =  $(0.3 \times 10\pi/2.0) = 1.5\pi$  rad (Figure 5.4.19).

We know that the high resisting torque of 1800 Nm acts for 15% of the time required for one complete cycle of events, this torque would need to

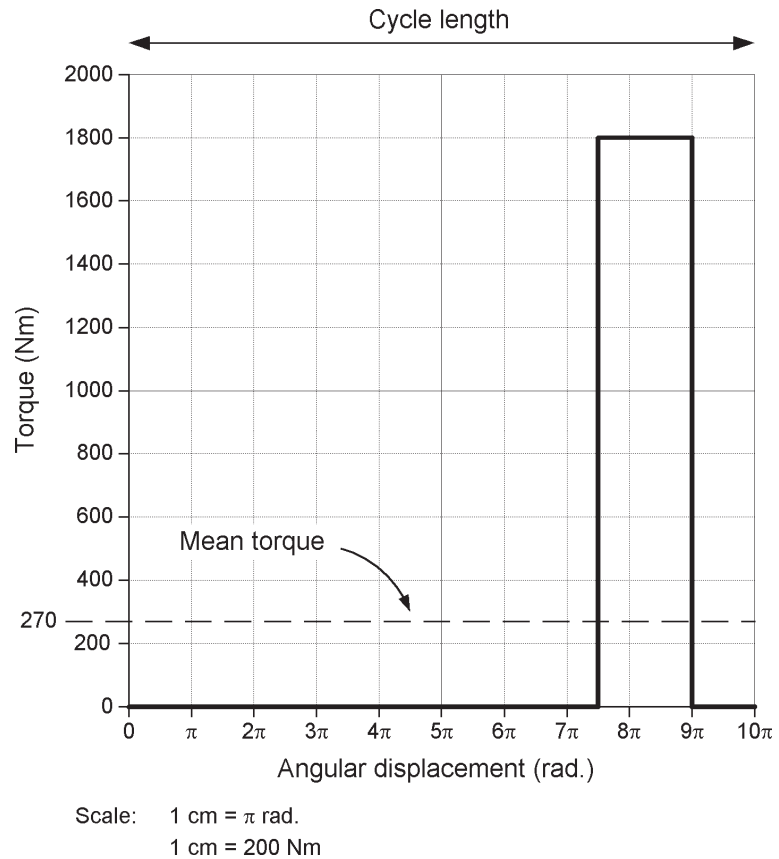


**Figure 5.4.19** Turning moment diagram for crank pressing machine

be provided by the prime mover, if no flywheel was included in the system. So, what power would we require from our prime mover under these circumstances?

Using an electric motor (the most likely source of power) we would need to provide gearing between the motor output and the input to the press crank, for the following reasons.

A suitable electric motor for such a task is likely to be three-phase AC. This type of motor is normally produced to run at synchronous speeds that depend on the number of pole pairs of the machine, 600, 900, and 1200 rpm, etc. We will choose a motor that runs at 1200 rpm, which requires a reduction gearbox of eight to one. What is the power required from our motor to provide the maximum torque, assuming that this peak torque needs to be *continuously* available from the motor?

*Without losses* we know that the power provided by the motor is equal to the power at the crankshaft, to drive the press. Then from the diagram we have already seen that the peak torque we require at the crank is 1800 Nm. Then we know that:

$$T_i \omega_i = T_o \omega_o$$

$$\frac{T_i}{T_o} = \frac{\omega_o}{\omega_i} \quad \text{so} \quad T_i = \frac{(1800)(60)}{1200} = 90 \text{ Nm.}$$

Note that we may leave the angular velocities in rpm because here they are a ratio.

Now if the motor has the capacity to deliver this torque *continuously*, then the *work* required from the motor in one complete cycle of events is  $10\pi \times 90 = 900\pi \text{ Nm}$ .

Now power is the rate of doing work, so in 1 s the motor shaft rotates 20 revolutions, the power required from our motor is  $20 \times 900\pi = 56.5 \text{ kW}$ .

It is obviously wasteful to provide such a large motor when the maximum torque requirements are only needed for a small fraction of the time. This is why our system *needs* the addition of a flywheel.

Let us assume that we require a flywheel to be fitted to our system, which keeps the press crank velocity between 140 and 160 rpm during pressing operations. In order to select such a flywheel we need to know the energy requirements of the system, and the total inertia that needs to be contributed by the flywheel during these pressing operations.

From our discussion so far I hope you recognise that the area under our crank torque–crank angle graph in Figure 5.4.19, represents the energy requirements for the press.

We assume as before that the energy in is equal to the energy out, and that there are no losses due to friction.

The total energy required during the pressing operation is given from our previous calculations as:

$$(1800 \text{ Nm})(1.5\pi) = 8482 \text{ J (area under graph for pressing operation)}.$$

This energy is required to be spread over the complete operation cycle, that is, spread over  $10\pi$  rad, so the uniform torque required from the motor is given by:

$$1800 \times 1.5\pi = T_m \times 10\pi \quad \text{where } T_m = \text{uniform motor torque}$$

$$T_m = 270 \text{ Nm (this is indicated on the diagram)}.$$

Then the *change* in energy is:

$$\Delta E = (T_{\max} - T_{\text{mean}}) \times \text{angular distance}$$

$$\Delta E = (1800 - 270)1.5\pi = 7210 \text{ J}.$$

So the flywheel must supply 7210 J and therefore the motor must supply approximately:

$$8482 - 7210 = 1272 \text{ J}.$$

Note the drastic reduction in the energy required by the motor.

We are now in a position to determine the inertia required by the flywheel, using Equation 5.4.34 we have:

$$7210 = \frac{1}{2}I(\omega_{\max}^2 - \omega_{\min}^2)$$

$$\text{where } \omega_{\max} = 160 \text{ rpm} = 16.76 \text{ rad/s and}$$

$$\omega_{\min} = 140 \text{ rpm} = 14.66 \text{ rad/s}.$$

Then the inertia of the flywheel is given by:

$$I = \frac{2 \times 7210}{(280.9 - 214.9)} = 218.5 \text{ kg m}^2.$$

This is where we end our discussion of this system. We could go on to make a specific motor selection, determine the requirements of a pulley

and belt drive system, and also determine the dimensional requirements of the flywheel, if we were so inclined.

We leave our discussion on the coupling of transmission and other systems, with one further example.

### Example 5.4.12

An electric motor drives a cylindrical rotational load through a clutch (Figure 5.4.20). The armature of the motor has an external diameter of 400 mm and a mass of 60 kg. It may be treated as being cylindrical in shape. The rotor has a diameter of 500 mm and a mass of 480 kg. If the motor armature is revolving at 1200 rev/min and the rotor at 450 rpm, both in the same direction. Assuming the moment of inertia of the driving clutch plate is  $0.25 \text{ kg m}^2$  and that after engagement the moment of inertia of the whole clutch is  $0.5 \text{ kg m}^2$ . Ignoring all other inertia effects, find the loss of kinetic energy when the clutch is engaged.

This problem requires us to find the angular momentum change, as a result of the impact. Angular momentum is given by the product of moment of inertia and the angular velocity. Knowing this, we are able to find the angular velocity after impact, and then the loss of kinetic energy.

The mass moment of inertia for a cylinder is determined simply from the  $I = \frac{1}{2}Mr^2$  in the normal way.

Then moment of inertia of motor armature and driving clutch assembly, where  $I_1$  = driving clutch assembly:

$$\begin{aligned} &= I_m + I_1 \\ &= (0.5)(60)(0.2)^2 + 0.25 \\ &= 1.45 \text{ kg m}^2. \end{aligned}$$

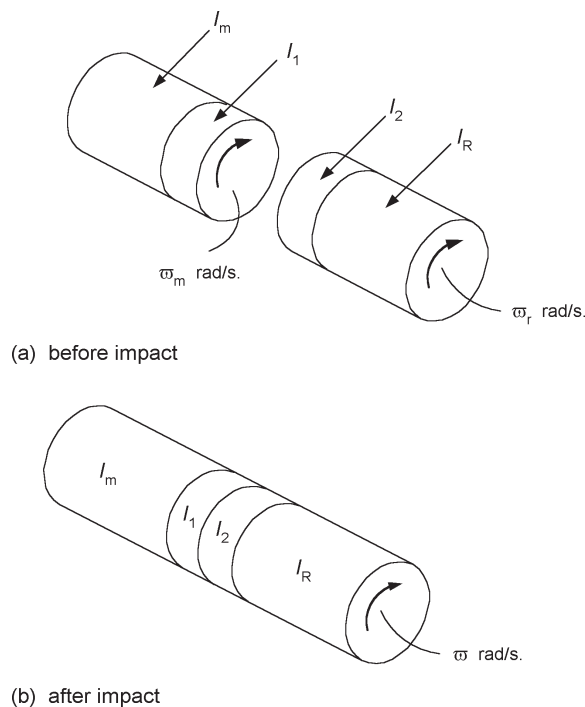


**Figure 5.4.20** Figure for Example 5.4.12

Moment of inertia of rotor end driven clutch assembly, where  $I_2$  = driven clutch assembly

$$= I_R + I_2.$$

$$= (0.5)(480)(0.25)^2 + 0.5$$

$$= 15.5 \text{ kg m}^2.$$

Also the angular velocity of the motor armature and driving clutch plate (same shaft) = 125.66 rad/s and the angular velocity of the rotor end driven clutch assembly (same shaft) = 47.12 rad/s.

Now from the conservation of angular moment we know that the total angular moment before impact equals the total angular moment after impact.

$$\text{Therefore, } (I_m + I_1)\omega_m + (I_R + I_2)\omega_R = (I_M + I_1 + I_R + I_2)\omega_f$$

where  $\omega_f$  = final velocity after impact.

$$\text{So } (1.45)(125.66) + (15.5)(47.12) = (1.45 + 15.5)\omega_f$$

$$912.567 = 16.95 \omega_f$$

$$\omega_f = 53.84 \text{ rad/s.}$$

Now to find the loss of kinetic energy at impact, then

Loss of KE = total KE before impact – total KE after impact

$$= (\text{KE of } I_m + I_1) + (\text{KE of } I_R + I_2) - \text{KE of } (I_M + I_1 + I_R + I_2)$$

$$= \frac{1}{2}(1.45)(125.66)^2 + \frac{1}{2}(15.5)(47.2)^2 - \frac{1}{2}(16.5)(53.84)^2$$

$$= 4799 \text{ J.}$$

**So loss of KE = 4.799 kJ.**

### Questions 5.4.1

- (1) A flat belt connects two pulleys with diameters of 200 and 120 mm. The larger pulley has an angular velocity of 400 rpm. Given that the distance  $C$  between centres is 400 mm, the belt maximum tension is 1200 N and the coefficient of friction between belt and pulleys is 0.3. Determine:
  - (i) the angle of lap of each pulley;
  - (ii) the maximum power transmitted.
- (2) A single V-belt with  $\phi = 18^\circ$  and a unit weight of 0.25 kg/m is to be used to transmit 14 kW of power from a 180 mm diameter drive pulley rotating at 2000 rpm to a driving pulley rotating at 1500 rpm. The distance between the centres of the pulleys is 300 mm.
  - (i) If the coefficient of friction is 0.2 and the initial belt tension is sufficient to prevent slippage, determine the values of  $T_1$  and  $T_2$ .



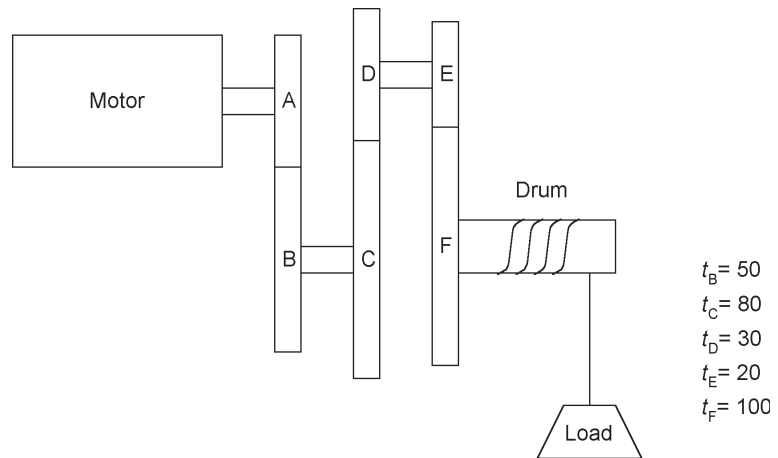


**Figure 5.4.21** Figure for Question 5.4.1(4) – motorised winch.

- (ii) Determine the maximum stress in the belt, if its density is  $1600 \text{ kg/m}^3$ . Note that the centripetal tension created by the centrifugal force acting on the belt is given by  $T_c = m'v^2$ , where  $m'$  = the mass per unit length of the belt.
- (3) A multiplate clutch consists of five discs, each with an internal and external diameter of 120 and 200 mm, respectively. The axial clamping force exerted on the discs is 1.4 kN. If the coefficient of friction between the rubbing surfaces is 0.25, and uniform wear conditions may be assumed. Determine the torque that can be transmitted by the clutch.
- (4) A motorised winch is shown below (Figure 5.4.21). The motor supplies a torque of 10 Nm at 1400 rpm. The winch drum rotates at 70 rpm and has a diameter of 220 mm. The overall efficiency of the winch is 80%. Determine the:
- number of teeth on gear A;
  - power input to the winch;
  - power output at the drum;
  - load being raised.
- (5) A flat belt 150 mm wide and 8 mm thick transmits 11.2 kW. The belt pulleys are parallel, and their axes are horizontally spaced 2 m apart. The driving pulley has a diameter of 150 mm and rotates at 1800 rev/min, turning so that the slack side of the belt is on top. The driven pulley has a diameter of 450 mm. If the density of the belt material is  $1800 \text{ kg/m}^3$ . Determine the:
- the tension in the tight and slack sides of the belt if the coefficient of friction between the belt and the pulleys is 0.3;
  - the length of the belt.

# 6

## Analytical methods

### Summary

The primary aim of this chapter is to introduce the mathematical principles, methods, and applications that underpin the Higher National Engineering units that you will meet in this book. The fundamental techniques presented here, should also prepare you for further study of Mathematics and other Engineering units which you may encounter during your future studies. You will not only cover aspects from the core analytical methods unit, but also selected topics from some of the more advanced units where these are considered necessary to aid your understanding.

This chapter is not intended to be a mathematical treatise! There are many excellent text books (see Appendix 1) which cover in great detail the foundations of mathematics, and provide an in-depth treatment of all the material introduced in the mathematical units of Higher National Engineering programmes.

We will, however, be covering in detail all the major areas of mathematics contained in the analytical method's core unit: that is fundamental algebra, trigonometry, calculus and statistics. In addition we will briefly cover: complex numbers, differential equations, matrices, Fourier series, and Laplace transforms (LTs), where their application is considered necessary to enhance the subject matter presented in this book.

Each topic will be presented in three parts. Starting with a formula sheet, the *formulae* being presented without proof. We will then look at the mathematical *methods* used for manipulating these formulae and establishing mathematical relationships. Finally, where appropriate, examples of their engineering *application* will be given drawn directly from the subject matter contained within this book. Since mathematics is a linear subject, some topics simply act as prerequisites for others, without having a direct engineering application. Under these circumstances only the appropriate formulae and methods will be presented.

In order to successfully study the subject matter presented in this way, you should be armed, already, with a number of fundamental mathematical tools. You *must* be completely familiar with elementary algebra, trigonometry, and number. In particular, you should be able to manipulate algebraic formulae and fractions, equations, indices, and logarithms, constants of proportionality, and inequalities. You should be able to calculate the areas and volumes of familiar solids, and solve problems involving the basic trigonometric ratios and their inverses. You should be familiar with estimation techniques and the manipulation and presentation of experimental data. Finally, you should be able to manipulate and use vulgar fractions!

Many people experience difficulties with the basics, identified above. You will notice as you study the analytical subjects presented in this book that much of your time will be spent manipulating expressions *using these basics*, this is why they are so important. The intention now, is to build on these basics by studying a number of additional analytical techniques. You will then be in a position to provide the solution to engineering problems more quickly, and hopefully with less effort!

## 6.1 ALGEBRA

In this section we will review those fundamental algebraic methods, which are considered essential to manipulate engineering expressions. These include, logarithms and logarithmic functions, exponential laws, and the solution of equations involving exponential and logarithmic functions.

New algebraic methods will be introduced primarily concerned with partial fractions (PFs), which we will need later to simplify expressions in order to *integrate* them more easily. PFs will also be used when we find LTs and their inverses.

We will finally take a look at the binomial, exponential, logarithmic, and hyperbolic series. These will be required for the power series and Fourier analysis which you will meet later.

### Algebraic fundamentals

#### Formulae

(1) *Factors:*

$$\begin{aligned}(a + b)^2 &= a^2 + 2ab + b^2 \\ (a - b)^2 &= a^2 - 2ab + b^2 \\ a^3 + b^3 &= (a + b)(a^2 - ab + b^2) \\ a^3 - b^3 &= (a - b)(a^2 + ab + b^2) \\ (a + b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3 \\ (a - b)^3 &= a^3 - 3a^2b + 3ab^2 - b^3.\end{aligned}$$

(2) *Indices:*

$$a^m \times a^n = a^{m+n} \quad (\text{i})$$

$$\frac{a^m}{a^n} = a^{m-n} \quad (\text{ii})$$

$$(a^m)^n = a^{mn} \quad (\text{iii})$$

$$\frac{1}{a^n} = a^{-n} \quad (\text{iv})$$

$$\sqrt[n]{a^m} = a^{m/n} \quad (\text{v})$$

$$a^0 = 1. \quad (\text{vi})$$

(3) *Logarithms:*

$$\text{If } a = b^c, \text{ then } c = \log_b a \quad (\text{i})$$

$$\log_a MN = \log_a M + \log_a N \quad (\text{ii})$$

$$\log_a \frac{M}{N} = \log_a M - \log_a N \quad (\text{iii})$$

$$\log_a (M^n) = n \log_a M \quad (\text{iv})$$

$$\log_b M = \frac{\log_a M}{\log_a b}. \quad (\text{v})$$

(4) *Exponential and logarithmic functions:*

$$y = a^x \quad \text{where } x \text{ is the exponent or index}$$

$$\exp x = e^x$$

$$\log_e x = \ln x \quad (\text{where } \ln \text{ is the log to base } e \text{ known as the natural or Naperian logarithm})$$

$$y = \log_a x \quad (\text{logarithmic function to any base}).$$

- (5) (i) *Remainder theorem*: When  $f(x)$  is divided by  $(x - a)$  the remainder is  $f(a)$ .  
 (ii) *Factor theorem*: If  $f(a) = 0$ ,  $(x - a)$  is a factor of  $f(x)$ .

## Methods

Simultaneous equations, and general algebraic manipulation have not been included in the fundamental algebraic methods that follow. However some methods, and engineering applications given after these methods, do require their use.

## Logarithms

Manipulation of expressions and formulae where the variable you require is part of an index, requires us to use the laws of logarithms to find the variable. For example, if

$a = b^c$  make  $c$  the subject of the formula.

Then from laws (i) and (iv), and taking logs to base  $b$  gives

$$\log_b a = c \log_b b$$

so  $\log_b a = c$  as required.

### Example 6.1.1

If  $U_2 = U_1 e^{\left(\frac{w}{pv}\right)}$  make  $w$  the subject of the formula then

$$\log_e \left( \frac{U_2}{U_1} \right) = \left( \frac{w}{pv} \right) \log_e e$$

(on division by  $U_1$  and applying logs)

$$\log \left( \frac{U_2}{U_1} \right) = \frac{w}{pv} \quad (\text{since } \log_e e = 1)$$

$$\text{and so } w = pv \log_e \left( \frac{U_2}{U_1} \right).$$

*Note*: It is normal to present the subject of a formula to the left of the equals sign.

Transposition of equations or expressions, when the new subject is part of a logarithmic expression, involves an extra step. This step is that of finding the antilogarithm of both sides of the equation.

### Example 6.1.2

Transpose  $b = \log_e t - a \log_e D$  to make  $t$  the subject:

$$b = \log_e t - \log_e D^a \quad (\text{law (iv)})$$

$$b = \log_e \left( \frac{t}{D^a} \right) \quad (\text{from law (iii)})$$

$$e^b = \frac{t}{D^a} \quad (\text{since } e^x \text{ is the inverse or antilogarithm of } \log_e x = \ln x)$$

$$t = e^b D^a.$$

Let us consider one more example where we are required to determine a numerical answer.

### Example 6.1.3

If  $\theta_f - \theta_i = \frac{R}{J} \log_e \left( \frac{U_2}{U_1} \right)$  find the value of  $U_2$  given that:

$$\theta_f = 3.5, \quad \theta_i = 2.5, \quad R = 0.315, \quad J = 0.4, \quad U_1 = 50.$$

Placing the values directly into the formula does not get us very far, because we will end up trying to take the logarithm, of an unknown.

Transposing gives:

$$\frac{J(\theta_f - \theta_i)}{R} = \log_e \left( \frac{U_2}{U_1} \right)$$

$$\text{and so } e^{\left( \frac{J(\theta_f - \theta_i)}{R} \right)} = \frac{U_2}{U_1} \quad (\text{antilogging})$$

then  $U_2 = U_1 e^{\left( \frac{J(\theta_f - \theta_i)}{R} \right)}$  which on substitution of the values gives  **$U_2 = 178$**  (check that you get this value).

### Exponential equations

These equations are simply an extension of our previous transposition methods where we used logarithms. They can sometimes be a little awkward, requiring a little algebraic manipulation.

### Example 6.1.4

Find the value of  $x$ , when

$$\log 8^x = \log 10$$

$$x \log 8 = \log 10 \quad (\text{using rule (iv) of logarithms})$$

$$x = \frac{\log 10}{\log 8} \quad (\text{and using calculator or tables})$$

$$\mathbf{x = 1.107} \quad \text{correct to three decimal places.}$$

**Example 6.1.5**

Solve the equation  $2(2^{2x}) - 5(2^x) + 2 = 0$  for  $x$ . Now at first glance taking logs appears to be no help! The trick is to simplify the equation using a substitution. In performing the substitution we need to be aware of the rules for indices as well as logs.

If we substitute  $y = 2^x$ , then we get:

$$2y^2 - 5y + 2 = 0 \quad (\text{where } 2^{2x} = 2^x \cdot 2^x = y \cdot y = y^2).$$

Now we solve a simple quadratic equation, which you should be familiar with. The simplest method is to factorise, if all else fails revert to the quadratic formula!

$$\text{Then } (2y - 1)(y - 2) = 0$$

$$\text{from which } 2y - 1 = 0 \quad \text{giving } y = \frac{1}{2}$$

$$\text{or } y - 2 = 0 \quad \text{giving } y = 2.$$

$$\text{So } 2^x = \frac{1}{2} \quad \text{or } 2^x = 2$$

then by inspection  $x = \pm 1$ .

If you could not see this straightaway look carefully at the rules given for indices, particularly where:

$$a^0 = 1 \quad \text{and} \quad \frac{1}{a^n} = a^{-n}.$$

**Example 6.1.6**

Solve the equations simultaneously for  $x$  and  $y$ :

$$\log x - 2 \log y = 1 \quad \text{(i)}$$

$$xy = 270. \quad \text{(ii)}$$

Then from the laws of logs (i) becomes

$$\log \frac{x}{y^2} = 1 \quad \text{and antilogging we get:}$$

$$\frac{x}{y^2} = 10 \quad (\text{logs to base 10})$$

so  $x = 10y^2$  (substituting in (ii)) we get:

$$10y^3 = 270$$

$$y^3 = 27$$

$$y = 3$$

Then  $x = 90, y = 3$ .

**Reduction of non-linear laws**

We will now look at a technique for reducing non-linear relationships to linear form. This process enables us to verify experimental laws and analyse experimental data in a convenient way. Two examples serve to illustrate the process.

### Example 6.1.7

The experimental values of  $p$  and  $s$  shown below, are believed to be related by the law  $s = ap^2 + b$ . By plotting a suitable graph verify this law and determine approximate values of  $a$  and  $b$ .

|     |      |       |       |      |      |
|-----|------|-------|-------|------|------|
| $p$ | 1    | 2     | 3     | 4    | 5    |
| $s$ | 6.73 | 11.92 | 20.57 | 32.7 | 48.3 |

If  $s$  is plotted against  $p$  a curve results and we are unable to determine the value of the constants  $a$  and  $b$ . By comparing the above relationship, with the law for a straight line where:

$$s = ap^2 + b \quad \text{and} \quad y = mx + c$$

it can be seen that a plot of  $s$  against  $p^2$ , should produce a straight line enabling us to estimate the constants  $a$ ,  $b$  and so verify or otherwise the suggested relationship.

To produce the graph (Figure 6.1.1) we draw up a table of values in the normal way.

|       |      |       |       |      |      |
|-------|------|-------|-------|------|------|
| $p$   | 1    | 2     | 3     | 4    | 5    |
| $p^2$ | 1    | 4     | 9     | 16   | 25   |
| $s$   | 6.73 | 11.92 | 20.57 | 32.7 | 48.3 |

The graph results in a straight line where slope (constant  $a$ ) is shown to be 1.73 and the intercept with the  $s$  axis, (constant  $b$ ) has a value close to 5.0.

So the law relating the data is  $s = 1.73p^2 + 5.0$  (you should verify that this is correct).

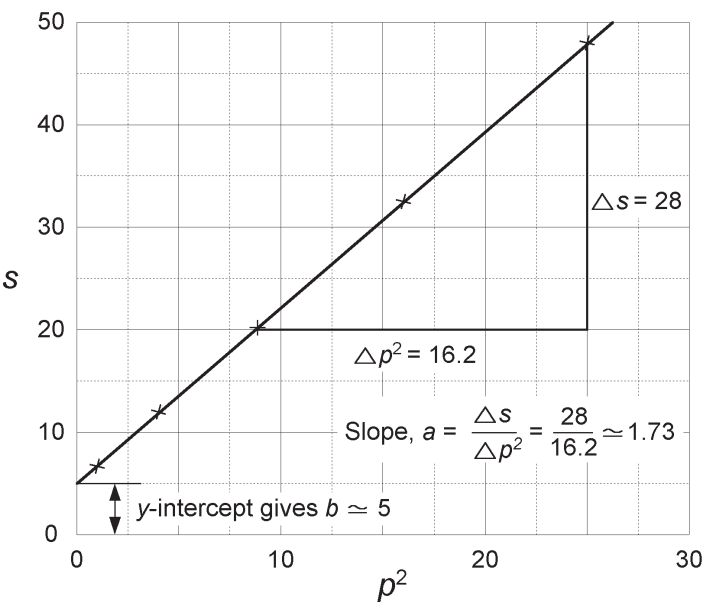


**Figure 6.1.1** Graph of  $s$  against  $p^2$

You should remember that the slope of the graph enables us to find the constant  $m$ , in the straight line law and, the constant  $c$  is found where the graph crosses the  $y$ -intercept at zero. Occasionally it is not convenient to start the graph at zero, under these circumstances, simultaneous equations may be used to determine the constants. One final example should ensure you are able to follow the method.

### Example 6.1.8

It is thought that the load ( $W$ ) and the effort ( $E$ ) for a machine are related by:

$$E = \frac{aW}{6} + b.$$

The results of an experiment carried out on the machine are given below. Determine the constants  $a$ ,  $b$  and verify the law of the machine.

|                |       |       |      |       |       |
|----------------|-------|-------|------|-------|-------|
| Load $W$ (N)   | 1000  | 2000  | 3000 | 4000  | 5000  |
| Effort $E$ (N) | 583   | 666   | 750  | 833   | 916   |
| $W/6$ (N)      | 166.7 | 333.3 | 500  | 666.7 | 833.3 |

Again, if we compare the supposed law of the machine with that of a straight line, then clearly all we need to do is plot effort  $E$  against  $W/6$  from which a straight line graph should result which verifies the law. The above table includes the row for  $E/6$ , so we can plot the graph (Figure 6.1.2) straightaway.

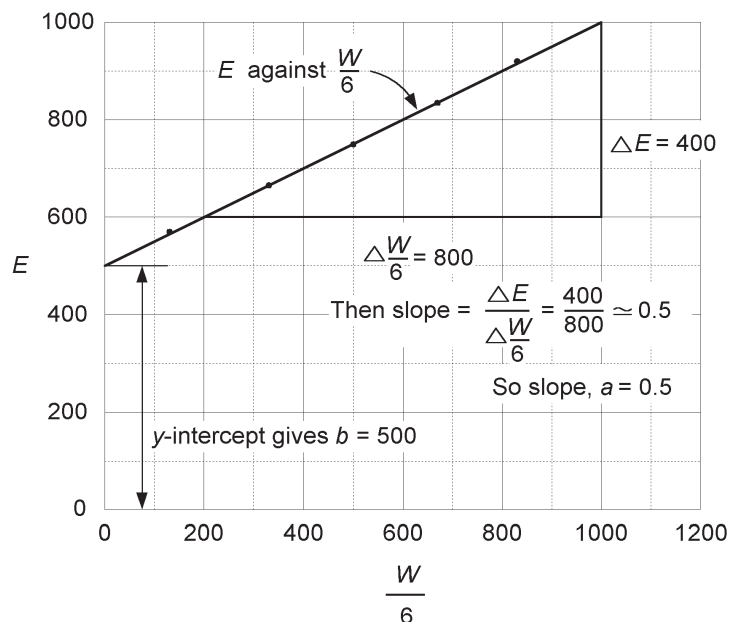


**Figure 6.1.2** Graph of  $W/6$  against  $E$



From the graph it can be seen that constant  $a = 0.5$  and constant  $b = 500$ . The law is verified because the graph yields a straight line and the law of the machine is given by:

$$E = \frac{0.5W}{6} + 500.$$

Any set of values may be used to check that you have determined the constants correctly.

The manipulation of experimental data and the verification of laws and other ideas, is an essential tool for experimental work. Sources of experimental error and error arithmetic are covered in the statistical applications section of this chapter. You should therefore study this section and become familiar with these topics if you have not already done so.

### **Engineering applications**

We now turn our attention to the application of fundamental algebra to engineering situations. You will see that a lot of your time is spent manipulating logarithmic expressions, solving simple equations and making substitutions. All the engineering applications are in the form of examples, with all necessary explanation forming part of the example.

#### **Example 6.1.9**

In information theory (Chapter 3) it can be shown that the information content of a message is given by  $I = \log_2(1/p)$ . Show using the laws of logs that the information content may be expressed as  $I = -\log_2(p)$  and find the information content of the message if the probability ( $p$ ) of receiving the code is  $1/16$ .

We need to show that:

$$I = \log_2\left(\frac{1}{p}\right) = -\log_2(p).$$

The left-hand side of the above expression may be written as  $\log_2(p^{-1})$  and comparing with law (iv) where,

$$\log_a(M^n) = n\log_a M$$

then in this case  $M = p$  and  $n = -1$  so

$$\log_2(p^{-1}) = -1 \log_2 p \text{ as required.}$$

Now substituting  $(1/16)$  into  $\log_2(1/p)$  gives  $\log_2(16)$ .

So using law (v) we get:

$$\log_2 16 = \frac{\log_{10} 16}{\log_{10} 2} = 4.$$

**Then the information content of the message = 4.**

**Example 6.1.10**

The Lamé equations relating radial and hoop stress in thick cylinders are given below (see Chapter 5.2). If the radial stress is  $\sigma_r = -24 \text{ N/mm}^2$  and the hoop stress  $\sigma_h = 54 \text{ N/mm}^2$ , when the internal radius  $r = 60 \text{ mm}$ . Find the value of the constants  $a$  and  $b$ .

$$\sigma_r = a - \frac{b}{r^2} \quad (1)$$

$$\sigma_h = a + \frac{b}{r^2}. \quad (2)$$

Now in order to find the constants  $a$  and  $b$ , we must first substitute the given values into Equations (1) and (2).

Then

$$-24 = a - \frac{b}{3600} \quad (1A)$$

$$54 = a + \frac{b}{3600}. \quad (2A)$$

So we need to solve these equations simultaneously. Multiplying both equations by 3600 to eliminate fractions gives:

$$-86\,400 = 3600a - b \quad (1B)$$

$$\text{and } 194\,400 = 3600a + b. \quad (2B)$$

‘Adding’ both equations eliminates  $b$  to give:

$$108\,000 = 7200a \quad \text{and} \quad a = 15.$$

Now substituting our value of ‘ $a$ ’ into any convenient equation (1B directly above) gives:

$$-86\,400 = (3600)(15) - b$$

$$\text{then } b = 140\,400.$$

So the required constants are  $a = 15$ ,  $b = 140\,400$ .

**Example 6.1.11**

In a belt and pulley drive system (Chapter 5.3) the relationships between belt pitch length ( $L$ ), distance between centres ( $C$ ), the pulley pitch diameters ( $D$ ), and the span length ( $S$ ) are given by the formulae below. Determine the belt pitch length and the span length for the system when  $D_1 = 0.2$ ,  $D_2 = 0.3$ , and  $B = 4.02$ .

$$L = 2C + 1.57(D_2 - D_1) + \frac{(D_2 - D_1)^2}{4C} \quad (1)$$

$$C = \frac{B + \sqrt{B^2 - 32(D_2 - D_1)^2}}{16} \quad (2)$$

$$S = \sqrt{C^2 - \left(\frac{D_2 - D_1}{2}\right)^2}. \quad (3)$$

We need to find the belt pitch length ( $L$ ) from Equation (1) above. Before we can do this we need to find the value of  $C$ . We can obtain  $C$  from Equation (2) since we are given  $B = 4.02$ .

Then

$$C = \frac{4.02 + \sqrt{(4.02)^2 - 32(0.3 - 0.2)^2}}{16}$$

$$C = 0.5$$

and substituting into Equation (1), to find  $L$  we have

$$L = (2)(0.5) + 1.57(0.1) + \frac{(0.1)^2}{(4)(0.5)}$$

$$\text{so } L = 1.162.$$

The span  $S$  is given by Equation (3) as:

$$S = \sqrt{(0.5)^2 - \left(\frac{0.1}{2}\right)^2}$$

$$\mathbf{S = 0.497.}$$

This is relatively simple, but beware of making arithmetic errors!

### Example 6.1.12

The resistors from two networks are connected by the equations given below. Use these equations to derive an equation for  $R_1$  and find its value when  $R_A = 20 \text{ k}\Omega$ ,  $R_B = 45 \text{ k}\Omega$ , and  $R_C = 15 \text{ k}\Omega$ .

$$R_1 + R_3 = \frac{R_A \times (R_B + R_C)}{R_A + R_B + R_C} \quad (1)$$

$$R_2 + R_3 = \frac{R_B \times (R_A + R_C)}{R_A + R_B + R_C} \quad (2)$$

$$R_1 + R_2 = \frac{R_C \times (R_A + R_B)}{R_A + R_B + R_C}. \quad (3)$$

By adding or subtracting the above equations from one another we should be able to leave  $R_1$  on its own on the right-hand side of an equation. So our first task is to eliminate  $R_2$  and  $R_3$  from the left-hand side of the above equations. If we subtract Equation (2) from (1), we first eliminate  $R_3$ .

So:

$$(R_1 + R_3) - (R_2 + R_3) = \left( \frac{R_A \times (R_B + R_C)}{R_A + R_B + R_C} \right) - \left( \frac{R_B \times (R_A + R_C)}{R_A + R_B + R_C} \right).$$

Now providing you know the rules for manipulating fractions, the above process is simple! We get:

$$R_1 - R_2 = \frac{R_A(R_B + R_C) - R_B(R_A + R_C)}{R_A + R_B + R_C}.$$

If you did not understand this step you need to revise your knowledge of fractions.

Writing the above a little more succinctly gives:

$$R_1 - R_2 = \frac{R_A(R_B + R_C) - R_B(R_A + R_C)}{R_A + R_B + R_C}. \quad (4)$$

Now I hope you can see that by adding Equations (3) and (4), we eliminate  $R_2$  and leave an expression for  $R_1$ , as follows:

$$(R_1 + R_2) + (R_1 - R_2) = \left( \frac{R_C \times (R_A + R_B)}{R_A + R_B + R_C} \right) + \left( \frac{R_A(R_B + R_C) - R_B(R_A + R_C)}{R_A + R_B + R_C} \right).$$

Then

$$2R_1 = \frac{R_C(R_A + R_B) + R_A(R_B + R_C) - R_B(R_A + R_C)}{R_A + R_B + R_C}$$

and simplifying gives:

$$2R_1 = \frac{R_C R_A + R_C R_B + R_A R_B + R_A R_C - R_B R_A - R_B R_C}{R_A + R_B + R_C}$$

$$\text{or } 2R_1 = \frac{R_A R_C + R_B R_C + R_A R_B + R_A R_C - R_A R_B - R_B R_C}{R_A + R_B + R_C}.$$

Note that putting multiplied terms into alphabetical order helps to identify like expressions.

$$\text{Then } 2R_1 = \frac{2R_A R_C}{R_A + R_B + R_C} \quad \text{or} \quad R_1 = \frac{R_A R_C}{R_A + R_B + R_C}$$

as required. Finally substituting the values gives:

$$R_1 = \frac{(20)(15)}{20 + 45 + 15} = 3.75\text{k}\Omega.$$

Compare the above example with the equations given for  $T$ ,  $\pi$ , star, and delta networks in Chapter 4.

**Example 6.1.13**

The formula set out below relates the power ( $P$ ), belt tension ( $T$ ), angle of lap ( $\theta$ ), linear velocity ( $v$ ), and the coefficient of friction ( $\mu$ ), for a flat belt drive system (Chapter 5.3). Determine an expression for the coefficient of friction and find its value when,  $P = 2500$ ,  $T = 1200$ ,  $v = 3$ , and  $\theta = 2.94$ .

$$P = T(1 - e^{-\mu\theta})v. \quad (1)$$

We are required to find an expression for ' $\mu$ '; in other words make ' $\mu$ ' the subject of the formula. We proceed as follows, from Equation (1).

$$\frac{P}{T} = (1 - e^{-\mu\theta})v \quad \text{or} \quad \frac{P}{Tv} = 1 - e^{-\mu\theta}$$

$$\text{and so} \quad \frac{P}{Tv} - 1 = -e^{-\mu\theta} \quad \text{or} \quad 1 - \frac{P}{Tv} = e^{-\mu\theta}.$$

Now taking logarithms to the base 'e' of both sides then

$$\ln\left(1 - \frac{P}{Tv}\right) = -\mu\theta.$$

$$\text{so} \quad \frac{-\ln\left(1 - \frac{P}{Tv}\right)}{\theta} = \mu.$$

Now substituting the given values, we get:

$$\begin{aligned} \frac{-\ln\left(1 - \frac{2500}{(1200)(3)}\right)}{2.94} &= \mu \\ \frac{-\ln(0.305)}{2.94} &= \mu \end{aligned}$$

**Coefficient of friction  $\mu = 0.4$ .**

There are many more engineering applications which require us to apply the fundamental rules of algebra, but space does not permit us to review them here. The problems set at the end of this section will give you the opportunity to put more of these basic algebraic skills into practice.

## Partial fractions

### Formulae

Provided that the numerator  $f(x)$  is second degree or lower

$$\frac{f(x)}{(x+a)(x+b)(x+c)} \equiv \frac{A}{(x+a)} + \frac{B}{(x+b)} + \frac{C}{(x+c)} \quad (1)$$

$$\frac{f(x)}{(x+d)(ax^2+bx+c)} \equiv \frac{A}{(x+d)} + \frac{Bx+C}{(ax^2+bx+c)} \quad (2)$$

$$\frac{f(x)}{(x+a)(x+b)} \equiv \frac{A}{(x+a)} + \frac{B}{(x+b)} + \frac{C}{(x+b)}. \quad (3)$$

## Methods

Relatively complex algebraic fractions can often be broken down into a series of smaller fractions which are much more easily manipulated, these are known as *partial fractions*, the formulae given above show examples. PFs, are very useful when trying to integrate complex expressions that are able to be broken down in this way. As mentioned previously, they are also useful when determining the inverse of non-standard LTs, which cannot be identified from tables.

You will be using PFs later, when we consider the calculus and the engineering applications of Laplace functions. In the mean time we will concentrate on the methods used to obtain them, covering their engineering applications at a later stage when we consider a number of more advanced topics.

The methods used for finding PFs, will depend on the types of factor that go to make up the original fraction.

### Linear factors

If the denominator consists only of linear factors, where the degree of the unknown is one, as in formula (1) above. Then for every linear factor, for example  $(x - a)$  in the denominator, there corresponds a PF of the form  $A/(x - a)$ . Then in order to find the PFs we need to find the values of the *coefficients* in the numerator. Remembering that in this instance a coefficient is any number which is used as a multiple of the unknown, for example in the expression  $2x$ , 2 is the coefficient of  $x$ . The following example illustrates the process for finding PFs, where only linear factors are involved. You will see later that the process developed here is similar to that required for determining all PFs, irrespective of their form.

#### Example 6.1.14

Find the PFs for the expression

$$\frac{3x}{(x-1)(x-2)(x-3)}.$$

From what has already been said, and following the format given in formula (1). Then we may write:

$$\frac{3x}{(x-1)(x-2)(x-3)} \equiv \frac{A}{(x-1)} + \frac{B}{(x-2)} + \frac{C}{(x-3)}.$$

This is the required form for *linear* factors. Now multiplying every term by  $(x-1)(x-2)(x-3)$ , that is the denominator and, cancelling where appropriate gives:

$$3x \equiv A(x-2)(x-3) + B(x-1)(x-3) + C(x-1)(x-2).$$

The sign  $\equiv$  means 'always equal to'.

Now we need to find values of the unknown 'x' which satisfies the above equation, in order to determine the coefficients  $A$ ,  $B$ , and  $C$ . This is, in fact, a lot easier than it looks! There are two methods often used, each has its merits. For linear factors the following simple short-cut method will be used. You will meet the other method next, when we consider quadratic factors.

All we need to do is substitute a value of 'x' into the bracketed expressions which, when possible, reduces them to zero. For example the brackets multiplied by the coefficient  $A$  are  $(x - 2)(x - 3)$  so choosing a value of  $x = 2$  or  $x = 3$  would reduce the whole expression, including the coefficient  $A$ , to zero. Whatever the value of 'x' is chosen, it must be applied to 'all x' throughout the whole expression.

So substituting  $x = 2$ , into the whole expression, gives:

$$6 \equiv 0 + B(2 - 1)(2 - 3) + 0$$

$$\text{then } 6 \equiv -B \quad \text{or} \quad B = -6.$$

We repeat this process by substituting appropriate values, then substituting  $x = 3$ , gives

$$9 \equiv 0 + 0 + C(3 - 1)(3 - 2)$$

$$\text{then } 9 \equiv 2C \quad \text{or} \quad C = 4.5.$$

One last substitution  $x = 1$ , gives

$$3 \equiv A(1 - 2)(1 - 3) + 0 + 0$$

$$\text{then } 3 \equiv 2A \quad \text{or} \quad A = 1.5.$$

So when  $A = 1.5$ ,  $B = -6$ , and  $C = 4.5$  we get:

$$\frac{3x}{(x - 1)(x - 2)(x - 3)} \equiv \frac{1.5}{(x - 1)} - \frac{6}{(x - 2)} + \frac{4.5}{(x - 3)}.$$

This solution can be checked by adding together all the PFs, which should result in the original expression.

Check:

$$\frac{1.5(x - 2)(x - 3) - 6(x - 1)(x - 3) + 4.5(x - 1)(x - 2)}{(x - 1)(x - 2)(x - 3)}$$

and on multiplication of brackets in the numerator

$$\begin{aligned} &= \frac{1.5x^2 - 7.5x + 9 - 6x^2 + 24x - 18 + 4.5x^2 - 13.5x + 9}{(x - 1)(x - 2)(x - 3)} \\ &= \frac{3x}{(x - 1)(x - 2)(x - 3)}. \end{aligned}$$

Correct, the left-hand side equals the right-hand side (LHS = RHS). Note the use of algebraic manipulation and fractions!

If the linear factors in the denominator are *repeated*, for example  $(x + b)^2$ , as in formula (3) above, then to every repeated factor there corresponds PFs to the value of the power of the repeated factor. So for example  $(x + b)^2$  in the denominator will have two partial factors of the form  $A/(x + b)$  and  $B/(x + b)^2$ . Similarly for factors like  $(x - a)^3$  we would have three PFs,  $A/(x - a)$ ,  $B/(x - a)^2$ , and  $C(x - a)^3$ .

So the expression  $\frac{(x - 2)}{(x - 1)(x - 3)(x + 1)^2}$  would have the following, corresponding PFs:

$$\frac{(x - 2)}{(x - 1)(x - 3)(x + 1)} \equiv \frac{A}{(x - 1)} + \frac{B}{(x - 3)} + \frac{C}{(x + 1)} + \frac{D}{(x + 1)}.$$

The coefficients in the above expression may be found in a similar way to the method described in Example 6.1.14, or by the method illustrated in Example 6.1.15 below.

### Example 6.1.15

Determine the PFs for the following expression:

$$\frac{1}{(x - 2)^2(x - 3)}.$$

In this case we have a 'repeated' linear factor, which requires two PFs. So following formula (3), above then

$$\frac{1}{(x - 2)^2(x - 3)} \equiv \frac{A}{(x - 2)} + \frac{B}{(x - 2)^2} + \frac{C}{(x - 3)}.$$

Now substituting appropriate values of  $x$ , as in our previous method would provide values of the coefficients very easily. However, there is another method known as '*equating coefficients*', which can be used independently, or in conjunction with method 1. Although rather more long winded, it has the advantage that it can *always* be applied. Method 1 becomes rather more hit and miss, when we deal with quadratic factors or higher. Here, we will combine method 1 with the new method of equating coefficients.

So clearing fractions we get:

$$1 \equiv A(x - 2)(x - 3) + B(x - 3) + C(x - 2)^2$$

$$\text{or } 1 \equiv A(x^2 - 5x + 6) + B(x - 3) + C(x^2 - 4x + 4). \quad (\text{i})$$

Substituting an appropriate value of ' $x$ ', as before say  $x = 2$  gives:

$$1 = -B \quad \text{or} \quad B = -1.$$

Now, we will equate coefficients. This requires us to set-up a very simple table containing the unknown in ascending powers, as follows.

Considering Equation (i) then the 'powers' of the unknown involved are:  $x^0$ ,  $x^1$ ,  $x^2$ . Note that  $x^0 = 1$  and this allows us to consider the *constants* in the equation, remembering,



for example, that  $6x^0 = 6 \times 1 = 6$ . So equating the coefficients of these powers of  $x$ , gives:

$$x^0: 1 = 6A - 3B + 4C \quad (1)$$

$$x^1: 0 = -5A + B - 4C \quad (2)$$

$$x^2: 0 = A + C \quad (3)$$

So for row (1) where we equate  $x^0$ , we note that on the right-hand side of the equation we have '1' =  $1 \times x^0$ , this is why it appears as a coefficient of  $x^0$  in (1). Also multiplying the first expression  $A(x^2 - 5x + 6)$  to get  $Ax^2 - 5Ax + 6A$ , shows why  $6A$ , appears as the coefficient of  $x^0$ . This process is repeated for coefficients of  $x^1$  and  $x^2$ , up to the highest power in the expression. What we are left with is a set of equations in which we know the value of  $B$ . So treating the equations simultaneously then, adding (1) to (2) gives:

$$1 = A - 2B \quad (\text{eliminating } C)$$

and since we know that  $B = -1$ , this immediately gives  $A$  as:

$$1 = A + 2 \quad \text{and} \quad A = -1.$$

Then from Equation (3)  $A = -C$  or  $C = 1$ .

Then:

$$\frac{1}{(x-2)^2(x-3)} \equiv -\frac{1}{(x-2)} - \frac{1}{(x-2)^2} + \frac{1}{x-3}.$$

You should check that the PFs are correct, by multiplication, as in the previous example.

### Quadratic factors

Formula (2) above, shows the way in which an expression involving linear and quadratic factors, should be split into PFs. To every *quadratic factor* like  $x^2 + ax + b$  there corresponds a PF  $(Cx + D)/(x^2 + ax + b)$ . Repeated quadratic factors require additional PFs in the same way as before. Thus a factor  $(ax^2 + ax + b)^2$  would require PFs  $(Cx + D)/(x^2 + ax + b)$  and  $(Ex + F)/(x^2 + ax + b)^2$ .

Our final example concerning PFs, illustrates the method you will need to adopt when dealing with both linear and quadratic factors, in the denominator of an expression. The system of linear equations which results from equating coefficients, is in this example, fairly easily solved. More complex systems require the use of matrix methods for their solution.

#### Example 6.1.16

Express  $\frac{x}{(x-1)^2(x^2+5x+1)}$  as a sum of PFs.

Then

$$\frac{x}{(x-1)^2(x^2+5x+1)} \equiv \frac{A}{(x-1)} + \frac{B}{(x-1)^2} + \frac{Cx+D}{x^2+5x+1}$$

and on multiplying every term by the left-hand side denominator and cancelling gives:

$$x = A(x-1)(x^2+5x+1) + B(x^2+5x+1) + (Cx+D)(x-1)^2$$

and after multiplying out brackets and simplifying we get:

$$x = A(x^3+4x^2-4x-1) + B(x^2+5x+1) + C(x^3-2x^2+x) + D(x^2-2x+1)$$

Now using the method of equating coefficients of different powers of 'x' on both sides of the equation, gives:

$$x^0: 0 = -A + B + D \quad (1)$$

$$x^1: 1 = -4A + 5B + C - 2D \quad (2)$$

$$x^2: 0 = +4A + B - 2C + D \quad (3)$$

$$x^3: 0 = A + C \quad (4)$$

We need to solve the above system of linear equations, to find the coefficients.

Starting with the simplest Equation (4) we see that  $A = -C$  and substituting  $A = -C$  into Equations (2) and (3), we get:

$$1 = -5A + 5B - 2D \quad (2')$$

$$0 = 6A + B + D \quad (3')$$

Now subtracting Equation (1) from (3')

$$\begin{array}{r} 0 = 6A + B + D \\ -(0 = -A + B + D) \\ \hline 0 = 7A \Rightarrow A = 0 \end{array} \quad \text{and from } A = -C, C = 0$$

Now substituting  $A = 0$ , into (2') and (3') gives:

$$1 = 5B - 2D$$

$$0 = B + D$$

Solving above simultaneous equations, we get:

$$1 = 5B - 2D$$

$$0 = 2B + 2D$$

$$\text{so } 1 = 7B \quad \text{and} \quad B = \frac{1}{7} \quad \text{and} \quad D = -\frac{1}{7}$$

Then required PFs are:

$$\frac{x}{(x-1)^2(x^2+5x+1)} = \frac{1}{7(x-1)^2} - \frac{1}{7(x^2+5x+1)}$$

Note that we could have found  $B$  initially using method (1). This would have simplified the manipulation of the equations.

If the numerator of an expression is of higher degree than the denominator then we have an *improper algebraic fraction*. To convert an improper fraction to a proper fraction, we need to divide the denominator into the

numerator, for algebraic expressions this requires the use of long division of algebra, with which you should be familiar.

## Series

As already mentioned, the algebraic manipulation of sequences and series has few direct applications to engineering, but does enable us to gain mathematical skills which will be of use later, when we study complex numbers, numerical integration, and Fourier analysis. For this reason only methods, not applications, will be covered in this section.

### Formulae

- (1) *Sigma notation*:  $\Sigma$  = 'sum of all terms such as' Range of summation depends on the highest and lowest terms above and below the sigma sign, respectively.

$$\sum_{n=1}^{n=\infty} a^n = \text{Sum of all terms } a^n \text{ between } n = 1 \text{ and } n = \infty$$

(an infinite series).

- (2) *Permutations and combinations*

$${}_nP_r = \frac{n!}{(n-r)!} \quad (\text{i})$$

$${}_nC_r = \frac{n!}{(n-r)! r!} = \binom{n}{r} \quad (\text{ii})$$

- (3) *Binomial theorem*

$$(a \pm b)^n = {}^nC_0 a^n + {}^nC_1 a^{n-1} b + {}^nC_2 a^{n-2} b^2 + \dots \quad (\text{i})$$

$$(a + b)^n = a^n + na^{n-1}b + \frac{n(n-1)a^{n-2}b^2}{2!} + \dots + b^n \quad (\text{ii})$$

$$(1 + x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \dots \quad (\text{iii})$$

If 'n' is a positive integer, the series terminates after (n + 1) terms and is valid for all x.

If 'n' is not a positive integer, the series terminates and is valid only for values of x such that  $-1 < x < 1$ .

- (4) *Hyperbolic functions*

$$e^x = \cosh x + \sinh x \quad (\text{i})$$

$$e^{-x} = \cosh x - \sinh x \quad (\text{ii})$$

$$\sinh x = \frac{e^x - e^{-x}}{2} \quad (\text{iii})$$

$$\cosh x = \frac{e^x + e^{-x}}{2} \quad (\text{iv})$$

$$\tanh x = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (\text{v})$$

$$\sinh^{-1}x = \ln\left(x + \sqrt{x^2 + 1}\right) \quad (\text{for all } x) \quad (\text{vi})$$

$$\cosh^{-1}x = \pm \ln\left(x + \sqrt{x^2 - 1}\right) \quad (|x| > 1) \quad (\text{vii})$$

$$\tanh^{-1}x = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) \quad (\text{viii})$$

### (5) Series

$$\text{valid for all } x \quad \exp x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad (\text{i})$$

$$\text{valid for } -1 < x \leq 1 \quad \ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \quad (\text{ii})$$

$$\text{valid for } -1 \leq x < 1 \quad \ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \dots \quad (\text{iii})$$

$$x \text{ in radians, valid for all } x \quad \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (\text{iv})$$

$$x \text{ in radians, valid for all } x \quad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \quad (\text{v})$$

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots \quad (\text{vi})$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots \quad (\text{vii})$$

## Methods

In this final section on algebra we will concentrate on exponential, logarithmic, and hyperbolic functions and their related series, in addition to investigating the concept of the limit and, the use of the binomial theorem as a mathematical tool.

You have already met the exponential function and its inverse (the natural logarithm), in our work on fundamentals. In this section we will be looking at their series and those for trigonometric and hyperbolic functions, to see how they interrelate. We start by considering a little notation.

### Sigma notation

The notation given in formula (1) above, requires a little explanation, particularly for those who may not be familiar with the way in which we can represent a series using this notation. Also, the following explanation should be useful for those who have not met the concept of a *limit*.

Using *Sigma* notation (symbol  $\Sigma$ ) for summing the terms of a series enables us to define the series more concisely.

The terms of a *sequence* when added form a *series*. For example the set of numbers: 1, 2, 4, 8, 16 are in a set order and form a sequence

because each number may be found from the previous number by applying an obvious rule. When this sequence of numbers is added a series is formed. This series *can be bounded* by stipulating the number of terms in the series using Sigma notation, that is:

$$\sum_{n=1}^{n=5} 2^n \quad \text{writing this out in full we get:}$$

$$\begin{aligned} \sum_{n=0}^{n=5} 2^n &= 2^0 + 2^1 + 2^2 + 2^3 + 2^4 + 2^5 \\ &= 1 + 2 + 4 + 8 + 16 + 32. \end{aligned}$$

So we have a succinct way of writing this series, which tells us the precise number of terms and in this case gives us a precise numerical value. It provides a lower boundary (1) and upper boundary (63) for the sum of the series.

Consider the following series, written in Sigma notation

$$\sum_{n=0}^{n=\infty} \frac{1}{2^n} \quad \text{Does this series have a limit?}$$

In other words does it have an upper and lower boundary. Expanding the series we get:

$$\begin{aligned} \sum_{n=0}^{n=\infty} \frac{1}{2^n} &= \frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} \dots \frac{1}{2^n} \\ \text{So } \sum_{n=0}^{n=\infty} \frac{1}{2^n} &= \frac{1}{1} + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} \dots \frac{1}{2^n}. \end{aligned}$$

I am sure you can see that if we sum just the first term of the series the lower bound would be 1. Also that as  $n \rightarrow \infty$  then  $\frac{1}{2^n} \rightarrow 0$ , so there must also be an upper bound, because successive terms approach zero.

The above series is known as *Geometric* because each successive term is obtained from the preceding term by multiplication of a common ratio. The common ratio ( $r$ ) for the above series is  $r = \frac{1}{2}$ . This series is known as a *convergent* series because  $r < 1$ .

The previous series  $\sum_{n=1}^{n=\infty} 2^n$  is known as a *divergent* series as written above it has no *upper bound*, the common ratio for successive terms is 2.

If a geometric series is *convergent* then an upper bound can be found from

$$S_{\infty} = \frac{a}{1-r} \quad \text{where } |r| < 1.$$

$|r|$  means the modulus of  $r$  or the positive value of  $r$ . Then for our series the sum to infinity  $S_{\infty}$  is equal to

$$S_{\infty} = \frac{1}{1 - \frac{1}{2}} = \frac{1}{\frac{1}{2}} = 2 \quad (\text{where } a = \text{the first term}).$$

So this series has a *limit*.

**Example 6.1.17**

Find the upper and lower limits of  $x$  so that the series

$$\sum_{n=0}^{\infty} \frac{(x-1)^n}{2^n} \quad \text{converges}$$

and determine the sum to infinity of the series when  $x = 2$ .

To see how the series might behave, we need to expand the series for the first few terms. Then

$$\sum_{n=0}^{\infty} \frac{(x-1)^n}{2^n} = 1 + \frac{x-1}{2} + \frac{(x-1)^2}{2^2} + \frac{(x-1)^3}{2^3} + \dots \quad (\text{i})$$

Then the common ratio ( $r$ ) for the series can be seen as  $(x-1)/2$ , so the condition for convergence is  $|r| < 1$  or in this case

$$\left| \frac{x-1}{2} \right| < 1 \quad (\text{ii})$$

then removing modulus sign, allows us to consider negative limit, that is:

$$-1 < \frac{x-1}{2} < 1$$

which on multiplication by 2 gives:

$$\begin{array}{lll} -2 < x-1 < 2 & \text{and} & x-1 < 2 \quad \text{means} \quad x < 3. \\ & & \text{and} \quad -2 < x-1 \quad \text{means} \quad -1 < x \end{array}$$

**So  $-1 < x < 3$  in other words  $x$  must lie between a lower value greater than  $-1$  and an upper value less than  $3$ .**

Then for  $x = 2$ , the series will converge. Using our formula

for sum to infinity  $S_{\infty} = \frac{a}{1-r}$  when  $x = 2$ ,  $r = \frac{1}{2}$  from (ii). So

$$S_{\infty} = \frac{1}{1-\frac{1}{2}} = 2. \quad (\text{Note that if you put } x = 2 \text{ into the series (i)})$$

we obtain the series  $1 + \frac{1}{2} + \frac{1}{4} \dots$ , etc. as before!)

**The binomial theorem**

Before we look at the methods for using the binomial theorem, we must be clear on how to use the combination formula 2(ii). This formula tells us how many combinations there are of selecting ( $r$ ) objects from ( $n$ ) objects. For *combinations* the selection is made *irrespective of order*, for *permutations*, formula 2(i) is the selection of ( $r$ ) objects from ( $n$ ) *in a specified order*.

We are primarily interested in combinations, because it is these that are used in the binomial theorem 3(i).

So the selection of four objects from six, *irrespective of order* is given by formula 2(ii) as:

$$\begin{aligned} {}^6C_4 &= \frac{n!}{(n-r)! r!} = \frac{6!}{(6-4)! 4!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1)(4 \times 3 \times 2 \times 1)} \\ &= \frac{720}{48} = 15. \end{aligned}$$

Note that a number followed by an exclamation mark (!) is a *factorial* number, that is a number which multiplies itself by all the numerals down to one, as shown in the example above.

We will now use the form for the binomial theorem given in formulae 3(i) and 3(ii).

### Example 6.1.18

Expand  $(x - y)^5$  using the binomial theorem, from formula 3(ii), we get:

$$\begin{aligned} (x - y)^5 &= x^5 + (5)(x)^4(-y)^1 + \frac{(5)(4)(x)^3(-y)^2}{2!} \\ &\quad + \frac{(5)(4)(3)(x)^2(-y)^3}{3!} + \frac{(5)(4)(3)(2)(x)^1(-y)^4}{4!} \\ &\quad + \frac{(5)(4)(3)(2)(1)(-y)^5}{5!} \end{aligned}$$

and simplifying gives:

$$(x - y)^5 = x^5 - 5x^4y + 10x^3y^2 - 10x^2y^3 + 5xy^4 - y^5.$$

The general interpretation of the binomial theorem considered in the previous example is restricted to an expansion which has integer powers of the index  $n$ . This restriction is removed when we consider the binomial expansion in *standard form*. If the expansion is modified to exclude the  $a^n$  term in formulae 3(i) and 3(ii), replacing it with the number 1. Then the theorem may be written to include values of the index which are negative, fractional, or both, provided certain limitations are observed.

So the standard binomial expansion may be written as:

$$(1 + x)^n = 1 + (n)(x) + \frac{(n)(n-1)(x)^2}{2!} + \frac{(n)(n-1)(n-2)(x)^3}{3!} + \dots$$

You will remember from our work on *limits*, that if we restrict values of  $x$  to  $-1 < x < +1$ , then the series will converge to a finite value. Under these circumstances the number of terms in the series is unrestricted for convergence. If the value of  $x$  is outside these limits, we still have

no restriction on the number of terms in the series but the series may diverge.

If the binomial expansion contains multiples of  $x$ , such as  $2x, 3x, \dots$  then the limits imposed on ' $x$ ' must be adjusted accordingly. For example:

$$\begin{aligned} \text{if } -1 < 2x < +1 \quad \text{then} \quad -\frac{1}{2} < x < \frac{1}{2} \\ \text{or if } -1 < 3x < +1 \quad \text{then} \quad -\frac{1}{3} < x < \frac{1}{3} \quad \text{and so on.} \end{aligned}$$

### Example 6.1.19

Expand  $(1 + 2x)^{-1/2}$  to four terms. Using 3(iii) we get:

$$\begin{aligned} (1 + 2x)^{-1/2} &= 1 + \left(-\frac{1}{2}\right)(2x) + \frac{\left(-\frac{1}{2}\right)\left(-\frac{1}{2} - 1\right)(2x)^2}{2!} \\ &\quad + \frac{\left(-\frac{1}{2}\right)\left(-\frac{1}{2} - 1\right)\left(-\frac{1}{2} - 2\right)(2x)^3}{3!} + \dots \\ &= 1 - x + \frac{3}{2}x^2 - \frac{5}{2}x^3 + \dots \end{aligned}$$

restriction  $-\frac{1}{2} < x < +\frac{1}{2}$  for convergence.

It can be seen from the Example 6.1.19 that a series can be produced for functions of the type  $(1 + x)^n$ , where  $n$  is a whole, fractional, or negative number, by using the binomial theorem to the required degree of accuracy. It is this use of the binomial, where we are able to find *approximations* for algebraic expressions in the form of a *power series* that will help us with our later work. Another, more subtle extension of the above approximation method may be used to calculate percentage changes.

### Example 6.1.20

Given that  $C = \sqrt{p}$ , find the percentage change in  $C$ , caused by a 3% change in  $p$ .

Then new value of

$$\begin{aligned} C &= \sqrt{1.03p} \\ &= \sqrt{p}(1 + 0.03)^{1/2} \quad (\text{laws of indices!}) \end{aligned}$$

and first approximation, using the binomial theorem (formula 3(iii)) gives:

$$\begin{aligned} C &= \sqrt{p}\left[1 + \left(\frac{1}{2}\right)(0.03) + \dots\right] \\ C &= \sqrt{p}[1 + 0.015] \\ C &= \sqrt{p}(1.015) \end{aligned}$$

therefore there will be a **1.5% increase in C**.



**Series**

We have already discussed the exponential and natural logarithm functions. If we use the binomial expansion we can derive a particularly useful series, the *Exponential series* (see Example 6.1.21). Estimates of exponential, logarithmic, and hyperbolic functions using *power series* (a series where each term is a simple power of the independent variable), are particularly useful when considering the calculus and other mathematical topics, which you will meet later.

**Example 6.1.21**

Expand  $\left(1 + \frac{x}{n}\right)^n$  using the binomial theorem and deduce what happens to the series as  $n \rightarrow \infty$ . Then

$$\begin{aligned}\left(1 + \frac{1}{n}x\right)^n &= 1 + x + \frac{n(n-1)}{2!}\left(\frac{1}{n}x\right)^2 \\ &\quad + \frac{n(n-1)(n-2)}{3!}\left(\frac{1}{n}x\right)^3 + \dots \\ &= 1 + x + \frac{\left(1 - \frac{1}{n}\right)}{2!}x^2 + \frac{\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)}{3!}x^3 + \dots\end{aligned}$$

Now you already know that for the terms  $\frac{1}{n}$ ,  $\frac{2}{n}$ , etc. as  $n \rightarrow \infty$ , these terms  $\rightarrow 0$ . Therefore the bracketed expressions, all approach 1, or in the *limit* equal 1.

$$\text{So } \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}x\right)^n = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

I hope you recognise this series as the *Exponential series*, that is formula 5(i), where

$$\exp(x) = e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

This series provides us with *the base* for natural or Napierian logarithms ( $\ln$  or  $\log_e$ ), when  $x = 1$ .

That is

$$\begin{aligned}e^1 &= 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} \dots \\ &= 1 + 1 + 0.5 + 0.16666' + 0.041666' + 0.008333' \\ e^1 &\approx 2.72 \quad \text{correct to two decimal places.}\end{aligned}$$

Continuing the series gives us a better and better approximation for 'e', correct to five decimal places

$$e \approx 2.71828.$$

**Example 6.1.22**

Formulae 5(ii) and 5(iii) give the power series for the natural logarithm functions  $\ln(1+x)$  and  $\ln(1-x)$ , respectively. Write down the expansion for the function:

$$\ln\left(\frac{1+x}{1-x}\right).$$

Remembering our laws of logarithms, in particular law (iii), where

$$\log_a \frac{M}{N} = \log_a M - \log_a N,$$

then the above may be written as:

$$\ln\left(\frac{1+x}{1-x}\right) = \log_e(1+x) - \log_e(1-x)$$

and since

$$\log_e(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} + \dots$$

$$\log_e(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5} - \dots$$

$$\text{then } \log_e(1+x) - \log_e(1-x) = 2x + \frac{2x^3}{3} + \frac{2x^5}{5} + \dots$$

$$\text{or } \ln\left(\frac{1+x}{1-x}\right) = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \dots\right).$$

Before we consider techniques for manipulating hyperbolic functions, it is important to understand what these functions represent. You are no doubt already familiar with the basic trigonometric functions: sine, cosine, and tangent, and will be aware that they can be represented by the radius of a circle of unit length, rotating anti-clockwise from the horizontal. As this radius rotates values for the basic trigonometric ratios can be determined by construction, using suitable axis graduated in radians. For this reason the trigonometric functions are sometimes known as *circular functions*. Now, if instead of a circle we use a hyperbola then, in a similar way, we can use this geometrical method to produce the hyperbolic sine, cosine, and tangent.

It is no coincidence that the power series for the sine and cosine functions, are very similar to their hyperbolic counterparts! Notice also, (formula 4(iii) and 4(iv)) that the hyperbolic functions can be represented in exponential form.

**Example 6.1.23**

Determine the value of  $\cosh 1.932$  using the definition given in formula 4(iv) and compare this value with that obtained by the first four terms of its power series (formula 5(vii)).

$$\cosh x = \frac{e^x + e^{-x}}{2} \quad \text{so} \quad \cosh(1.932) = \frac{e^{1.932} + e^{-1.932}}{2}$$

let  $y = e^{1.932}$  then  $\log_e y = 1.932$  and so  $y = 6.9033$  therefore

$$\frac{1}{y} = e^{-1.932} = 0.1449$$

then  $\cosh x = \frac{6.9033 + 0.1449}{2} = 3.524$  (three decimal places).

Now using

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \dots$$

$$\begin{aligned} \cosh(1.932) &= 1 + \frac{(1.932)^2}{2} + \frac{(1.932)^4}{24} + \frac{(1.932)^6}{720} + \dots \\ &= 1 + \frac{3.7326}{2} + \frac{13.9325}{24} + \frac{52.0047}{720} \\ &= 1 + 1.8663 + 0.58055 + 0.0722 \\ &= \mathbf{3.52.} \end{aligned}$$

An accuracy of three decimal places would require a minimum of five terms.

### Example 6.1.24

Find the value of  $\tanh^{-1} 0.825$  for *real* values of  $x$ . We could use formula 4(viii) and solve directly, but we will take a more convoluted route to incorporate a little algebraic manipulation! The trigonometric identity  $\frac{\sin x}{\cos x} = \tan x$  is well known,

in a similar manner  $\frac{\sinh x}{\cosh x} = \tanh x$

$$\text{or} \quad \frac{\frac{e^x - e^{-x}}{2}}{\frac{e^x + e^{-x}}{2}} = \left( \frac{e^x - e^{-x}}{2} \right) \left( \frac{2}{e^x + e^{-x}} \right) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

We can say therefore that

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4ix)$$

Now if we let  $x = \tanh^{-1} 0.825$ , then we find the value of  $x$  such that  $\tanh x = 0.825$ . Using 4(ix) we get:

$$\frac{e^x - e^{-x}}{e^x + e^{-x}} = 0.825$$

$$e^x - e^{-x} = 0.825(e^x + e^{-x})$$

$$e^x - 0.825e^x = e^{-x} + 0.825e^{-x}$$

$$0.175e^x = 1.825e^{-x} \quad (\text{and on multiplication by } e^x)$$

$$0.175(e^x)^2 = 1.825$$

$$(e^x)^2 = \frac{1.825}{0.175}$$

$$e^x = \sqrt{\frac{1.825}{0.175}} = \pm 3.229$$

( $e^x$  cannot be *negative* for real values of  $x$ ) so  $e^x = 3.229$ .

Therefore,  $x = \ln 3.229 = 1.172$  then  **$\tanh^{-1} 0.825 = 1.172$** .

Check this solution using formula 4(viii)!

There are many other useful results which can be acquired by manipulating the above functions and their associated series, we will return to these series later.

We leave our study of fundamental algebra with a number of problems which will enable you to practice the methods presented here, and in some cases, to apply them to engineering situations.

### Questions 6.1.1

- (1) The thermal resistance of a composite wall is related by the equation:

$$R_T = \frac{x_1}{k_1 A} + \frac{x_2}{k_2 A} + \frac{x_3}{k_3 A}.$$

Transpose the formula for 'A' and find the value of A when  $k_1 = k_2 = 0.2$ ,  $k_3 = 0.3$ ,  $x_1 = 40 \times 10^{-3}$ ,  $x_2 = 125 \times 10^{-3}$ ,  $x_3 = 60 \times 10^{-3}$ , and  $R_T = 625$ .

- (2) The heat flow rate through a cylindrical wall is given by:

$$Q = \frac{2k(t_1 - t_2)}{\ln(r_2/r_1)},$$

also the thermal resistance of a cylinder wall is given by:

$$R = \frac{\ln(r_2/r_1)}{2\pi k}.$$

Obtain a formula for 'R' in terms of Q, and the temperature difference ( $t_1 - t_2$ ) and find its value when  $t_1 = 250$ ,  $t_2 = 20$ , and  $Q = 16.3$ .

- (3) The Bernoulli equation may be written as:

$$\frac{P_1}{\gamma} + \frac{v_1^2}{2g} + h_1 = \frac{P_2}{\gamma} + \frac{v_2^2}{2g} + h_2$$

given that  $(h_1 - h_2) = 2$ ,  $(v_1^2 - v_2^2) = 8.4$ ,  $P_1 = 350$ ,  $\gamma = 10$ , transpose the formula in a suitable way to find the value of the pressure  $P_2$  ( $g = 9.81$ ).

- (4) The signal-to-noise ratio is given by the formula

$$\frac{S}{N} = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right)$$

where  $P$  is signal and noise power in watts. Determine

$P_{\text{noise}}$  given  $\frac{S}{N} = 20$  and  $P_{\text{signal}} = 2$ .

- (5) Values for a reciprocal motion problem result in the following simultaneous equation being formulated

$$12 = \omega \sqrt{r^2 - 0.08^2} \quad (1)$$

$$3 = \omega \sqrt{r^2 - 0.24^2} \quad (2)$$

Find the value for the radius ' $r$ '.

- (6) Solve the simultaneous equations

$$\log_x y = 2 \quad (1)$$

$$xy = 8 \quad (2)$$

- (7) It is believed that the law connecting two variables
- $H$
- and
- $V$
- is of the form
- $H = aV^n$
- , where
- $a$
- and
- $n$
- are constants. The results of an experiment for
- $V$
- and
- $H$
- are given below. By plotting a graph using suitable axes, find the law relating
- $H$
- and
- $V$
- .

|     |       |      |       |       |     |       |
|-----|-------|------|-------|-------|-----|-------|
| $V$ | 9     | 12   | 15    | 18    | 20  | 24    |
| $H$ | 36.45 | 86.4 | 168.8 | 291.6 | 400 | 691.2 |

- (8) Find the PFs for the expressions given below:

$$(i) \frac{1}{(x+2)(x-3)}$$

$$(ii) \frac{x}{(x+1)(x^2+2x+6)}$$

$$(iii) \frac{x^2}{(x-1)(x+2)^2}$$

- (9) The exponential function
- $f(x) = Ae^{-bx}$
- satisfies the conditions
- $f(0) = 2$
- and
- $f(1) = 0.5$
- . Find the constants
- $A$
- and
- $b$
- .
- 
- (10) If the general condition for 'convergence' of a series is

$$\lim_{n \rightarrow \infty} \left| \frac{u_{n+1}}{u_n} \right| < 1 \text{ where } u_n \text{ is } n\text{th term of a series.}$$

Determine whether or not the following series converge:

$$\sum_{n=1}^{\infty} \frac{1}{n^2+1} \quad (i)$$

$$\sum_{n=1}^{\infty} \frac{n}{n^2+1} \quad (ii)$$

- (11) Expand the following binomial expressions and extract their coefficients:

$$(a + b)^0, (a + b)^1, (a + b)^2, (a + b)^3, (a + b)^4, (a + b)^5$$

By placing the coefficients of each expansion sequentially one underneath the other (starting with the lowest), study the triangular pattern formed and determine the coefficients of  $(a + b)^6$ , *without* carrying out the expansion.

- (12) Given that  $Q = \frac{1}{R} \left( \frac{L}{C} \right)^{1/2}$ . Use the binomial theorem to

find the percentage change in  $Q$  caused by a 2% increase in  $L$ , a 3% decrease in  $R$  and a 4% decrease in  $C$ .

- (13) Find the first five terms in the expansion of  $(x^2 - 1)e^{-x}$ .  
 (14) Sketch the graphs of the hyperbolic functions given by formulae 4(iii), 4(iv), and 4(v), choosing a suitable scale. Comment on the characteristics of these functions when  $x = 0$ ,  $x = 1$ ,  $x \rightarrow -\infty$ , and  $x \rightarrow +\infty$ .  
 (15) Using appropriate formulae, find the value of  
 (i)  $\sinh x = 1.5$   
 (ii)  $\cosh x = 1.875$   
 (iii)  $\sinh^{-1} 1.375 = x$   
 (iv)  $\tanh x = 0.32$ .

## 6.2 TRIGONOMETRY

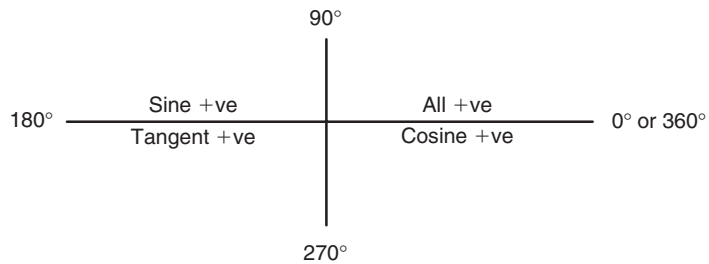
In this short section on trigonometry, we will treat the subject matter in an identical way to the algebra you meet previously. We start with a review of some fundamental trigonometric methods concerned with radian measure and basic trigonometric functions and their associated graphs. In the second part of this section we will concentrate on the identification and use of trigonometric identities to simplify expressions and solve trigonometric equations, which will be of use when we study some of the more advanced mathematical topics.

### Trigonometric fundamentals

#### Formulae

- (1) *Radian measure:*

- (i)  $\pi$  radians =  $180^\circ$ ,
- (ii) if  $\theta$  is *small* and measured in radians, then  $\sin \theta = \theta$ ,  $\cos \theta = 1$ ,  $\tan \theta = \theta$ ,
- (iii) arc length  $s = r\theta$  ( $\theta$  in radians),
- (iv) area of sector =  $\frac{1}{2}r^2\theta$  ( $\theta$  in radians).

(2) *Trigonometric ratios for angles of any magnitude:*

For all values of  $\theta$ :  $\sin(-\theta) = -\sin \theta$   
 $\cos(-\theta) = \cos \theta$   
 $\tan(-\theta) = -\tan \theta$

(3) *Equivalent ratios:*

| $90^\circ \leq \theta \leq 180^\circ$     | $180^\circ \leq \theta \leq 270^\circ$ | $270^\circ \leq \theta \leq 360^\circ$ |
|-------------------------------------------|----------------------------------------|----------------------------------------|
| $\sin \theta = \sin(180^\circ - \theta)$  | $-\sin(\theta - 180^\circ)$            | $-\sin(360^\circ - \theta)$            |
| $\cos \theta = -\cos(180^\circ - \theta)$ | $-\cos(\theta - 180^\circ)$            | $\cos(360^\circ - \theta)$             |
| $\tan \theta = -\tan(180^\circ - \theta)$ | $\tan(\theta - 180^\circ)$             | $-\tan(360^\circ - \theta)$            |

(4) *Triangle formulae:*

(i) Sine rule  $\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2r$

(ii) Cosine rule  $a^2 = b^2 + c^2 - 2bc \cos A$

(iii) Area  $= \frac{1}{2}bc \sin A$   
 $= \sqrt{s(s-a)(s-b)(s-c)}$  where  $s = \frac{(a+b+c)}{2}$ .

(5) *Polar and Cartesian co-ordinate system:*

(i)  $x = r \cos \theta$        $y = r \sin \theta$

(ii)  $r = +\sqrt{x^2 + y^2}$ ,  $\theta = \tan^{-1} \left[ \frac{y}{x} \right]$ ,  $\tan \theta = \frac{y}{x}$ .

(6) *Equation of a circle, centre at the origin, radius (a):*

$$x^2 + y^2 = a^2$$

$$x^2 + y^2 + 2gx + 2fy + c = 0$$

where the centre is at  $(-g, -f)$  and radius  $= \sqrt{g^2 + f^2 - c}$ .

(7) *Superposition:*

(i)  $a \sin \omega x + b \cos \omega x = c \sin(\omega x + \phi)$  where point  $(a, b)$  has polar co-ordinates  $r, \theta$  and  $c = r = \sqrt{a^2 + b^2}$  and  $\tan \phi = \frac{b}{a}$ .

**Methods**

We will look at a variety of methods using examples, as we did previously. These will cover radian measure, polar and Cartesian co-ordinates, graphs of sinusoidal functions, and the solution of triangles using trigonometric ratios.

### Radian measure

Circular measure using degrees has been with us since the days of the Babylonians, when they divided a circle into 360 equal parts, corresponding to what they believed were the days in a year. The *degree*, being an arbitrary form of circular measurement, has not always proved an appropriate unit for mathematical manipulation.

Another less arbitrary unit of measure has been introduced, which you will know as the *radian*, the advantage of this unit is its relationship with the arc length of a circle.

#### Example 6.2.1

An arc  $AB$  of length 4.5 cm is marked on a circle of radius 3 cm. Find the area of the sector bounded by this arc and the centre of the circle, using formulae 1(iii) and 1(iv).

Then arc length  $S = r\theta$

$$\theta = \frac{S}{r} = \frac{4.5}{3} = 1.5 \text{ rad}$$

So area of sector  $= \frac{1}{2}r^2\theta = (0.5)(3)^2(1.5)$  ( $\theta$  in radian)

**Area = 6.75 cm<sup>2</sup>.**

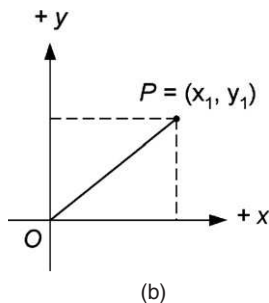
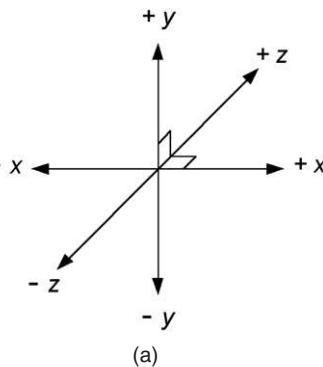


**Figure 6.2.1** Cartesian co-ordinates

#### Example 6.2.2

The area of a sector of a circle is 20 cm<sup>2</sup>. If the circle has a diameter of 9.5 cm. What is the length of the arc of the sector?

Then again from  $A = \frac{1}{2}r^2\theta$

$$\theta = \frac{2A}{r^2} = \frac{40}{4.75^2} = 1.77 \text{ rad}$$

So required arc length  $S = (4.75)(1.77)$

**S = 8.41 cm.**

### Polar and Cartesian co-ordinates

The ability to use both Polar and Cartesian co-ordinate systems and, to be able to change from one co-ordinate system to another is important for your later work, especially when we deal with complex numbers.

You will be familiar with the *Cartesian co-ordinate* system (Figure 6.2.1), where two or more mutually perpendicular axes are used to identify a point or position in space.

In Figure 6.2.1(b) the point  $P$  in the plane is located using the  $xy$  co-ordinates. The length of the line  $OP$  is found using Pythagoras, that is,  $(OP)^2 = x^2 + y^2$  so  $OP = \sqrt{x^2 + y^2}$ .

The position of the point  $P$  in Figure 6.2.1(b) is thus defined by the  $x$  and  $y$  values. Equally if the angle  $POQ$  is known in conjunction with the length of  $OP$ , then point  $P$  can be defined in terms of the  $x$ -axis only. We then have a *polar co-ordinate* system (Figure 6.2.2), where in this case  $OP = r$ .

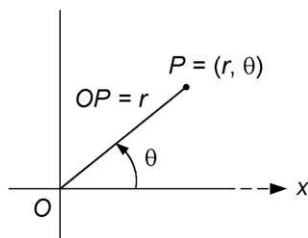


**Figure 6.2.2** Polar co-ordinates



**Example 6.2.3**

Convert the point (3, 4) to polar co-ordinates and convert the point (3, 30°) to Cartesian co-ordinates.

- (i) Using the formulae given in 5(ii), we have for point (3, 4),

$$r = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

$$\text{and } \theta = \tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}\left(\frac{4}{3}\right) = \tan^{-1} 1.333 = 53.13^\circ.$$

So point (3, 4) in polar form is **(5, 53.13°)**.

- (ii) Using the formulae given in 5(i), we have for point (3, 30°),

$$x = r \cos \theta = 3 \cos 30 = \mathbf{2.6}$$

$$\text{and } y = r \sin \theta = 3 \sin 30 = \mathbf{1.5}.$$

**Periodic functions**

Circular sinusoidal functions such as  $y = A \sin nx$  and  $y = B \cos nx$  may be represented graphically as shown in Figure 6.2.3. The important points to notice are:

- (i) These curves have a repeating pattern that occurs at period or wavelength  $2\pi/n$ .
- (ii) They have amplitude,  $A$  or  $B$ .
- (iii) The curves first reach their positive maximum value from a position or time zero at different points, this difference when measured as angular distance is known as *phase shift*. They are out of phase by some angular distance  $\phi$ .
- (iv) The general sinusoidal wave form may be written  $y = A \sin(nx + \phi)$ , where  $n$  is sometimes referred to as the *harmonic* and  $\phi$  is the *phase angle* which may lead or lag.

When the above graphs take the form  $y = A \sin \omega x$  or  $y = B \cos \omega x$  (formula 7(iv)), then if the variable ( $x$ ) is time, the time period is given by  $2\phi\omega$ , where  $\omega$  is the *angular frequency measured in rad/s* and the reciprocal of the time period  $\omega/2\pi$  is the *frequency measured in cycles per second* or

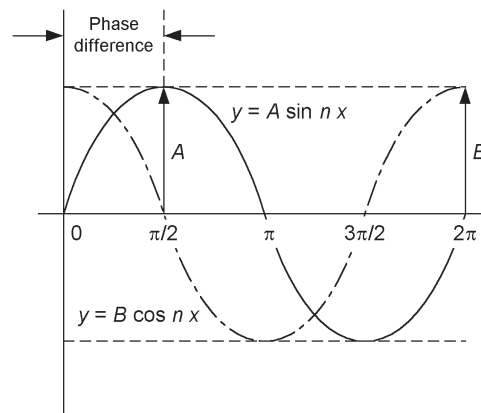


**Figure 6.2.3** Circular sinusoidal functions

*Hertz.* In this form we are able to use formula 7(i), which is known as the superposition formula, to sum two or more sinusoidal functions. This relationship is important in the study of waves and vibration analysis.

### Example 6.2.4

Find the sum of the two trigonometric functions  $y_1 = 4 \sin \omega x$ ,  $y_2 = 3 \cos \omega x$ , which both have the same period.

Then using the superposition formula

$$4 \sin \omega x + 3 \cos \omega x = c \sin(\omega x + \phi)$$

$$\text{where } c = \sqrt{4^2 + 3^2} = \sqrt{25} = 5$$

$$\text{and } \tan \phi = \frac{B}{A} = \frac{3}{4} \quad \text{giving } \theta = 36.9^\circ$$

$$\text{then } y_3 = 5 \sin(\omega x + 36.9).$$

### Example 6.2.5

- (i) Write down the amplitude, order of harmonic and phase angle of the sine function

$$y = 30 \sin(3\theta + 20^\circ)$$

then **Amplitude = 30**

**Harmonic = 3**

**Phase angle = 20° leading.**

- (ii) Find the instantaneous value of the sinusoidal function

$$y = 4 \sin \left( \omega t + \frac{\pi}{4} \right) \text{ given } \omega = 214 \text{ rad/s and } t = 7.5 \text{ ms}$$

$$y = 4 \sin \left( 1.605 + \frac{\pi}{4} \right) \quad \text{remember } 1.605 \text{ radians}$$

$$\text{so } y = 4 \sin(2.39 \text{ rad})$$

$$y = 4(0.6825)$$

$$y = 2.73.$$

### Solution of triangles

This is very much revision to remind you of the use of the fundamental ratios for the solution of triangles.

### Example 6.2.6

If the ratios of the sides of a triangle are such that  $\sin \theta = 3/5$  find the ratios for  $\cos \theta$  and  $\tan \theta$ , without referring to tables or using a calculator. We can apply the fundamental trigonometric identity:

$$\sin^2 \theta + \cos^2 \theta = 1 \quad \text{as follows}$$

$$\cos^2 \theta = 1 - \sin^2 \theta$$

$$\text{then } \cos \theta = \sqrt{1 - \sin^2 \theta}$$

$$\cos \theta = \sqrt{1 - \left(\frac{3}{5}\right)^2} = \sqrt{\frac{16}{25}} = \frac{4}{5}.$$

We also know that

$$\tan \theta = \frac{\sin \theta}{\cos \theta}$$

$$\text{therefore } \tan \theta = \frac{\frac{3}{5}}{\frac{4}{5}} = \frac{3}{4}.$$

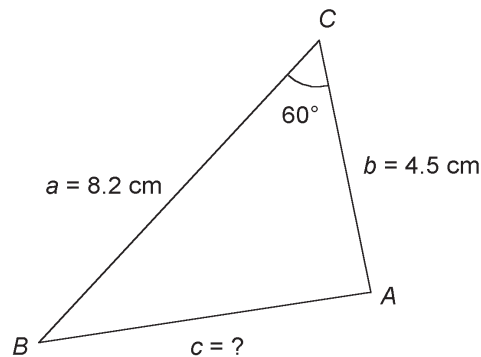
Then required ratios are:

$$\sin \theta = \frac{3}{5}, \cos \theta = \frac{4}{5} \text{ and } \tan \theta = \frac{3}{4}.$$

### Example 6.2.7

In the triangle  $ABC$ ,  $BC = 8.2$  cm,  $AC = 4.5$  cm and  $\angle ACB = 60^\circ$ . Calculate:

- (i) the other angles of the triangle;
- (ii) the area of the triangle.



First we sketch the triangle to establish the sides  $a$ ,  $b$ , and  $c$ . Then using the cosine rule (formula 4(ii)), then

$$c^2 = a^2 + b^2 - 2ab \cos C.$$

$$\text{So } c^2 = (8.2)^2 + (4.5)^2 - (2)(8.2)(4.5)\cos 60$$

$$c^2 = 67.24 + 20.25 - 36.9$$

$$c = 7.11 \text{ cm.}$$

Now using the sine rule

$$\frac{\sin A}{8.2} = \frac{\sin B}{4.5} = \frac{\sin 60}{7.11}$$

$$\text{so } \sin B = \frac{4.5 \sin 60}{7.11} = 0.5379$$

and  $\angle B = 32.54^\circ$  therefore  $\angle A = 87.46^\circ$

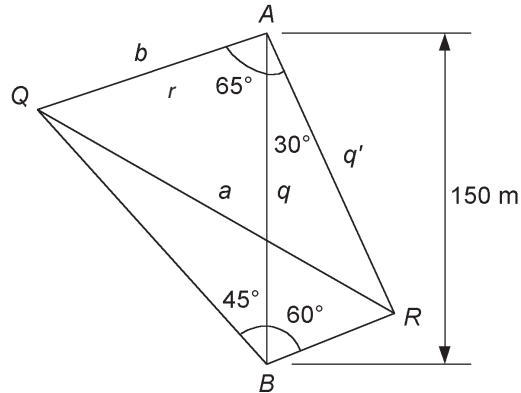
and then the area of the triangle  $= \frac{1}{2}ab \sin C$ .

$$\text{So area} = (0.5)(8.2)(4.5) \sin 60$$

$$\text{required area} = 15.98 \text{ cm}^2.$$

**Example 6.2.8**

The diagram below shows a plot to be used as the footprint for a factory site. Determine the length of the diagonal  $QR$ .



In triangle  $ARB$

$$\angle ARB = 90^\circ.$$

So length  $AR$  is given by:

$$\begin{aligned} AR &= 150 \cos 30^\circ \\ &= 129.9 \text{ m.} \end{aligned}$$

Also in triangle  $AQB$

$$\angle AQB = 70^\circ$$

$$\text{So } \frac{q}{\sin Q} = \frac{b}{\sin B} \quad \text{or} \quad \frac{150}{\sin 70} = \frac{b}{\sin 45}$$

$$\text{and } b = \frac{(150)(\sin 45)}{\sin 70}$$

$$\text{side } b = 112.9 \text{ m.}$$

Now considering triangle  $QAR$  and using cosine rule where sides are  $q, q', r$ , that is,  $AR = q', AQ = r, a = QR$

$$a^2 = q'^2 + r^2 - 2q'r \cos 95$$

$$a^2 = (129.9)^2 + (112.9)^2 - (2)(129.9)(112.9)(-0.087)$$

$$a^2 = 168\,74.01 + 127\,46.41 + 2551.8$$

$$\text{and } QR = a = 179.4 \text{ m.}$$

**Engineering applications**

The applications of trigonometric fundamentals are numerous. We will look specifically at the use of trigonometry for manipulating formulae concerned with shear strain and simple harmonic motion, in addition to AC and signal wave forms. As before, each application is given in the form of an example.

**Example 6.2.9**

An element of material is subject to shear strain (refer to Figure 5.1.8). Using the figure and the cosine rule for triangle

*dab*, derive the relationship:

$$1 + 2\varepsilon_{\max} + (\varepsilon_{\max})^2 = 1 + \sin\gamma.$$

Also, if  $\varepsilon_{\max}$  and  $\gamma$  are very small show that  $\varepsilon_{\max} = \frac{\gamma}{2}$ .

Applying the cosine rule as suggested to triangle *dab*, where the longest side is written in terms of  $s\sqrt{2}(1 + \varepsilon_{\max})$  then

$$[s\sqrt{2}(1 + \varepsilon_{\max})]^2 = s^2 + s^2 - 2s^2 \cos\left(\frac{\pi}{2} + \gamma\right)$$

$$\text{So } 2s^2(1 + \varepsilon_{\max})^2 = 2s^2 - 2s^2 \cos\left(\frac{\pi}{2} + \gamma\right)$$

which on division by  $2s^2$  gives

$$(1 + \varepsilon_{\max})^2 = 1 - \cos\left(\frac{\pi}{2} + \gamma\right).$$

Now using the trigonometric identity (formula 3) that  $\cos 90 + x = -\sin x$  then  $\cos \frac{\pi}{2} + \gamma = -\sin \gamma$ . So we get  $(1 + \varepsilon_{\max})^2 = 1 + \sin \gamma$  and on expansion of the left-hand side:

$$1 + 2\varepsilon_{\max} + (\varepsilon_{\max})^2 = 1 + \sin \gamma \quad \text{as required:}$$

Also if  $\varepsilon_{\max}$  and  $\gamma$  are very small then  $\sin \gamma \simeq \gamma$  (formula 1(ii)). We may also ignore the squared terms, that is,  $(\varepsilon_{\max})^2$ ,

$$\text{then } 1 + 2\varepsilon_{\max} = 1 + \gamma$$

$$\text{and } \varepsilon_{\max} = \frac{\gamma}{2} \quad \text{as required.}$$

In the next example we use the trigonometric identity you have already met, that is,  $\sin^2 \theta + \cos^2 \theta = 1$ .

### Example 6.2.10

In simple harmonic motion the variation of velocity of a point *Q* is given by  $Q = -\omega r \sin \omega t$ . With reference to Figure 3.2.12, show that  $Q = \pm \omega \sqrt{r^2 - x^2}$ . You need to know that  $\theta = \omega t$  and that the linear velocity *v* is equal to the product of the angular velocity  $\omega$  and its radius from the centre, *r*. So  $v = \omega r$ .

$$\text{Then } Q = -\omega r \sin \omega t$$

$$\text{so } Q = -\omega r \sin \theta \quad \text{and from } \sin^2 \theta = 1 - \cos^2 \theta$$

$$\text{then } Q^2 = \omega^2 r^2 (1 - \cos^2 \theta)$$

$$\text{or } Q = \pm \omega r \sqrt{1 - \cos^2 \theta}$$

Now  $\cos \theta = \frac{x}{r}$  (from Figure 3.2.12a)

$$\text{so } Q = \pm \omega r \sqrt{1 - \left(\frac{x}{r}\right)^2} \quad \text{and placing } r \text{ inside}$$

$$\text{gives } Q = \pm \omega \sqrt{r^2 - x^2}.$$

Note that in the previous example you should also be able to obtain the original expression for the velocity of  $Q$ , from the information given in Figure 3.2.12. The component of  $P$ 's velocity which is parallel and equal to  $OQ$  is given in Figure 3.2.12c as,  $\theta = -\omega r \sin \theta$ , make sure you can derive this from the geometry of the situation. The minus sign is present, because velocity is a vector quantity (speed in a given direction) and acts in opposition to the velocity deemed to be positive.

In the following example we look at a method for graphically adding two sinusoidal waveforms and analysing the resultant waveform.

### Example 6.2.11

Sketch the graphs of

$$y_1 = 10 \sin \left( 100\pi t + \frac{\pi}{4} \right)$$

$$\text{and } y_2 = 5 \sin \left( 200\pi t - \frac{2\pi}{3} \right)$$

on the same axis and hence sketch the graph of:

$$y_3 = 10 \sin \left( 100\pi t + \frac{\pi}{4} \right) + 5 \sin \left( 200\pi t - \frac{2\pi}{3} \right).$$

In order to sketch the resultant graph ( $y_3$ ), the following procedure may help. Choose points:

- (i) Where the original two graphs cross.
- (ii) Where there is a maximum and minimum value for each waveform.
- (iii) Where each waveform crosses the  $x$  axis.

Now the format for each of the waves is slightly different because the velocity is given in cycle/s not rad/s. So general form is:

$$y = a \sin(2\pi f t + \phi) \quad \text{where } 2\pi f = \text{frequency}$$

in cycles per second. Taking  $y_1 = 10 \sin \left( 100\pi t + \frac{\pi}{4} \right)$  then the frequency  $f = 50$  Hz (cycles per second). Therefore, the time period  $t = 1/50$  s = 20 ms (that is the time to complete

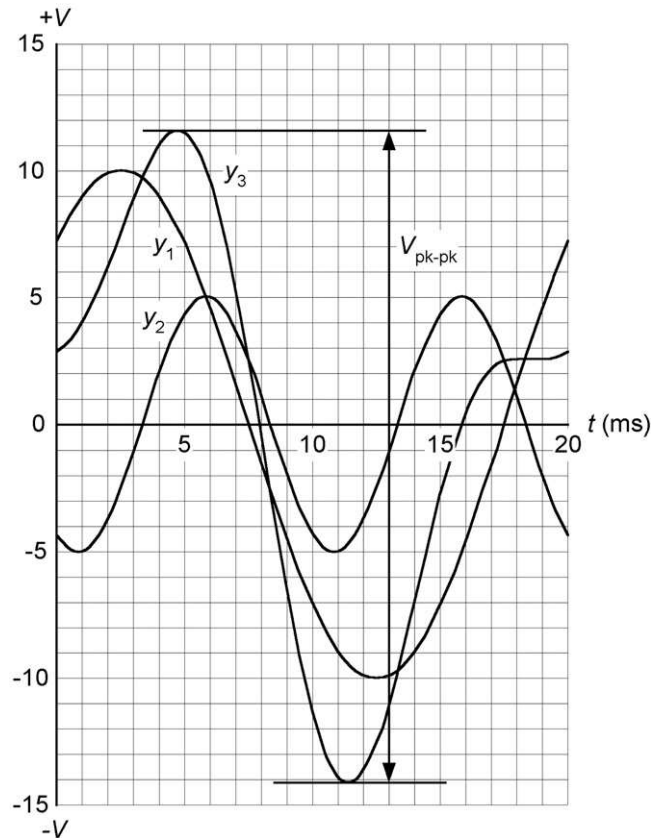


**Figure 6.2.4** Graphical solution

one cycle). We already know that the constant  $a$  = the amplitude, which in the case of  $y_1$ ,  $a = 10$ . So we have amplitude = 10, frequency ( $f$ ) = 50 Hz and the periodic time  $t = 20$  ms (milliseconds).

In a similar manner for  $y_2$  we have amplitude = 5, frequency = 100 Hz, periodic time =  $1/100 = 10$  ms. Now using these values we are able to set-up the axis of our graph, where  $y$  is amplitude (which may be voltage, current or power) and the  $x$  axis is time (in this case).

The appropriate waveforms are shown in Figure 6.2.4. The waveform for  $y_3$  is sketched by graphically adding  $y_1$  and  $y_2$  as suggested earlier. If the  $y$  axis is graduated in volts, you can see that the peak-to-peak voltage ( $V_{pp}$ ) for  $y_3$  approximately = 26 V.

You will find this technique very useful when looking at AC circuits and determining complex waveforms, which may be other than sinusoidal.

The final example in this section is rather contrived, and is concerned with determining the length of the bracing members used for the jib of a crane. It does, however, give us the opportunity to use the sine rule.

**Example 6.2.12**

Figure 6.2.5 shows the jib of a crane, consisting of three members plus the pulley wire supporting the load. Find the length of the strut  $AC$  and determine its angle of inclination *from* the vertical.

Now we have two sides and an included angle so we need to use the cosine rule, then

$$b^2 = a^2 + c^2 - 2ac \cos B \quad \text{which gives}$$

$$b^2 = (7.5)^2 + (3.5)^2 - (2)(7.5)(3.5) \cos 125^\circ$$

$$b^2 = 56.25 + 12.25 + 30.0825$$

$$b = 9.93 \text{ m.}$$

The angle of inclination  $\theta$  may now be found using the sine rule.

$$\frac{7.5}{\sin A} = \frac{9.93}{\sin 125^\circ} \quad \text{or} \quad \frac{(7.5)(\sin 125^\circ)}{9.93} = \sin A.$$

$$\sin A = 0.6727 \quad \text{and} \quad \text{so } A = \theta = 42.28^\circ.$$

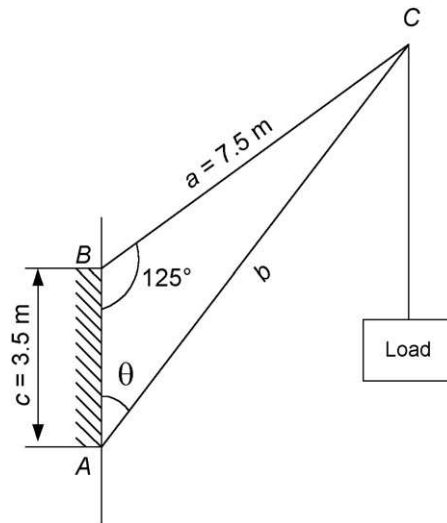


**Figure 6.2.5** Crane jib

There are numerous applications of fundamental trigonometry to engineering situations, space has permitted us to look at just a few. You will be able to gain further practice by attempting the problems which will be found at the end of this section on trigonometry.

## Trigonometric identities

### Formula

(1) General identities:

$$(i) \quad \operatorname{cosec} \theta = \frac{1}{\sin \theta}, \quad \sec \theta = \frac{1}{\cos \theta}, \quad \cot \theta = \frac{1}{\tan \theta}.$$



$$(ii) \tan \theta = \frac{\sin \theta}{\cos \theta}, \quad \sin^2 \theta + \cos^2 \theta = 1.$$

$$(iii) \tan^2 \theta + 1 = \sec^2 \theta, \quad \cot^2 \theta + 1 = \operatorname{cosec}^2 \theta.$$

$$(iv) \sin(A \pm B) = \sin A \cos B \pm \cos A \sin B.$$

$$(v) \cos(A \pm B) = \cos A \cos B \pm \sin A \sin B.$$

$$(vi) \tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B}.$$

(2) *Products to sums:*

$$(i) \sin A \cos B = \frac{1}{2} [\sin(A + B) + \sin(A - B)].$$

$$(ii) \cos A \sin B = \frac{1}{2} [\sin(A + B) - \sin(A - B)].$$

$$(iii) \cos A \cos B = \frac{1}{2} [\cos(A + B) + \cos(A - B)].$$

$$(iv) \sin A \sin B = -\frac{1}{2} [\cos(A + B) - \cos(A - B)].$$

(3) *Sums to products:*

$$(i) \sin A + \sin B = 2 \sin \frac{A+B}{2} \cos \frac{A-B}{2}.$$

$$(ii) \sin A - \sin B = 2 \cos \frac{A+B}{2} \sin \frac{A-B}{2}.$$

$$(iii) \cos A + \cos B = 2 \cos \frac{A+B}{2} \cos \frac{A-B}{2}.$$

$$(iv) \cos A - \sin B = -2 \sin \frac{A+B}{2} \sin \frac{A-B}{2} \quad \text{where } A > B.$$

(4) *Doubles and squares:*

$$(i) \sin 2A = 2 \sin A \cos A.$$

$$(ii) \cos 2A = \cos^2 A - \sin^2 A = 2 \cos^2 A - 1 = 1 - 2 \sin^2 A.$$

$$(iii) \tan 2A = \frac{2 \tan A}{1 - \tan^2 A}.$$

$$(iv) \cos^2 A + \sin^2 A = 1.$$

$$(v) \sec^2 A = 1 + \tan^2 A$$

$$(vi) \operatorname{cosec}^2 A = 1 + \cot^2 A.$$

(5) *Hyperbolic doubles and squares:*

$$(i) \sinh 2A = 2 \sinh A \cosh A.$$

$$(ii) \cosh 2A = \cosh^2 A + \sinh^2 A = 2 \cosh^2 A - 1 = 1 + 2 \sinh^2 A.$$

$$(iii) \tanh 2A = \frac{2 \tanh A}{1 + \tanh^2 A}.$$

$$(iv) \cosh^2 A - \sinh^2 A = 1.$$

$$(v) \operatorname{sech}^2 A = 1 - \tanh^2 A.$$

$$(vi) \operatorname{cosech}^2 A = \coth^2 A - 1.$$

(6) *Negatives:*

- (i)  $\sin(-A) = -\sin A.$
- (ii)  $\cos(-A) = \cos A.$
- (iii)  $\tan(-A) = -\tan A.$
- (iv)  $\sinh(-A) = -\sinh A.$
- (v)  $\cosh(-A) = \cosh A.$
- (vi)  $\tanh(-A) = -\tanh A.$

(7) *Halves:*

When  $t = \tan \frac{\theta}{2}$  then:

- (i)  $\sin \theta = \frac{2t}{1+t^2}.$
- (ii)  $\cos \theta = \frac{1-t^2}{1+t^2}.$
- (iii)  $\tan \theta = \frac{2t}{1-t^2}.$

When  $t = \tanh \frac{\theta}{2}$  then:

- (iv)  $\sinh Q = \frac{2t}{1-t^2}.$
- (v)  $\cosh Q = \frac{1+t^2}{1-t^2}.$
- (vi)  $\tanh Q = \frac{2t}{1+t^2}.$

## Methods

We now look at ways in which expressions can be manipulated and simplified using trigonometric identities. One or two simple identities were needed to solve the application examples given previously. Our study here, is concerned with the simplification and rearrangement of expressions in order to obtain solutions using the calculus. Their engineering applications will form part of our later study.

### Example 6.2.13

Solve the following trigonometric equations

- (i)  $4 \sin^2 \theta + 5 \cos \theta = 5.$
- (ii)  $3 \tan^2 \theta + 5 = 7 \sec \theta.$

(i) The most difficult problem when manipulating identities, is to know where to start! In this equation, we have two unknowns (sine and cosine) so the most logical approach is to try and get the equation in terms of one unknown, this leads us to the use of an appropriate identity. We can in this case use our old friend,

$$\sin^2 \theta + \cos^2 \theta = 1 \quad \text{from which} \quad \sin^2 \theta = 1 - \cos^2 \theta$$

which when substituted into equation (i) gives

$$4(1 - \cos^2 \theta) + 5 \cos \theta = 5$$

$$\text{or} \quad 4 - 4 \cos^2 \theta + 5 \cos \theta = 5.$$

This is a quadratic expression, which may be solved in a number of ways but it helps if you can factorise!

$$\text{Then} \quad (-4 \cos \theta + 1)(\cos \theta - 1) = 0$$

$$\Rightarrow -4 \cos \theta = -1 \quad \text{or} \quad \cos \theta = 1$$

$$\Rightarrow \cos \theta = \frac{1}{4} \quad \text{or} \quad \cos \theta = 1$$

$$\text{so} \quad \theta = 75.5^\circ \quad \text{or} \quad 0^\circ.$$

(ii) Proceeding in a similar manner to (i), we need a trigonometric identity which relates  $\tan \theta$  and  $\sec \theta$  (look at formula 1(iii)).

$$\text{then} \quad 3 \tan^2 \theta + 5 = 7 \sec \theta$$

$$\text{and using} \quad \sec^2 \theta = 1 + \tan^2 \theta \quad \text{or} \quad \tan^2 \theta = \sec^2 \theta - 1$$

$$\text{we get} \quad 3(\sec^2 \theta - 1) + 5 = 7 \sec \theta$$

$$\Rightarrow 3 \sec^2 \theta - 3 + 5 = 7 \sec \theta$$

$$\Rightarrow 3 \sec^2 \theta - 7 \sec \theta + 2 = 0$$

(again a quadratic expression) factorising gives

$$(3 \sec \theta - 1)(\sec \theta - 2) = 0$$

$$\text{so} \quad 3 \sec \theta = 1 \quad \text{or} \quad \sec \theta = 2$$

$$\left( \text{remembering that } \sec \theta = \frac{1}{\cos \theta} \right)$$

$$\sec \theta = \frac{1}{3} \quad \text{or} \quad \sec \theta = 2$$

$$\text{so} \quad \cos \theta = 3 \quad \text{or} \quad \cos \theta = \frac{1}{2}$$

$$\cos \theta = 3 \quad (\text{not permissible}) \quad \text{so there is only one solution}$$

$$\cos \theta = 0.5 \quad \text{so} \quad \theta = 60^\circ.$$

Example 6.2.14, is intended to show one or two techniques that may be used to verify trigonometric identities involving compound angle and double angle formulae. Example 6.2.15 shows how trigonometric identities may be used for evaluating trigonometric ratios.

### Example 6.2.14

Verify the following identities by showing that each side of the equation is equal in all respects:

$$(i) \quad (\sin \theta + \cos \theta)^2 \equiv 1 + \sin 2\theta$$

$$(ii) \frac{\sin 3\theta - \sin \theta}{\cos \theta - \cos 3\theta} \equiv \cot 2\theta.$$

(i) Simply requires the LHS to be manipulated algebraically to equal the RHS. So multiplying out gives

$$(\sin \theta + \cos \theta)^2 \equiv 1 + \sin 2\theta$$

$$\sin^2 \theta + 2 \sin \theta \cos \theta + \cos^2 \theta \equiv$$

$$\sin^2 \theta + \cos^2 \theta + 2 \sin \theta \cos \theta \equiv (\text{and from } \sin^2 \theta + \cos^2 \theta = 1)$$

$$\text{then } 1 + 2 \sin \theta \cos \theta \equiv (\text{and from formula 4(i)})$$

$$1 + \sin 2\theta \equiv 1 + \sin 2\theta \text{ as required.}$$

(ii) Again considering LHS and using sums to products (formulae 3)

$$\text{where (3ii)} = \sin A - \sin B = 2 \cos \frac{A+B}{2} \sin \frac{A-B}{2}$$

$$(3iv) = \cos A - \cos B = -2 \sin \frac{A+B}{2} \sin \frac{A-B}{2}$$

where  $A > B$  then

$$\sin 3\theta - \sin \theta = 2 \cos \left( \frac{3+\theta}{2} \right) \sin \left( \frac{3-\theta}{2} \right) \theta = 2 \cos 2\theta \sin \theta$$

and

$$\cos \theta - \cos 3\theta = -2 \sin \left( \frac{1-3}{2} \right) \theta \sin \left( \frac{1+3}{2} \right) \theta$$

$$\cos \theta - \cos 3\theta = -2 \sin(-\theta)(\sin 2\theta)$$

$$\text{and from formula 6(i) } \sin(-\theta) = -\sin \theta$$

$$\text{so } \cos \theta - \cos 3\theta = 2 \sin 2\theta \sin \theta$$

$$\text{therefore } \frac{\sin 3\theta - \sin \theta}{\cos \theta - \cos 3\theta} \equiv \frac{2 \cos 2\theta \sin \theta}{2 \sin 2\theta \sin \theta} \equiv \cot 2\theta.$$

### Example 6.2.15

If  $A$  is an acute angle and  $B$  is obtuse, where  $\sin A = \frac{3}{5}$  and  $\cos B = -\frac{5}{13}$ , find the values of:

$$(i) \sin(A + B)$$

$$(ii) \sin 2A$$

$$(iii) \tan(A + B)$$

$$(i) \sin(A + B) = \sin A \cos B + \cos A \sin B. \quad (1)$$

In order to use this identity we need to find the ratios for  $\sin B$  and  $\cos A$ . So again we need to choose an identity that allows us to find  $\sin \theta$  or  $\cos \theta$ , in terms of each other.

We know that  $\sin^2 B + \cos^2 B = 1$

$$\text{so } \sin^2 B = 1 - \cos^2 B$$

$$= 1 - \left(-\frac{5}{13}\right)^2$$

$$\sin^2 B = 1 - \frac{25}{169} = \frac{144}{169}$$

$$\sin B = \frac{12}{13}$$

(since  $B$  is obtuse  $90^\circ < B < 180^\circ$  and sign ratio is positive in second quadrant. Then only positive values of ratio need be considered). Similarly:

$$\sin^2 A + \cos^2 A = 1$$

$$\text{so } \cos^2 A = 1 - \sin^2 A$$

$$= 1 - \frac{9}{25}$$

$$\cos A = \frac{4}{5}$$

(since angle  $A$  is  $< 90^\circ$ , that is acute only the positive value is considered).

Now using equation (1) above

$$\sin(A + B) = \sin A \cos B + \cos A \sin B$$

$$= \left(\frac{3}{5}\right)\left(-\frac{5}{13}\right) + \left(\frac{4}{5}\right)\left(\frac{12}{13}\right)$$

$$= -\frac{15}{65} + \frac{48}{65}$$

$$\text{then } \sin(A + B) = \frac{33}{65}.$$

Note the use of fractions to keep exact ratios!

(ii) Now to find the value for  $\sin(2A)$ , we again need an identity which relates 'double' and single angles formula 4(i) provides just such a relationship.

$$\sin 2A = 2 \sin A \cos A$$

(for which we already have values)

$$\text{then } \sin 2A = 2\left(\frac{3}{5}\right)\left(\frac{4}{5}\right) = \frac{24}{25}.$$

(iii) For this part of the question we simply need to remember that  $\frac{\sin A}{\cos A} = \tan A$  and use formula 1(vi) in a similar way as before.

$$\text{then } \tan A = \frac{\sin A}{\cos A} = \frac{\frac{3}{5}}{\frac{4}{5}} = \left(\frac{3}{5}\right)\left(\frac{5}{4}\right) = \frac{3}{4}$$

$$\text{and } \tan B = \frac{\sin B}{\cos B} = \frac{\frac{12}{13}}{-\frac{5}{13}} = \left(\frac{12}{13}\right)\left(-\frac{13}{5}\right) = -\frac{12}{5}$$

and using formula 1(vi),

$$\tan(A + B) = \frac{\tan A + \tan B}{1 - \tan A \tan B} = \frac{\frac{3}{4} - \frac{12}{5}}{1 - \left(\frac{3}{4}\right)\left(-\frac{12}{5}\right)}$$

multiplying *every* term in the expression by 20 (the lowest common multiple), we get

$$\tan(A + B) = \frac{15 - 48}{20 + 36} = -\frac{33}{56}.$$

The final example in this section requires the use of the half-angle identities, which we have not used up till now.

### Example 6.2.16

Solve the equation  $2\sin\theta + \cos\theta = 2$  for angles between 0 and  $2\pi$  (radian). Then using formulae 7(i) and (ii) where

$$t = \tan\frac{\theta}{2}$$

$$2\sin\theta + \cos\theta = 2$$

$$\frac{4t}{1+t^2} + \frac{1-t^2}{1+t^2} = 2 \quad \text{and on multiplication by } 1+t^2$$

$$4t + 1 - t^2 = 2 + 2t^2$$

$$\text{or } -3t^2 + 4t - 1 = 0 \quad (\text{yet another quadratic!})$$

again factorising gives:

$$(-3t + 1)(t - 1) = 0$$

$$\Rightarrow -3t + 1 = 0 \quad \text{giving } t = \frac{1}{3}$$

$$\text{or } \Rightarrow t - 1 = 0 \quad \text{giving } t = 1 \quad \text{where } t = \tan\frac{\theta}{2}$$

$$\text{so } \tan\frac{\theta}{2} = \frac{1}{3} \quad \text{or } 1$$

so we require values of  $\frac{\theta}{2}$  between 0 and  $\pi$

$$\text{so } \tan\frac{\theta}{2} = \frac{1}{3} \quad \text{gives } \frac{\theta}{2} = 18.43^\circ$$

$$\text{or } \tan\frac{\theta}{2} = 1 \quad \text{gives } \frac{\theta}{2} = 45^\circ.$$

So in our range  $0-360^\circ$ ,

$$\theta = 36.87^\circ \quad \text{or } 90^\circ.$$

Note the next values of  $\theta/2$  corresponding to  $\frac{1}{3}$  and 1 are  $198.43^\circ$  and  $225^\circ$  which when doubled are outside the required range.

This concludes our brief study of trigonometry, now try the problems!

**Question 6.2.1**

- (1) Find the length of arc and area of sector of a circle when:  
 (i)  $\theta = 30^\circ$ ,  $r = 18$  cm; (ii)  $\theta = 135^\circ$ ,  $r = 24$  cm.

- (2) Convert the points  $(-3, 4)$  and  $(-2, -3)$  to polar co-ordinates.

- (3) Convert the points  $(4, \pi/4)$  and  $(2, -30^\circ)$  to Cartesian co-ordinates.

- (4) Sketch the graphs of the following sinusoidal functions:

(i)  $y_1 = 2 \sin \left( 2\theta + \frac{\pi}{2} \right)$

(ii)  $y_2 = 10 \cos (\theta - 60)$ .

What is the 'phase angle' between the functions  $y_1$  and  $y_2$ ?

- (5) State the amplitude and phase angle for the following functions:

(i)  $y_1 = 6 \sin \left( 3\theta - \frac{\pi}{6} \right)$

(ii)  $y_2 = 4 \sin \left( 20\theta - \frac{\pi}{3} \right)$ .

- (6) For each of the functions:

(i)  $i = 3 \sin \left( 200\pi t - \frac{\pi}{4} \right)$

(ii)  $v = 0.7 \sin \left( 400\pi t + \frac{\pi}{3} \right)$ .

Calculate the amplitude, frequency, periodic time, phase angle and the time taken for each function to reach its first positive maximum value.

- (7) On the same axes sketch the graphs:

$$y_1 = 3 \sin \left( x + \frac{\pi}{3} \right) \quad \text{and} \quad y_2 = 2 \sin \left( x - \frac{\pi}{3} \right),$$

between 0 and  $2\pi$ . On the same axes graphically sum these functions to produce the graph of

$$y_3 = 3 \sin \left( x + \frac{\pi}{3} \right) + 2 \sin (x - 60).$$

- (8) For an AC  $i = 40 \sin(100\pi t - 0.32)$ , find:

- (i) the amplitude, periodic time, frequency and phase angle with respect to  $40 \sin 100\pi t$ ;  
 (ii)  $i$  when  $t = 0$  and  $i$  when  $t = 4$ ;  
 (iii) find  $t$  when the current is 25 A.

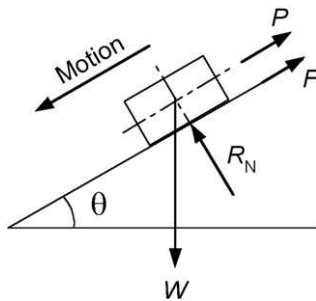


Figure 6.2.6 Force diagram

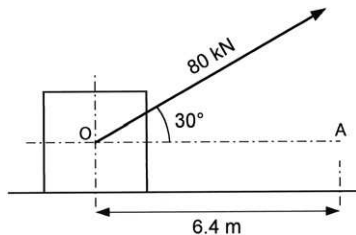


Figure 6.2.7 Force diagram

- (9) A complex waveform is described by the equation

$$v = 100 \sin(100\pi t) + 60 \sin\left(200\pi t - \frac{\pi}{4}\right).$$

Determine graphically the shape of the waveform and estimate the peak-to-peak voltage.

- (10) The instantaneous current ( $i$ ) and the instantaneous voltage ( $v$ ) in a pure resistance AC circuit is given by  $i = I_{\max} \sin(\omega t)$  and  $v = V_{\max} \sin(\omega t)$ . Since power  $P = IV$  show that an equation for instantaneous power is

$$P = \frac{I_{\max}^2 R}{2(1 - \cos(2\omega t))}.$$

- (11) For the situation shown in the force diagram (Figure 6.2.6), find, using the resolution of forces and trigonometric ratios, equations that relate  $P$ ,  $F$ ,  $W$  and  $R_N$  and so show that

$$\frac{F}{R_N} = \tan \theta.$$

- (12) Given that  $\sin(\theta + \phi) = 0.6$  and  $\cos(\theta + \phi) = 0.9$ , find a value for ' $\mu$ ' when  $\mu = \tan \phi$ .
- (13) Calculate the component of force parallel to the horizontal axis (Figure 6.2.7) and the work done given that work = force  $\times$  distance moved in direction of force.
- (14) Verify the following identities:

$$(i) \tan 3\theta = \frac{\sin \theta + \sin 3\theta + \sin 5\theta}{\cos \theta + \cos 3\theta + \cos 5\theta};$$

$$(ii) \tan 2\theta = \frac{1}{1 - \tan \theta} - \frac{1}{1 + \tan \theta}.$$

- (15) Express the following as ratios of single angles:

$$(i) \sin 5\theta \cos \theta + \cos 5\theta \sin \theta;$$

$$(ii) \cos 9t \cos 2t - \sin 9t \sin 2t.$$

## 6.3 CALCULUS

Calculus is a branch of mathematics involved with the quantitative analysis of continually varying functions. It falls conveniently into two parts, *differential calculus*, which is used by engineers mainly to model instantaneous rates of change, while *integral calculus* is used primarily as a summation tool.

You will already have met some of the arithmetic of the calculus, when you *differentiated* (found the differential coefficient) of simple algebraic and trigonometric expressions. You may also be familiar with the *anti-derivative*, when you *integrated* similar expressions.



This section is not intended to offer a rigorous treatment of the derivation of the calculus but rather, to provide a vocabulary of methods needed for its engineering application. The methods are subdivided into differentiation and integration, as discussed above. The engineering applications of the calculus will be dealt with separately, after our study of the methods. Algebraic and trigonometric techniques needed to manipulate the calculus will also be highlighted, as required.

## Differential calculus

### Formulae

#### (1) Standard derivatives

| $y$                         | $\frac{dy}{dx}$                              |
|-----------------------------|----------------------------------------------|
| $ax^n$                      | $nax^{n-1}$                                  |
| $\sin f(x)$                 | $f'(x) \cos f(x)$                            |
| $\cos f(x)$                 | $-f'(x) \sin f(x)$                           |
| $\tan f(x)$                 | $f'(x) \sec^2 f(x)$                          |
| $\operatorname{cosec} f(x)$ | $-f'(x) \operatorname{cosec} f(x) \cot f(x)$ |
| $\sec f(x)$                 | $f'(x) \sec f(x) \tan f(x)$                  |
| $\cot f(x)$                 | $-f'(x) \operatorname{cosec}^2 f(x)$         |
| $\log_e f(x)$               | $\frac{f'(x)}{f(x)}$                         |
| $e^f(x)$                    | $f'(x) e^f(x)$                               |
| $a^x$                       | $a^x \log_e a$                               |
| $\sin^{-1} \frac{x}{a}$     | $\frac{1}{\sqrt{a^2 - x^2}}$                 |
| $\tan^{-1} \frac{x}{a}$     | $\frac{a}{a^2 + x^2}$                        |

#### (2) Rules of differentiation

(i) Function of a function rule  $\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx}$ .

Where  $u$  and  $v$  are functions of  $x$ :

(ii) Product rule  $\frac{d}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}$ .

(iii) Quotient rule  $\frac{d}{dx} \left[ \frac{u}{v} \right] = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$ .

#### (3) Conditions for maxima and minima

$f(x)$  has a **maximum value** at  $x = a$

if  $f'(a) = 0$  and  $f''(x)$  changes sign from +ve to -ve as  $x$  goes through the value  $a$ ,

**or** if  $f'(a) = 0$  and  $f''(a)$  is negative.

$f(x)$  has a **minimum value** at  $x = a$

if  $f'(a) = 0$  and  $f''(x)$  changes sign from -ve to +ve as  $x$  goes through the value  $a$ .

**or** if  $f'(a) = 0$  and  $f''(a)$  is positive.

## Methods

We start with a review of the differing terminology used to describe the differential coefficient and then we review the standard derivatives, with which you should already be familiar. We will then look at the rules needed to differentiate more complex functions. These will include the use of the product, quotient, and function of a function rules. When applied to algebraic, trigonometric, exponential, and hyperbolic functions.

### Standard derivatives

The terminology used for determining the derivative of an expression, differs from textbook to textbook.

For example, we may say:

- (a) find the derivative of ...
- (b) find the differential coefficient for ...
- (c) differentiate ...
- (d) find the rate of change of ...
- (e) find the tangent to the function ...
- (f) find the gradient of the function at a point ...

This differing terminology is often confusing to beginners. It is further complicated by the fact that different symbols are used for the differentiation process, based on the convention chosen.

We can, for example, use Leibniz notation, where the differential coefficient for the function  $y(x)$  is given as,  $\frac{dy}{dx}$ .

This is asking for the function 'y' (with independent variable  $x$ ) to be differentiated once.

In this notation the *second derivative* is expressed as  $\frac{d^2y}{dx^2}$ , in this case we differentiate the function  $y(x)$  twice, while  $\frac{d^3y}{dx^3}$  is the differential coefficient of *degree three*.

Similarly, we may use *functional notation* where for a function  $f(x)$  the first derivative is given as  $f'(x)$ , the second derivative  $f''(x)$  and so on.

One other notation which is often used in mechanics is *dot notation*. For example  $\dot{v}$ ,  $\ddot{s}$  etc. which means the variable is differentiated once ( $\dot{v}$ ) or twice ( $\ddot{s}$ ) and so on.

The terminology in this chapter may vary according to the application. The standard derivative formulae are given in functional notation,  $f(x)$ , etc. but the rules for differentiation are given in Leibniz notation.

Remember also that the notation differs according to the variable! So  $\frac{dy}{dx}$ ,  $\frac{ds}{dt}$ ,  $\frac{du}{dx}$ , each require the first derivative of the functions  $y$ ,  $s$ , and  $u$ , respectively.

It is useful to aid understanding, to state in words the meaning of expressions like  $\frac{ds}{dt}$ . This is saying 'differentiate the function  $s$  with respect to the variable  $t$ .'

Let's look at a few examples of direct application of the standard derivatives, to remind you of the arithmetic process required.

Differentiate the following functions:

$$(i) \ y = 6x^3 \quad \text{then} \quad \frac{dy}{dx} = (6)(3)x^{3-1} = 18x^2$$

$$(ii) \ y = e^{3x} \quad \frac{dy}{dx} = 3e^{3x}$$

$$(iii) \ y = \sin 4x \quad \frac{dy}{dx} = 4 \cos 4x$$

$$(iv) \ y = 3 \cos 3x \quad \frac{dy}{dx} = (3)(-3 \sin 3x) = -9 \sin 3x.$$

Our first example will act as revision for differentiating using standard derivatives.

### Example 6.3.1

Differentiate the following functions with respect to the variables given.

$$(i) \ \frac{x^3}{3} - 6x^{-2} + \frac{12x^4}{32}$$

$$(ii) \ (\sqrt{x})^3 - (x^{-3})^2$$

$$(iii) \ (2x + 3)^2$$

$$(iv) \ \sin 4t - 6 \cos 2t + e^{-3t}.$$

The *golden rule* is to *simplify* whenever possible *before* differentiating.

$$\text{So for (i) if } y = \frac{x^3}{3} - 6x^{-2} + \frac{3x^4}{8}$$

$$\text{then } \frac{dy}{dx} = \frac{3x^2}{3} + 12x^{-3} + \frac{12x^3}{8}$$

$$\frac{dy}{dx} = x^2 + 12x^{-3} + \frac{3x^3}{2}.$$

(ii) Let the function

$$f(x) = (\sqrt{x})^3 - (x^{-3})^2$$

then simplifying  $f(x) = x^{3/2} - x^{-6}$  (laws of indices!)

$$\text{so } f'(x) = \frac{3}{2}x^{1/2} + 6x^{-7}$$

(iii)  $(2x + 3)^2$ . We need to expand the bracket in order to apply a standard derivative:

$$\text{So } (2x + 3)(2x + 3) = 4x^2 + 12x + 9$$

$$\text{and if } y = 4x^2 + 12x + 9$$

$$\text{then } \frac{dy}{dx} = 8x + 12.$$

(iv) Here, we are required to use the standard derivatives for trigonometric and exponential functions. They are all differentiated with respect to the same variable  $t$ , then:

$$\text{if } f(t) = \sin 4t - 6 \cos 2t + e^{-3t}$$

$$f'(t) = 4 \cos 4t + 12 \sin 2t - 3e^{-3t}.$$

Note also that the standard derivative for the Napierian logarithm ( $\ln x$ ) or ( $\log_e x$ ) is  $\log_e f(x) = \frac{f'(x)}{f(x)}$  so for example if

$y = \log_e f(t)$  and  $f(t) = x^2 + 1$ , then:

$$y = \log_e (x^2 + 1) \quad \text{and} \quad \frac{dy}{dx} = \frac{2x}{x^2 + 1}.$$

### Function of a function rule

We will now look at examples of the use of the function of a function rule, which is really differentiation by substitution.

#### Example 6.3.2

Find  $\frac{dy}{dx}$  if  $y = (x^2 - x)^9$ .

Now the function of a function rule

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$$

is a rule that requires us to make a substitution, in the above function where  $y = (x^2 - x)^9$  if we let the bracketed expression =  $u$ , then we would get:

$$y = u^9 \quad (\text{where } u = x^2 - x).$$

We cannot, however, directly differentiate this expression because it involves a different variable to that required, that is, we want to differentiate with respect to  $x$ , but we have  $y$  in terms of  $u$ .

So to complete the function of a function formula we need

$$\frac{dy}{du} \quad \text{and} \quad \frac{du}{dx}$$

we have  $y = u^9$  and  $u = x^2 - x$

$$\text{So } \frac{dy}{du} = 9u^8 \quad \text{and} \quad \frac{du}{dx} = 2x - 1$$

$$\text{Then } \frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx} = (9u^8)(2x - 1).$$

The differentiation is complete, all that is required is to put  $u$  in terms of  $x$  by using the original substitution.

$$\text{So } \frac{dy}{dx} = 9(x^2 - x)^8(2x - 1).$$

To ensure that you are able to master the substitution method for differentiating functions. You should study the example given next, very carefully. Make sure that you are able to follow the standard procedure used for all the worked solutions.

### Example 6.3.3

Differentiate the following functions with respect to the variables given using the function of a function (substitution) rule.

(i)  $y = \sqrt{1 - 5x^3}$

(ii)  $\frac{d}{d\theta} (\sin 7\theta)$

(iii)  $\frac{d}{dt} \left[ \cos \left( 2t - \frac{3\pi}{2} \right) \right]$

(iv)  $\frac{d}{dx} [\log_e (x^2 + 5)]$

(i) Let  $y = \sqrt{1 - 5x^3}$  that is  $y = (1 - 5x^3)^{1/2}$ . Then  $y = u^{1/2}$  where  $u = 1 - 5x^3$ .

$$\text{Therefore } \frac{dy}{du} = \frac{1}{2}u^{-1/2} \quad \text{and} \quad \frac{du}{dx} = -15x^2.$$

$$\text{But } \frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx},$$

$$\text{Therefore } \frac{dy}{dx} = \frac{1}{2}u^{-1/2} \times (-15x^2)$$

$$\begin{aligned} \text{So } \frac{dy}{dx} &= \frac{1}{2}(1 - 5x^3)^{-1/2} (-15x^2) \\ &= -\frac{15}{2}x^2(1 - 5x^3)^{-1/2} \end{aligned}$$

$$\frac{dy}{dx} = -\frac{15x^2}{2\sqrt{1 - 5x^3}}.$$

$$\text{Hence } \frac{d}{dx} (\sqrt{1 - 5x^3}) = -\frac{15x^2}{2\sqrt{1 - 5x^3}}.$$

(ii)  $\frac{d}{d\theta} (\sin 7\theta)$ . Let  $y = \sin 7\theta$ .

$$\text{Then } y = \sin u \quad \text{where } u = 7\theta$$

$$\frac{dy}{du} = \cos u \quad \text{and} \quad \frac{du}{d\theta} = 7.$$

$$\text{But } \frac{dy}{d\theta} = \frac{dy}{du} \times \frac{du}{d\theta}.$$

$$\text{Therefore } \frac{dy}{d\theta} = \cos u \times 7 = 7 \cos u = 7 \cos 7\theta.$$

Therefore  $\frac{d}{d\theta} (\sin 7\theta) = 7 \cos 7\theta$ .

$$(iii) \frac{d}{dt} \left( \cos \left( 2t - \frac{3\pi}{2} \right) \right). \quad \text{Let } y = \cos \left( 2t - \frac{3\pi}{2} \right).$$

Then  $y = \cos u$  where  $u = 2t - \frac{3\pi}{2}$ .

Therefore  $\frac{dy}{du} = -\sin u$  and  $\frac{du}{dt} = 2$ .

$$\text{But } \frac{dy}{dt} = \frac{dy}{du} \times \frac{du}{dt},$$

Therefore  $\frac{dy}{dt} = (-\sin u) \times 2 = -2 \sin u = -2 \sin \left( 2t - \frac{3\pi}{2} \right)$

$$\text{hence } \frac{d}{dt} \left( \cos \left( 2t - \frac{3\pi}{2} \right) \right) = -2 \sin \left( 2t - \frac{3\pi}{2} \right).$$

$$(iv) \frac{d}{dx} (\log_e (x^2 + 5)). \quad \text{Let } y = \log_e (x^2 + 5).$$

Then  $y = \log_e u$  where  $u = x^2 + 5$ .

Therefore  $\frac{dy}{du} = \frac{1}{u}$  and  $\frac{du}{dx} = 2x$

$$\text{But } \frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}.$$

$$\text{Therefore } \frac{dy}{dx} = \frac{1}{u} \times 2x = \frac{1}{x^2 + 5} \times 2x.$$

$$\text{Hence } \frac{d}{dx} (\log_e (x^2 + 5)) = \frac{2x}{x^2 + 5} = \frac{f'(x)}{f(x)}!$$

With practice the analysis of a function becomes a mental process and with it the differentiation of a function of a function. For example:

$$\text{If } y = (3x^3 - 3x)^4$$

$$\frac{dy}{dx} = 4(3x^3 - 3x)^3 (9x^2 - 3).$$

$$\text{If } y = \cos^3 5x = (\cos 5x)^3$$

$$\frac{dy}{dx} = (3 \cos^2 5x)(-\sin 5x)5$$

$$= -15 \cos^2 5x \sin 5x.$$

$$\text{If } y = \log_e (3x^2 - 6x)^3$$

$$\frac{dy}{dx} = \frac{1}{(3x^2 - 6x)^3} 3(3x^2 - 6x)^2 (6x - 6)$$

$$= \frac{3(6x - 6)}{(3x^2 - 6x)}.$$

If this mental process is unclear, you should stick to the full process already covered for differentiation using the function of a function (substitution) rule.

### Product and quotient rules

These rules (formula 2(ii) and 2(iii)) enable us to differentiate more complex functions which are made up, as their name suggests from products and quotients, for example the function:

$y = x^3 \sin 2x$  is the product of  $x^3$  and  $\sin 2x$

while  $y = \frac{e^{2x}}{x + 3}$  is  $e^{2x}$  divided by  $x + 3$ .

The rules for products and quotients require the use of similar algebraic manipulation to that required when using the function of a function rule. So if  $y = u \times v$  where  $u$  and  $v$  are functions of any other variable (say  $x$ ), then we must use the product rule (2ii) to find the differential of the function. Similarly, if  $y = u/v$  where  $u$  and  $v$  are functions of any variable (say  $x$ ) then we must use the quotient rule (2iii) to find the differential of the function.

The following examples illustrate the similarities between the use of the product and quotient rules. The quotient rule is considered slightly more difficult to use than the product rule, because it is dependent on order, as you will see. Quotients can often be represented as products by manipulating the function using the laws of indices. The quotient may then be treated as a product and differentiated using the product rule.

#### Example 6.3.4

(a) Find  $\frac{d}{dx} (x^3 \sin 2x)$ .

(b) Differentiate  $(x^2 + 1) \log_e x$ .

Both of the above functions are products so rule 2ii is appropriate.

(a) Let  $y = x^3 \sin 2x$

then  $y = u \times v$  where  $u = x^3$  and  $v = \sin 2x$ .

Therefore  $\frac{du}{dx} = 3x^2$  and  $\frac{dv}{dx} = 2 \cos 2x$ .

But from the rule:

$$\frac{dy}{dx} = v \frac{du}{dx} + u \frac{dv}{dx}.$$

after substituting our values from above, we get:

$$\frac{dy}{dx} = (\sin 2x) (3x^2) + (x^3) (2 \cos 2x)$$

then  $\frac{d}{dx} (x^3 \sin 2x) = x^2 (3 \sin 2x + 2x \cos 2x)$ .

(b) Again letting  $y = (x^2 + 1) \log_e x$ , then we are required to find  $\frac{dy}{dx}$ .

$$y = u \times v \quad \text{where} \quad u = x^2 + 1 \quad \text{and} \quad v = \log_e x.$$

$$\text{Therefore} \quad \frac{du}{dx} = 2x \quad \text{and} \quad \frac{dv}{dx} = \frac{1}{x}$$

$$\text{but} \quad \frac{dy}{dx} = v \frac{du}{dx} + u \frac{dv}{dx} \quad \text{and on substitution:}$$

$$\begin{aligned} \frac{dy}{dx} &= (\log_e x) (2x) + (x^2 + 1) \left( \frac{1}{x} \right) \\ &= 2x(\log_e x) + x + \frac{1}{x}. \end{aligned}$$

### Example 6.3.5

(a) Find  $\frac{dy}{dx}$  if  $y = \frac{e^{2x}}{x+3}$ .

(b) Find  $\frac{d}{d\theta} (\tan \theta)$ .

These are both quotients, although for  $b$  this is not obvious!

(a)  $y = \frac{e^{2x}}{x+3}$  let  $y = \frac{u}{v}$  where  $u = e^{2x}$  and  $v = x+3$ .

$$\text{Therefore} \quad \frac{du}{dx} = 2e^{2x} \quad \text{and} \quad \frac{dv}{dx} = 1.$$

$$\text{But} \quad \frac{dy}{dx} = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$$

**When selecting  $u$  and  $v$  this rule requires  $u$  to be the numerator and  $v$  to be the denominator of the function.**

Then substitution gives:

$$\begin{aligned} \frac{dy}{dx} &= \frac{(x+3)(2e^{2x}) - (e^{2x} \times 1)}{(x+3)^2} \\ &= \frac{e^{2x}(2x+6-1)}{(x+3)^2} \end{aligned}$$

$$\frac{dy}{dx} = \frac{(2x+5)(e^{2x})}{(x+3)^2}.$$

(b) Let  $y = \tan \theta$ , we require  $\frac{dy}{d\theta}$ . Now to turn  $\tan \theta$  into a quotient we use the trigonometric identity:

$$\tan \theta = \frac{\sin \theta}{\cos \theta} \quad \text{so} \quad y = \frac{\sin \theta}{\cos \theta}$$

$$\text{where} \quad u = \sin \theta \quad \frac{du}{d\theta} = \cos \theta$$



$$v = \cos \theta \quad \frac{dv}{d\theta} = -\sin \theta.$$

$$\text{But } \frac{dy}{d\theta} = \frac{v \frac{du}{d\theta} - u \frac{dv}{d\theta}}{v^2}$$

$$= \frac{(\cos \theta)(\cos \theta) - (\sin \theta)(-\sin \theta)}{\cos^2 \theta}$$

$$= \frac{\cos^2 \theta + \sin^2 \theta}{\cos^2 \theta} \quad (\text{our old friend!})$$

$$\frac{dy}{d\theta} = \frac{1}{\cos^2 \theta} = \sec^2 \theta.$$

$$\text{So } \frac{d(\tan \theta)}{d\theta} = \sec^2 \theta.$$

We will now consider two general methods for the differential calculus which enable us to find *rates of change* and *turning points* (TPs). These methods may then be adopted for specific engineering use.

### The derivative and rate of change

You saw in the beginning of this section that we can define the differential coefficient  $dy/dx$  as the rate at which  $y$  changes as  $x$  changes. Suppose we have a function where both  $y$  and  $x$  depend on time ( $t$ ), then we may use the *chain* (function of a function) *rule* to write this as:

$$\frac{dy}{dx} = \frac{dy}{dt} \cdot \frac{dt}{dx} \quad (\text{dot notation here means multiplication})$$

$$\text{then } \frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt}$$

$$\text{or } \frac{dx}{dt} = \frac{dy}{dt} \cdot \frac{dx}{dy}.$$

So from the above the rate of change of  $y$  with respect to  $t$ , or the rate of change of  $y$  with respect to  $x$  may be found, as required.

This method of calculating the rate of change of some variable or function with respect to time has wide practical use, since many situations that occur in real life depend upon time as their independent variable. The exponential growth of bacteria, the decay of charge in a capacitor, or the change in velocity of a car, are all dependent on time.

#### Example 6.3.6

Calculate the rate at which the area of a circle increases with respect to time ( $t$ ), when the radius of the circle increases at 0.2 cm/s. Now from your previous work the area of a circle is:

$$A = \frac{\pi d^2}{4} = \pi r^2.$$

Therefore, the rate of change of area,  $A$ , with respect to the radius,  $r$  is given by:

$$\frac{dA}{dr} \quad \text{so when} \quad A = \pi r^2, \quad \frac{dA}{dr} = 2\pi r$$

and remembering that  $\frac{dr}{dt} = 0.2 \text{ cm/s}$ , then using function of a function rule

$$\frac{dA}{dr} = \frac{dA}{dt} \cdot \frac{dt}{dr}$$

so  $\frac{dA}{dt}$  (the change in area with respect to time) is given by:

$$\frac{dA}{dt} = \frac{dA}{dr} \cdot \frac{dr}{dt}$$

$$\text{so } \frac{dA}{dt} = (2\pi r)(0.2) = 0.4\pi r \text{ cm}^2/\text{s}.$$

Then for example when  $r = 2 \text{ cm}$ , the rate of change of area

$$\frac{dA}{dt} = 0.8\pi \text{ cm}^2/\text{s}.$$

### Example 6.3.7

Suppose an empty spherical vessel is filled with water. As the water level rises the radius of the water in the vessel and the volume of water will change. Now if the radius of water in the sphere increases at  $0.5 \text{ cm/s}$ , find the rate of change of volume, when the radius is  $5 \text{ cm}$ .

We know that:

$$V = \frac{4}{3}\pi r^3 \quad \text{and} \quad \frac{dr}{dt} = 0.5 \text{ cm/s},$$

$$\text{so } \frac{dV}{dr} = 4\pi r^2$$

and using function of a function rule:

$$\frac{dV}{dt} = \frac{dV}{dr} \cdot \frac{dr}{dt}$$

$$\text{then } \frac{dV}{dt} = (4\pi r^2)(0.5) = 2\pi r^2$$

$$\text{and when } r = 5, \quad \frac{dV}{dt} = 50\pi = 157 \text{ cm}^3/\text{s}.$$

### TPs

Consider the curve of the function  $y = f(x)$ , shown in Figure 6.3.1. Each cross represents a TP for the function where, the gradient of the tangent at these points is zero. So a TP is defined as *the point on any continuous function where the gradient of the tangent of the point is zero*. So to find TPs

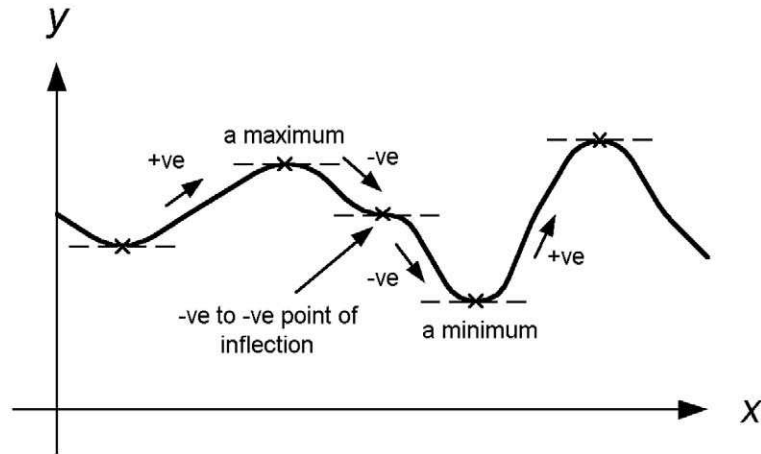


**Figure 6.3.1** Curve showing turning points

values (independent variable  $x$  in this case), we solve the gradient equation:

$$\frac{dy}{dx} = 0.$$

To find the corresponding  $y$ -ordinate of the TP, we substitute the  $x$ -values back into the original function, that is  $y = f(x)$ , these values are called *stationary values* (SVs). The following example illustrates the process.

### Example 6.3.8

Find the TPs for the function  $y = 2x^3 - 3x^2 - 12x + 4$ . The requirement for a TP is that the rate of change of the function equals zero, that is:

$$\frac{dy}{dx} = 0 \quad \text{or} \quad \frac{dy}{dx} = 6x^2 - 6x - 12 = 0. \quad (a)$$

Solving this equation will produce the  $x$  value for which there is a TP.

Dividing equation (a) by 6 gives:  $\frac{dy}{dx} = x^2 - x - 2 = 0$  which has factors  $(x + 1)(x - 2) = 0$ .

**So TPs exist at  $x = 2$  or  $-1$ .**

Now substituting these values back into the original function  $y$  we get:

$$y = 2(2)^3 - 3(2)^2 - 12(2) + 4 = -16$$

$$\text{or } y = 2(-1)^3 - 3(-1)^2 - 12(-1) + 4 = 11.$$

So TPs and SVs are at the points **(2, -16) and (-1, 11)**.

To decide on the nature of the TPs, that is, whether they are a *maximum*, or *minimum* or *point of inflection*, two methods are often adopted (see formula 3).

### Method 1

Determine the rate of change of the gradient function, in other word find the value for the second derivative of the function  $\frac{d^2y}{dx^2}$  at the TP. If this

value is *positive* the TP is a *minimum*, if this value is *negative* then the TP is a *maximum*.

## Method 2

Consider the *gradient of the curve* close to the TPs, that is, near to either side. Then for a *minimum* the gradient goes from *negative to positive* (Figure 6.3.1) and for a *maximum* the gradient goes from *positive to negative*. Both methods are illustrated in the following two examples.

### Example 6.3.9

Using Method 1 find the maximum and minimum values of  $y$  given that:

$$y = x^3 + 3x^2 - 9x + 6$$

$$\text{then } \frac{dy}{dx} = 3x^2 + 6x - 9 \text{ and } \frac{d^2y}{dx^2} = 6x + 6.$$

At a TP  $\frac{dy}{dx} = 0$ , therefore  $3x^2 + 6x - 9 = 0$  or  $x^2 + 2x - 3 = 0$  and  $(x - 1)(x + 3) = 0$ . Therefore, TPs at  $x = 1$  and  $x = -3$ . Now we must test for maximum or minimum. So at TP where  $x = 1$

$$\frac{d^2y}{dx^2} = 6x + 6 = 6(1) + 6 = +12.$$

This is *positive* and so the TP at  $x = 1$  is a *minimum* and the SV at this point is found by substituting  $x = 1$  into the *original* equation for  $y$ ,

$$\text{That is } y_{\min} = (1)^3 + 3(1)^2 - 9(1) + 6 = 1.$$

**So a minimum at point (1, 1).**

Similarly at the point where  $x = -3$ :

$$\frac{d^2y}{dx^2} = 6(-3) + 6 = -12.$$

This is *negative* and so at  $x = -3$  there is a *maximum* TP.

$$y_{\max} = (-3)^3 + 3(-3)^2 - 9(-3) + 6 = +13.$$

**So a maximum at point (-3, 13).**

### Example 6.3.10

For the function  $y = x^3 - 3x$  find the TPs and determine their nature.

For TPs we require:

$$\frac{dy}{dx} = 3x^2 - 3 = 0$$

and so  $3x^2 - 3 = 0$  or  $x^2 - 1 = 0$

so  $x = \pm 1$ .

SV values corresponding to TP at  $+1$  and  $-1$  are found by substitution into original equation:

$$y = x^3 - 3x \quad \text{then} \quad y = (+1)^3 - 3(+1) = -2$$

$$\text{and also} \quad y = (-1)^3 - 3(-1) = +2.$$

**So TPs at  $(1, -2)$  and  $(-1, +2)$ .**

Now for point  $(1, -2)$  we consider values above and below the value of  $x = 1$  so choose  $x = 0$  and  $x = 2$ , then using the *gradient* equation:

$$\frac{dx}{dy} = 3x^2 - 3,$$

when  $x = 0$ ,  $\frac{dy}{dx}$  is negative;

when  $x = 2$ ,  $\frac{dy}{dx}$  is positive.

Gradient function goes from  $-ve$  to  $+ve$ . **Then point  $(1, -2)$  is a minimum.**

Similarly at point  $(-1, 2)$  at values of  $x = -2$  and  $x = 0$ ,

then when  $x = -2$ ,  $\frac{dy}{dx} = 3x^2 - 3$  is  $+ve$ .

and when  $x = 0$ ,  $\frac{dy}{dx} = 3x^2 - 3$  is  $-ve$ .

Gradient function goes from  $+ve$  to  $-ve$ . **Then point  $(-1, -2)$  is a maximum.**

This concludes our short study on methods for using the differential calculus, its application to engineering will be found later after our study of the integral calculus.

## Integral calculus

### Formulae

(1) *Standard integrals (constant of integration omitted)*

| $y$               | $\int y \, dx$        |
|-------------------|-----------------------|
| $x^a (a \neq -1)$ | $\frac{x^{a+1}}{a+1}$ |
| $\frac{1}{x}$     | $\ln x $              |
| $\ln x$           | $x \ln x - x$         |

(continued)

| $y$                      | $\int y \, dx$                                                                                 |
|--------------------------|------------------------------------------------------------------------------------------------|
| $a^x (a > 0)$            | $\frac{a^x}{\ln a}$                                                                            |
| $(ax + b)^n; n \neq -1$  | $\frac{1}{a} \frac{(ax + b)^{n+1}}{n + 1}$                                                     |
| $\frac{1}{ax + b}$       | $\frac{1}{a} \ln(ax + b)$                                                                      |
| $\tan x$                 | $\sec^2 x$                                                                                     |
| $\sin(ax + b)$           | $-\frac{1}{a} \cos(ax + b)$                                                                    |
| $\cos(ax + b)$           | $\frac{1}{a} \sin(ax + b)$                                                                     |
| $\tan(ax + b)$           | $\frac{1}{a} \ln[\sec(ax + b)]$                                                                |
| $e^{ax+b}$               | $\frac{1}{a} e^{ax+b}$                                                                         |
| $\sec x$                 | $\begin{cases} \ln  \tan(\frac{1}{2}x + \frac{1}{4}\pi)  \\ \ln  \sec x + \tan x  \end{cases}$ |
| $\operatorname{cosec} x$ | $\begin{cases} \ln  \tan \frac{1}{2}x  \\ \ln  \operatorname{cosec} x - \cot x  \end{cases}$   |
| $\cot x$                 | $\ln  \sin x $                                                                                 |
| $\sinh x$                | $\cosh x$                                                                                      |
| $\cosh x$                | $\sinh x$                                                                                      |
| $\tanh x$                | $\ln \cosh x$                                                                                  |
| $\sin^2 x$               | $\frac{1}{2}x - \frac{1}{4}\sin 2x$                                                            |
| $\cos^2 x$               | $\frac{1}{2}x + \frac{1}{4}\sin 2x$                                                            |
| $\tan^2 x$               | $\tan x - x$                                                                                   |

(2) *Integration by substitution (or change of variable)*

$$\int f(x) \, dx = \int f(g(t))g'(t) \, dt \quad \text{when } x = g(t).$$

(3) *Integration by parts*

$$\int u \frac{dv}{dx} \, dx = uv - \int v \frac{du}{dx} \, dx \quad \text{or} \quad \int uv' \, dx = uv - \int vu' \, dx.$$

(4) *Numerical integration*

*Simpson's rule:* If a plane area is divided into an even number of strips of equal width, then:

$$\text{Area} = \frac{\text{Common width}}{3} \left[ \begin{array}{l} \text{sum of first and last ordinates} \\ + 4 \times (\text{sum of even ordinates}) \\ + 2 \times (\text{sum of remaining ordinates}) \end{array} \right]$$

(5) *Trapezoidal rule*

If a plane area is divided into strips of equal width, then the area = common width  $\times$  (half the sum of the first and the last ordinates + the sum of the other ordinates).

**Methods**

In this section we will briefly review the standard integrals and their use in the determination of areas by direct integration. Next we will look at the techniques for integrating functions by substitution and, by parts. Finally we will use the Trapezoidal rule and Simpson's rule to illustrate the very powerful technique of numerical integration.

**Standard integrals**

Finding the prime function using standard integrals, having been given its derivative, is often referred to as *anti-differentiation*. A glance at the rule for finding the prime function (F) for a simple polynomial expression shows that we are finding the antiderivative. The following examples show the straightforward use of some standard integrals.

**Example 6.3.11**

Find the *integrals* of the following prime functions, using standard integrals.

- (i)  $3x + 2$ , (ii)  $(x + 2)^6$ , (iii)  $3e^{2x}$ ,  
 (iv)  $3 \cos 2x$ , (v)  $3 \cosh 2x$ , (vi)  $(x + 2)^{-1} x > -2$ ,  
 (vii)  $3e^{2x+1}$ , (viii)  $3 \cos(2x + 1)$ , (ix)  $3 \cosh(2x + 1)$ ,  
 (x)  $(3x + 2)^{-1} x > -\frac{2}{3}$ .

Using the table the *integrals* for the above functions are:

(i)  $\frac{3}{2}x^2 + 2x + c$

The constant 'c', results from the process of anti-differentiation where it may or may not be present it is known as the *constant of integration*, with which you should already be familiar.

- (ii)  $\frac{(x + 2)^7}{7}$ , (iii)  $\frac{3}{2}e^{2x}$ , (iv)  $\frac{3}{2}\sin 2x$ ,  
 (v)  $\frac{3}{2}\sinh 2x$ , (vi)  $\log(x + 2)$ , (vii)  $\frac{3}{2}e^{2x+1}$ ,  
 (viii)  $\frac{3}{2}\sin(2x + 1)$ , (ix)  $\frac{3}{2}\sinh(2x + 1)$ , (x)  $\log(x + 2)$ .

Make sure you can integrate elementary functions using standard integrals!

**Example 6.3.12**

- (i)  $\int (x^2 + 2x - 3)dx$ ,  
 (ii)  $\int (\sin 7\theta + 2 \cos 5\theta)d\theta$

$$(iii) \int \left( e^{6t} - \frac{1}{e^{3t}} \right) dt.$$

$$(i) \int (x^2 + 2x - 3) dx = \frac{x^3}{3} + \frac{2x^2}{2} - 3x + c$$

$$= \frac{x^3}{3} + x^2 - 3x + c$$

(applying the rule to each term sequentially).

$$(ii) \int (\sin 7\theta + 2 \cos 5\theta) d\theta = -\frac{1}{7} \cos 7\theta + \frac{2}{5} \sin 5\theta + c.$$

$$(iii) \int \left( e^{6t} - \frac{1}{e^{3t}} \right) dt = \int (e^{6t} - e^{-3t}) dt \quad (\text{indices!})$$

$$= \frac{1}{6} e^{6t} - \frac{1}{(-3)} e^{-3t} + c$$

$$= \frac{1}{6} e^{6t} + \frac{1}{3} e^{-3t} + c.$$

### Example 6.3.13

Evaluate  $\int_2^3 (1 + \cos 2\phi) d\phi$ .

The above integral is known as a *definite integral* because it has definite *limits*. It is telling us to find the integral of the function and then find a numerical value, when the limiting values are substituted for the variable ( $\phi$  in this case). So we first integrate in the normal manner.

$$\int_2^3 (1 + \cos 2\phi) d\phi = [\phi + \frac{1}{2} \sin 2\phi]_2^3$$

We put the integral in square brackets, with the upper and lower limit of  $\phi$ , as indicated. Now the final numerical value for the function will be the difference between the upper and lower limits, when the values of  $\theta$  are substituted. In other words:

$$= [3 + \frac{1}{2} \sin(2)(3)] - [2 + \frac{1}{2} \sin(2)(2)]$$

$$= 3 + \frac{1}{2} \sin 6 - 2 - \frac{1}{2} \sin 4.$$

Now remembering that when limits are substituted into trigonometrical functions, we deal in **radian** values. So evaluating, using a calculator, gives:

$$\int_2^3 (1 + \cos 2\phi) d\phi = 1.24.$$

The above examples were all straightforward, all we had to do was use standard integrals either referring to the formulae given in the table or rely on our memory. We will shortly be looking at methods to help us integrate some of the more complex expressions. Before we do, let us look at an example where we need to simplify the integral in some manner, prior to using standard integrals.



**Example 6.3.14**

Find: (i)  $\int \frac{e^x + e^{6x} + e^{3x}}{e^{4x}} dx$ ;

(ii)  $\int \tan^2 x dx$ .

(i) This is simple when you realise that all we need to do is divide each term by  $e^{4x}$  and then integrate each term sequentially. It does, once again, require us to use indices.

$$\begin{aligned} \int \frac{e^x + e^{6x} + e^{3x}}{e^{4x}} dx &= \int e^{-3x} + e^{2x} + e^{-x} dx \\ &= -\frac{1}{3}e^{-3x} + \frac{1}{2}e^{2x} - e^{-x} + c. \end{aligned}$$

(ii)  $\int \tan^2 x dx$ .

This is fairly straightforward if we use a trigonometric identity, the problem is which one? A general rule is that whenever you see  $\tan^2 x$ , think of this as  $\sec^2 x - 1$  (look at your standard integrals to see why).

The above may now be written:

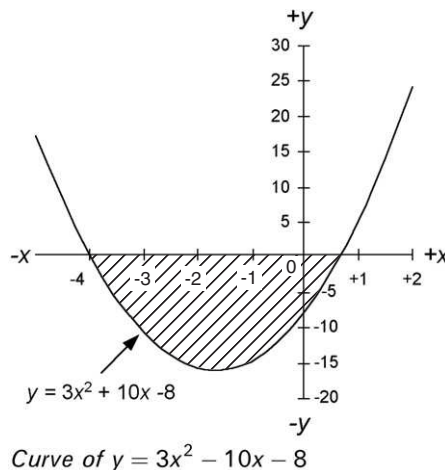
$$\int \tan^2 x dx = \int (\sec^2 x - 1) dx = \tan x - x + c.$$

This is why it is so useful to know your trigonometric identities!

The following example illustrates the technique for finding the area of a given function, when we know the rule of the function. In practice, this seldom occurs, and we would need to use one of the numerical methods, you will meet later.

**Example 6.3.15**

Determine the shaded area shown below (see figure), between the function and the x-axis.



Now in order to find this area by integration, we first need to find the 'limits of integration'. The function crosses the  $x$ -axis when  $y = 0$ , so the required limits are at  $y = 0$ .

$$\text{Then } 3x^2 + 10x - 8 = 0$$

and factorising we get:

$$(3x - 2)(x + 4) = 0, \text{ so } x = \frac{2}{3} \text{ and } x = -4.$$

So we have our limits. To obtain the required area, we need to integrate between the limits  $-4$  to  $0$  and  $0$  to  $\frac{2}{3}$ . In other words the required area is given by:

$$\begin{aligned} & \int_0^{-4} 3x^2 + 10x - 8 \, dx + \int_0^{2/3} 3x^2 + 10x - 8 \, dx \\ &= \left[ x^3 + 5x^2 - 8x + c \right]_0^{-4} + \left[ x^3 + 5x^2 - 8x + c \right]_0^{2/3} \\ &= \left[ (-4)^3 + 5(-4)^2 - 8(-4) + c \right]_0^{-4} \\ &\quad + \left[ \left(\frac{2}{3}\right)^3 + 5\left(\frac{2}{3}\right)^2 - 8\left(\frac{2}{3}\right) + c \right]_0^{2/3} \\ &= [(48 + c) - (0 + c)] + [(0.259 + c) - (0 + c)]. \end{aligned}$$

Notice the constants of integration are eliminated.

$$\text{Then } = 48 + 0.259.$$

**So required area = 48.259 sq. units.**

Example 6.3.16 shows how a relatively complex expression may be simplified using partial fractions (PFs) and then the integral found with relative ease. You are now able to reap the benefits of your previous work on PFs!

### Example 6.3.16

$$\text{Find } \int \frac{dx}{(x-2)^2(x-3)}.$$

At first sight, this looks a rather daunting problem, but reference to the table of standard integrals, gives us an idea of what is

required. Functions of the type  $\int \frac{1}{x+a} dx$  have solutions that involve the  $\ln$  function.

If we can find terms similar to this for our integral, it can be solved. To do this we use PFs.

The above integral will have PFs of the form:

$$\frac{A}{x-2} + \frac{B}{(x-2)^2} + \frac{C}{(x-3)}$$

In fact reference to Example 6.1.15 gives us the PFs.

Then

$$\int \frac{dx}{(x-2)^2(x-3)} = \int -\frac{1}{(x-2)} - \frac{1}{(x-2)^2} + \frac{1}{(x-3)} dx$$

So

$$\int \frac{dx}{(x-2)^2(x-3)} = -\ln(x-2) + \frac{1}{(x-2)} + \ln(x-3) + C.$$

### Integration by substitution

You have already met a technique involving substitution, when you used the function of a function rule for differentiation. Unfortunately, unlike derivatives, when considering substitutions to simplify integrals we do not always know what to substitute. With practice and a little trial and error, the appropriate substitution can normally be found. The following examples illustrate some of the more important substitutions, that provide useful results.

#### Example 6.3.17

Find  $\int \frac{\arctan x}{1+x^2} dx$ .

The clue is in the  $\arctan$  (the angle whose tangent is). We could let  $u = 1 + x^2$ , but then the integral still involves  $\arctan x$ . A good course of action is to try and eliminate that which we know least about. So let's try  $u = \arctan x$ , then  $x = \tan u$ . If we substitute now, we have:

$$\int \frac{u}{1 + \tan^2 u} dx$$

(this is nonsense because we are told to integrate with respect to  $x$ , but the integral has another variable) so we

look again at  $x = \tan u$  and note that  $\frac{dx}{du} = \sec^2 u$ . So

$dx = \sec^2 u du$  (now we are getting somewhere). We have a substitute for  $dx$  in terms of  $u$ :

$$\begin{aligned} \text{Then } \int \frac{\arctan x}{1+x^2} dx &= \int \frac{u}{1+\tan^2 u} \sec^2 u du \\ &= \int u du \left( \begin{array}{l} \text{because} \\ \sec^2 u = 1 + \tan^2 u \end{array} \right) \\ &= \frac{u^2}{2} \text{ and using our original substitution} \end{aligned}$$

$$\text{then } \int \frac{\arctan x}{1+x^2} dx = \frac{(\arctan x)^2}{2}.$$

Look back carefully over this example and make sure you follow the logic of the process.

**Example 6.3.18**

Find  $\int \frac{dx}{\sqrt{1+x^2}}$ .

Integrals of this type are best dealt with using a trigonometric or hyperbolic substitution, since the expression under the square root is of the form  $a^2 + x^2$ , or  $a^2 - x^2$ , or  $x^2 - a^2$ . In our case  $a^2 + x^2$ , we use  $x = \sinh u$ .

Substitutions for the other variations are given at the end of this example.

For our integral we have:

$$x = \sinh u, \quad \frac{dx}{du} = \cosh u \quad \text{or} \quad dx = \cosh u \, du.$$

So our integral becomes:

$$\int \frac{\cosh u}{\sqrt{1 + \sinh^2 u}} \, du$$

and remembering yet again our old friend in hyperbolic form  $\cosh^2 u - \sinh^2 u = 1$  then  $\cosh^2 u = 1 + \sinh^2 u$

and so  $\cosh u = \sqrt{1 + \sinh^2 u}$

$$\text{so} \quad \int \frac{\cosh u}{\cosh u} \, du = \int 1 \, du = u = \sinh^{-1} x.$$

For other integrals of this type, try the following substitutions.

For integrals involving:

(i)  $\sqrt{a^2 - x^2}$  try  $x = a \sin u$  or  $x = a \cos u$ .

(ii)  $\sqrt{x^2 - a^2}$  try  $x = a \cosh u$  or  $x = a \sec u$ .

For integrals involving odd power of trigonometric ratios like:

$$\int \cos^n x \, dx \, (n \text{ odd}) \quad \text{try} \quad u = \sin x$$

$$\int \sin^n x \, dx \, (n \text{ odd}) \quad \text{try} \quad u = \cos x.$$

It would be useful to memorise these substitutions!

**Integration by parts**

This method (formula 3) is used mainly where the function under consideration is a *product*, in a similar way to the product rule for differentiation (the reverse process). In comparison with the previous substitution methods, it is relatively straightforward to use. The following examples illustrate the method very well.

**Example 6.3.19**

Find  $\int \ln x \, dx$ .

On first sight it seems that this function is not a product and a substitution might be appropriate. Although the integral can

be solved in this way, it is much easier to integrate by parts when we consider the integral as:

$$\int \ln x(1) dx.$$

Now applying the rule (in functional notation)

$$\int uv' dx = uv - \int vu' dx$$

$$\text{with } u = \ln x \text{ so } u' = \frac{1}{x} \text{ or } u = \ln x \text{ so } \frac{du}{dx} = \frac{1}{x}$$

$$\text{and } v' = 1 \text{ so } v = x \text{ or } \frac{dv}{dx} = 1 \text{ so } v = x$$

$$\text{then } \int uv' dx = x \ln x - \int x \frac{1}{x} dx$$

$$\int \ln x dx = x \ln x - x + c$$

(constant of integration is often assumed).

Check carefully that you can follow this process noting that

we started by letting  $\ln x = u$  and  $1 = \frac{dv}{dx}$  or  $v'$ .

### Example 6.3.20

Find  $\int x \cos^2 x dx$ .

This is quite clearly a product, however the integral of  $\cos^2 x$  is not straightforward, but by using a trigonometric substitution we can simplify the integral before we start. Referring back to formula 4(ii) in the trigonometry where:

$$\cos 2A = 2 \cos^2 A - 1 \quad \text{or} \quad \cos 2A + 1 = 2 \cos^2 A.$$

Then for above integral:

$$\begin{aligned} \int x \cos^2 x dx &= \int x \left( \frac{1 + \cos 2x}{2} \right) dx \\ &= \int \frac{x}{2} dx + \int \frac{x}{2} \cos 2x \\ &= \frac{x^2}{4} + \int \frac{x}{2} \cos 2x \end{aligned}$$

By parts

$$\begin{aligned} \int x \cos^2 x dx &= \frac{x^2}{4} + \left[ \left( \frac{x}{2} \frac{\sin 2x}{2} \right) - \int \frac{1}{2} \frac{\sin 2x}{2} dx \right] \\ &= \frac{x^2}{4} + \frac{x \sin 2x}{4} - \left( -\frac{\cos 2x}{8} \right) \\ &= \frac{x^4}{4} + \frac{x \sin 2x}{4} + \frac{\cos 2x}{8} + c. \end{aligned}$$

**Example 6.3.21**

Find  $\int x e^x dx$ .

Let  $u = x$  and  $v' = e^x$

then  $u' = 1$  and  $v = e^x$ .

Using  $\int uv' dx = uv - \int vu' dx$ .

Then  $\int x e^x dx = x e^x - \int (e^x)(1) dx$

$$\int x e^x dx = x e^x - e^x + c.$$

What about finding  $\int x^2 e^x dx$  ?

Lets try the same process.

$u = x^2$  and  $v' = e^x$  then  $u' = 2x$  and  $v = e^x$ .

So by parts:

$$\int x^2 e^x dx = x^2 e^x - \int 2x e^x dx.$$

The right-hand integral can be integrated by parts again! With

$$u = 2x \quad u' = 2 \quad v' = e^x \quad v = e^x$$

$$\int x^2 e^x dx = x^2 e^x - [2x e^x - 2 \int e^x dx]$$

taking constant behind integral sign.

$$\text{So } \int x^2 e^x dx = x^2 e^x - [2x e^x - 2e^x]$$

$$\int x^2 e^x dx = e^x (x^2 - 2x + 2)$$

constant assumed.

If you were able to follow the process illustrated in examples 6.3.19–6.3.21, you should be able to integrate products involving polynomial, trigonometric, and exponential functions. There is no space here to show some of the more intricate techniques, which you may require. Integration, unlike differentiation, is often considered to be something of an art rather than a science. With *practice* you will be able to master the art!

**Numerical integration**

Up till now, the integrals we have been dealing with have all been clearly defined by a *rule*. In many engineering situations we are required to make sense of data which may result from experimentation or, be produced during system/component operation. If we have no clearly defined rule governing the summing process or the integral cannot be evaluated using analytical methods. Then, we need another method of solution, we use a *numerical summing process*. Two simple, but very powerful techniques

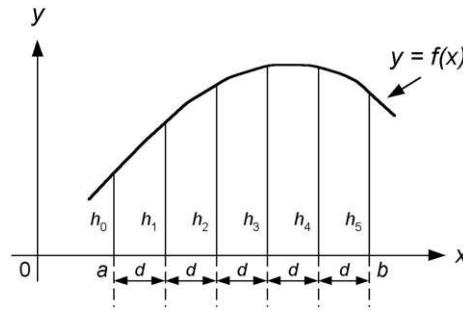


Figure 6.3.2 Trapezoidal rule

of numerical integration (summing) are given next. These techniques involve the use of Simpson's rule and the Trapezoidal rule (formulae 4 and 5).

The Trapezoidal rule. This rule is illustrated in Figure 6.3.2.

If the area represented by  $\int_a^b f(x) dx$  is divided into strips (Figure 6.3.2), each of equal width, then each strip is approximately a trapezium. The area of a trapezium is equal to half the sum of the parallel sides  $h$ , multiplied by the distance  $d$ , between them. If we use the sum of these areas as an approximation for the actual value of the area under the integral sign, then:

$$\begin{aligned} \int_a^b f(x) dx &\simeq \frac{1}{2}d(h_0 + h_1) + \frac{1}{2}d(h_1 + h_2) + \frac{1}{2}d(h_2 + h_3) \\ &\quad + \frac{1}{2}d(h_3 + h_4) + \frac{1}{2}d(h_4 + h_5) \\ &\simeq \frac{1}{2}d[h_0 + 2h_1 + 2h_2 + 2h_3 + 2h_4 + h_5]. \end{aligned}$$

Now in general the rule depends on the number of ordinates chosen so the rule (formula 5) may be written as:

$$\int_a^b f(x) dx \simeq \frac{1}{2}d[h_0 + 2h_1 + \dots + 2h_{n-2} + \dots + h_{n-1}].$$

### Example 6.3.22

Use the Trapezoidal rule, with five ordinates to evaluate

$$\int_0^1 e^{x^2} dx.$$

This integral could be solved by integration, using a substitution, but it serves as a simple example of the process.

We have five ordinates  $h_0, h_1, h_2, h_3, h_4$ , evenly spaced between  $h_0 = 0$  and  $h_4 = 1$ , so  $d = 0.25$ .

Then when  $h = e^{x^2}$

|                    |            |       |            |
|--------------------|------------|-------|------------|
| $h_0 = e^0$        | $= 1$      | where | $d = 0$    |
| $h_1 = e^{0.0625}$ | $= 1.0645$ | where | $d = 0.25$ |
| $h_2 = e^{0.25}$   | $= 1.2840$ | where | $d = 0.5$  |
| $h_3 = e^{0.5625}$ | $= 1.7551$ | where | $d = 0.75$ |
| $h_4 = e^1$        | $= 2.7183$ | where | $d = 1$ .  |

So the Trapezoidal rule gives:

$$\begin{aligned}
 \int_0^1 e^{x^2} &\approx \frac{1}{2} d[h_0 + 2h_1 + 2h_2 + 2h_3 + h_4] \\
 &\approx \left(\frac{1}{2}\right)(0.25)[1 + (2)(1.0645) + 2(1.284) \\
 &\quad + 2(1.7551) + 2.7183] \\
 &\approx 0.125(1 + 2.129 + 2.568 + 3.5102 + 2.7183) \\
 &\approx \mathbf{1.366}.
 \end{aligned}$$

#### *Simpson's rule*

Simpson's rule will not be proved. It can however, be easily used by applying the version of the rule given in formula 4.1.

Then for Simpson's rule use:

$$\frac{1}{3} d\{(I + F) + 4(ODD) + 2(EVEN)\}$$

where  $I$  = the initial ordinate ( $h_0$ ),

$F$  = the final ordinate ( $h_{2n}$ ).

ODD = the sum of the intervening odd ordinates,  $h_1, h_3 \dots$

EVEN = the sum of the intervening even ordinates,  $h_2, h_4 \dots$

### **Example 6.3.23**

Use Simpson's rule to evaluate the integral of Example 6.3.22, with five ordinates.

Then we evaluate  $\int_0^1 e^{x^2} dx$ .

Simpson's rule written out in full is:

$$\begin{aligned}
 \int_a^b f(x) dx &\approx \frac{1}{3} d[h_0 + 4h_1 + 2h_2 + 4h_3 + 2h_4 \\
 &\quad + \dots + 4h_{2n-1} + h_{2n}] \\
 \int_0^1 e^{x^2} dx &\approx \frac{1}{3} d[(I + F) + 4(ODD) + 2(EVEN)].
 \end{aligned}$$

Now we have already calculated the values in Example 6.3.22, they are:

$$\begin{aligned}
 h_0 &= 1 & d &= 0 \\
 h_1 &= 1.0645 & d &= 0.25 \\
 h_2 &= 1.2840 & d &= 0.5 \\
 h_3 &= 1.7551 & d &= 0.75 \\
 h_4 &= 2.7183 & d &= 1.
 \end{aligned}$$

Then using Simpson's rule:

$$\begin{aligned}
 \int_0^1 e^{x^2} dx &\approx \frac{1}{3}(0.25)[(1 + 2.7183) + 4(2.8196) + 2(1.2840)] \\
 &\approx 0.0833(3.7183 + 11.2784 + 2.568) \\
 &\approx \mathbf{1.463}.
 \end{aligned}$$

Compare this value with that obtained using the Trapezoidal rule. Simpson's gives a more accurate solution.



## Engineering applications

We now turn our attention to the numerous engineering applications of the calculus. These will include examples on rates of change, centroids, second moment of area, second moment of mass, root mean square (rms) values of waves, and numerical summation of engineering functions. Some important formulae concerned with engineering applications are also given in this section, for convenience.

### Formulae

#### (1) Areas and volumes of revolution

$$\text{Area } A = \int_a^b f(x) \, dx$$

Volume of revolution obtained by rotating area  $A$  through four right-angles.

$$(i) \text{ about } 0x = \pi \int_a^b [f(x)]^2 \, dx$$

$$(ii) \text{ about } 0y = 2\pi \int_a^b x f(x) \, dx.$$

#### (2) Centroids of plane areas

For the figure, if  $\bar{x}$ ,  $\bar{y}$  denote the co-ordinates of the centroid of area  $A$ , then:

$$\bar{x} = \frac{\int_a^b x f(x) \, dx}{\int_a^b f(x) \, dx}; \quad \bar{y} = \frac{\frac{1}{2} \int_a^b [f(x)]^2 \, dx}{\int_a^b f(x) \, dx}.$$

#### (3) Second moment of area

If  $I_x$  and  $I_y$  denote second moments of area  $A$  about  $0x$  and  $0y$  respectively, then:

$$I_x = \frac{1}{3} \int_a^b y^3 \, dx \quad \text{and} \quad I_y = \int_a^b x^2 y \, dx$$

#### (4) Second moment of mass (moment of inertia)

If the mass per unit volume of the volume of revolution generated by the rotation of area  $A$  about  $0x$  is  $m$ , then the moment of inertia of the solid about  $0x$  is:

$$I_{0x} = \frac{1}{2} \int_a^b m\pi y^4 \, dx.$$

#### (5) Theorem of Pappus

Volume of revolution obtained by rotating area  $A$  through four right-angles:

$$(i) \text{ about } 0x = 2\pi yA;$$

$$(ii) \text{ about } 0y = 2\pi xA.$$

#### (6) Mean and rms values

If  $y = f(x)$ , then the **mean value** of  $y$  over the range  $x = a$  to  $x = b$  is given by:

$$\text{Mean value} = \frac{1}{b-a} \int_a^b f(x) \, dx$$

and the **rms value** over the same range is given by:

$$\text{rms value} = \sqrt{\frac{1}{b-a} \int_a^b [f(x)]^2 dx}.$$

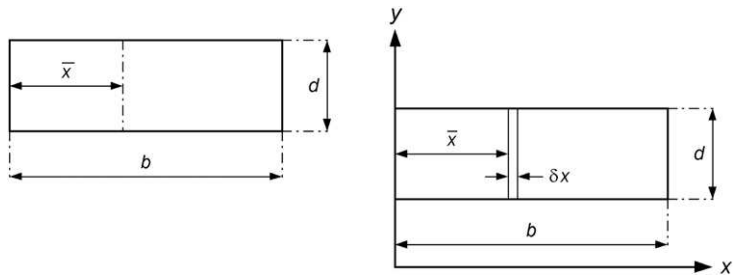
### Centroids of area

The *centroid* of an area is the point at which the total area is considered to be situated for calculation purposes. It is needed in the calculation of second moments of area, which follows. For simple shapes such as rectangles and circles, the centroid is easily found. There are however, many more complex shapes which cannot be divided into these standard shapes. For these areas, we need a mathematical summing technique to establish their centroid. The following two examples illustrate the process.

*Note:* The centroid is found from conveniently positioned axes.

#### Example 6.3.24

Find by integration the value of  $\bar{x}$  (i.e. the distance of the centroid from left-hand edge) for the rectangle shown below.



*Determination of second moment of area of rectangle*

The rectangle must be set up on suitable axes as shown above. We include a very small elemental strip.

Now the area of the rectangle =  $\sum A$

where

$\sum A$  = the sum of the elementary strip areas

$$\begin{aligned} &= \sum_{x=0}^{x=b} d \cdot dx \\ &= \int_0^b d \cdot dx = d \int_0^b 1 \cdot dx \\ &= d[x]_0^b = d(b - 0) = b \cdot d. \end{aligned}$$

Also the first moment of the rectangular area about the  $y$ -axis:

$$\begin{aligned} &= \sum A \cdot x \\ &= \text{sum of the first moment of area of each of the elementary strips} \\ &= \text{sum of: (area of strip} \times \text{distance of its centroid from the } y\text{-axis)} \end{aligned}$$

$$\begin{aligned}
 &= \sum_{x=0}^{x=b} d \, dx \cdot x \\
 &= \int_0^b d \, dx \cdot x = d \int_0^b x \cdot dx = d \left[ \frac{x^2}{2} \right]_0^b \\
 &= d \left( \frac{b^2}{2} - \frac{0^2}{2} \right) = \frac{b^2 \cdot d}{2}.
 \end{aligned}$$

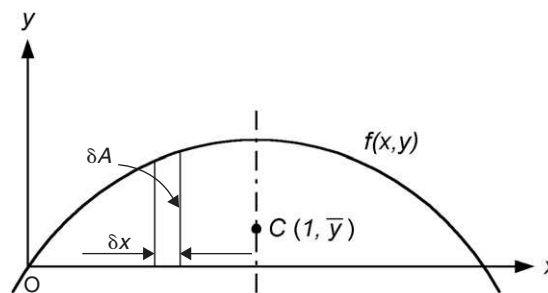
$$\text{and so } \bar{x} = \frac{\sum A \cdot x}{\sum A} = \frac{b^2 d / 2}{bd} = \frac{b}{2}.$$

This of course is what we would expect from symmetry!

### Example 6.3.25

Consider the area bounded by the curve  $y = x(2 - x)$  and the  $x$ -axis.

Always draw a sketch and set up axes.



Area bounded by curve  $y = x(2 - x)$  and the  $x$ -axis

From symmetry the centroid  $C$  lies on  $x = 1$ , that is  $\bar{x} = 1$ . To find  $\bar{y}$  we consider the first moment about  $Ox$ . Taking a vertical strip as shown, we have:

$$\delta A \approx y \, \delta x.$$

The centroid of the element is approximately a distance of  $y/2$  from  $Ox$ . Therefore, the first moment of  $\delta A$  about  $Ox \approx (y \, \delta x) \frac{1}{2} y$

$$\text{so } \bar{y} \approx \frac{\sum_{x=0}^2 \frac{1}{2} y^2 \, dx}{\sum_{x=0}^2 y \, dx}$$

$$\text{or } \bar{y} = \frac{\int_0^2 \frac{1}{2} y^2 \, dx}{\int_0^2 y \, dx} \quad \text{where } y = x(2 - x)$$

$$\bar{y} = \frac{\frac{1}{2} \int_0^2 x^2 (2-x)^2 dx}{\int_0^2 x(2-x) dx} = \frac{\left[ \frac{4x^3}{3} - x^4 + \frac{x^5}{5} \right]_0^2}{\left[ x^2 - \frac{x^3}{3} \right]_0^2}$$

$$= \frac{\frac{8}{15}}{\frac{4}{3}} = \frac{2}{5}$$

So centroid at  $\bar{x} = 1$ ,  $\bar{y} = \frac{2}{5}$ .

### Second moment of area

In engineers bending theory you will use the relationship:

$$\frac{M}{I} = \frac{\sigma}{y} = \frac{E}{R}$$

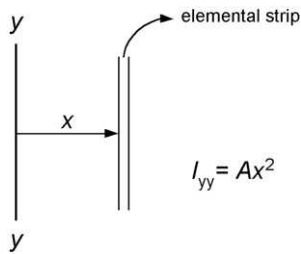


**Figure 6.3.3** Method for calculating second moment of area

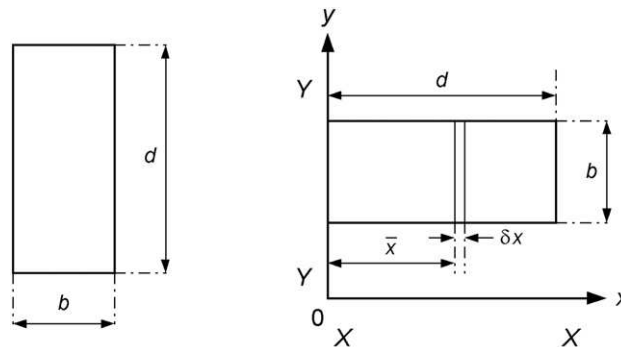
where  $I$  is known as the second moment of area about the cross-section, it is in fact a measure of the resistance to the bending of a structural member, based on the geometry of its cross-section. It is an important engineering property and is calculated in the following way.

Figure 6.3.3 shows an elemental strip with very small width of area  $A$ , which is distance  $x$  from the  $YY$ -axis.

Now from our work on centroids we know that the first moment of the area  $AA$  about  $YY$  is given by  $Ax$ . Then the *second moment* of the area  $A$  about  $YY$  is given by  $Ax^2$ . The second moment of area is always stated with reference to an axis or datum line. So in the above case  $I_{yy} = Ax^2$ .

### Example 6.3.26

Find the second moment of area of the rectangle shown, about its base edge.



Second moment of area for rectangle

The rectangle is shown set up on suitable reference axes. It is in this case convenient to turn the rectangle through  $90^\circ$  and let the base edge lie on the  $y$ -axis ( $YY$ ). The diagram shows a typical elemental strip area parallel to the reference axis ( $YY$ ), whose area is  $b \cdot \delta x$ .

Now second moment of the rectangular area:

$$\text{about the } y\text{-axis} = \sum A \cdot x^2 = \sum_{x=0}^{x=d} b \cdot \delta x \cdot x^2 = \int_0^d b \cdot dx \cdot x^2.$$

$$\text{So } I_{yy} = b \int_0^d x^2 dx = b \left[ \frac{x^3}{3} \right]_0^d = \frac{bd^3}{3}.$$

In engineers theory of torsion where:

$$\frac{T}{J} = \frac{\tau}{R} = \frac{G\theta}{L}$$

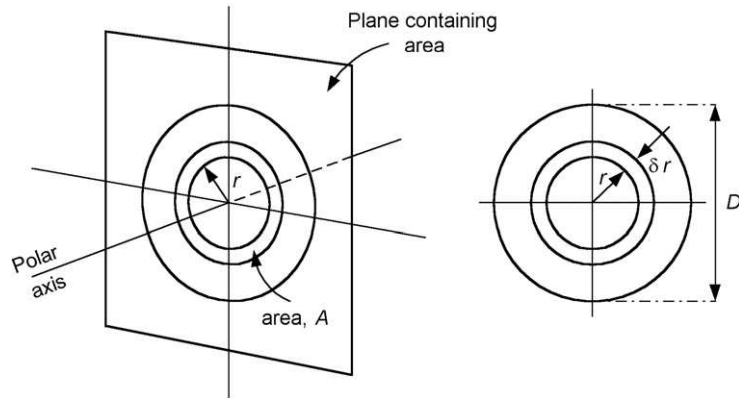
$J$  is known as the *polar second moment of area* of the cross-section.

It is a measure of the resistance of a shaft to torsional loads, an expression for its value may be found in a similar manner to the method illustrated in Example 6.3.26.

### Example 6.3.27

Find the polar second moment of area ( $J$ ) of the circular area shown in the figure.

As before we consider the second moment of area of the elemental strip, set-up on the polar axis as shown.



*Polar second moment of area*

Then  $J$  for the circular area  $= \sum Ar^2$ . Now approximate area  $A$  of the elemental strip  $=$  (circumference of strip  $\times$  width of strip).

$$\text{So } A \approx 2\pi r \cdot \delta r.$$

$$\begin{aligned}
 \text{Therefore } J &= \sum_{r=0}^{r=D/2} 2\pi r \cdot \delta r \cdot r^2 \\
 &= \int_0^{D/2} 2\pi r \cdot dr \cdot r^2 \\
 &= 2\pi \int_0^{D/2} r^3 dr \\
 &= 2\pi \left[ \frac{r^4}{4} \right]_0^{D/2} \\
 &= \frac{\pi D^4}{32}.
 \end{aligned}$$

### Moment of inertia

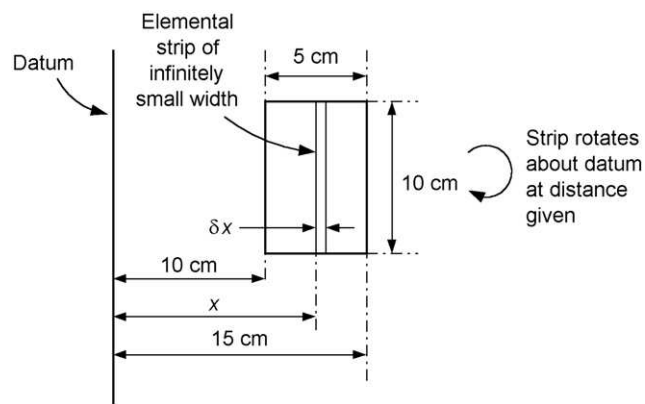
The second moment of mass (*moment of inertia*) of a rotating body is found in a similar way to the second moment of area. You met the moment of inertia of rotating masses in Chapter 3 when you studied angular motion.

The moment of inertia  $I$ , of a body about a given axis is the sum of the products of each element of mass and the square of its distance from a given axis (same as for moment of area except mass is taken into account). The procedure we adopt is therefore the same as that for second moment of area.

### Example 6.3.28

Find the *moment of inertia* of a rectangular lamina (solid area without depth) of length 10 cm and breadth 5 cm about an axis parallel to the 10 cm side and 10 cm from it. Take the area density of the lamina to be  $\rho \text{ kg/cm}^2$ .

The figure illustrates the situation.



*Determination of moment of inertia*

Then mass of element =  $10\rho\delta x \text{ kg}$

$I$  (second moment) for element =  $10\rho x^2 \delta x \text{ m}^2$

$$\begin{aligned}
 I \text{ for rectangle} &= \int_{10}^{15} 10\rho x^2 \, dx \\
 &= 10\rho \left[ \frac{x^3}{3} \right]_{10}^{15} \\
 &= 10\rho \left[ \frac{15^3}{3} - \frac{10^3}{3} \right] \\
 &= 10\rho \left[ 1125 - 333\frac{1}{3} \right] \\
 &= 7916\frac{2}{3} \rho \text{ kg cm}^2.
 \end{aligned}$$

### Mean and rms values

We have already used the integral calculus to find areas under curves, when you integrated between limits. The average or *mean value* of the height of such curves above the limits of integration may easily be found by dividing the area obtained (between the limits), by the distance between these limits (Figure 6.3.4). From this process we obtain the expression:

$$\text{Mean value of height} = \frac{\int_a^b f(x) \, dx}{b - a} = \frac{\text{Area}}{\text{Horizontal distance}}$$

$$\text{so mean value} = \frac{1}{b - a} \int_a^b f(x) \, dx.$$

So, for example if the function represented in Figure 6.3.4 was given by:

$$\int_2^4 e^x \, dx$$

Then the mean value would be given by:

$$\begin{aligned}
 \frac{1}{4 - 2} \int_2^4 e^x \, dx &= \frac{1}{2} [e^x]_2^4 \\
 &= \frac{1}{2} [e^4 - e^2] \\
 &= \frac{1}{2} [54.6 - 7.39] \\
 &= 23.6 \text{ units of length.}
 \end{aligned}$$

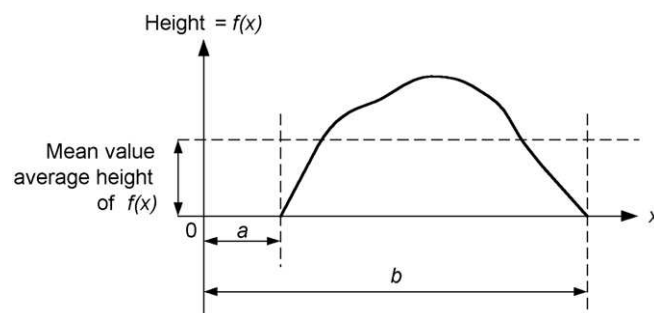


**Figure 6.3.4** Mean value of curve

When finding average values of sinusoidal wave forms, we need to consider the area above and below the  $x$ -axis. If we square the mean value of the sinusoidal wave its height ( $y$ ) will always be positive and, if we then square root this value we obtain the average or the *rms value*.

So considering our formula for the mean value of a function we get:

$$\text{Mean value } f(x) = \frac{1}{b-a} \int_a^b f(x) dx.$$

$$\text{So squaring gives } (f(x))^2 = \frac{1}{b-a} \int_a^b (f(x))^2 dx$$

and taking the square root gives:

$$\text{rms value} = \sqrt{\frac{1}{b-a} \int_a^b (f(x))^2 dx}.$$

The rms value of sinusoidal waveforms is particularly useful when determining power and other parameters from AC voltages and currents. The following example shows how we calculate the rms value for sinusoidal wave forms.

### Example 6.3.29

Find the rms value of the function  $f(x) = \sin x$  between 0 and  $\pi$  (radians).

$$\text{Then using } \text{rms} = \sqrt{\frac{1}{b-a} \int_a^b (f(x))^2 dx}$$

$$\text{then } \text{rms} = \sqrt{\frac{1}{\pi-0} \int_0^\pi \sin^2 \theta d\theta}$$

and using the identity  $\cos 2A = 1 - 2 \sin^2 A$

$$\text{then } \sin^2 \theta = \frac{1}{2}(1 - \cos 2\theta).$$

$$\begin{aligned} \text{So } \text{rms} &= \sqrt{\frac{1}{\pi-0} \int_0^\pi \frac{1}{2}(1 - \cos 2\theta) d\theta} \\ &= \sqrt{\frac{1}{2\pi} \left[ \theta - \frac{\sin 2\theta}{2} \right]_0^\pi} \quad \left( \begin{array}{l} \text{constant } \frac{1}{2}, \text{ put outside} \\ \text{integral sign} \end{array} \right) \\ &= \sqrt{\frac{1}{2\pi} \left[ \left( \pi - \frac{\sin 2\pi}{2} \right) - \left( 0 + \frac{\sin 0}{2} \right) \right]} \\ &= \sqrt{\frac{1}{2\pi} (\pi)} = \sqrt{\frac{1}{2}} = 0.7071. \end{aligned}$$

So **rms = 0.7071** (I hope this is what you expected!)

Since the sinusoidal function is squared and then square rooted, the wave repeats itself every  $\pi$  rad, so the rms value for the whole sinusoidal wave form is always 0.7071.

### Numerical method for buoyancy

The distance and velocity of falling bodies in water, or other resistive mediums can be related by the integral calculus. The following example



illustrates the use of the Trapezoidal rule and Simpson's rule to solve a problem, where the relationship has been established by experiment.

### Example 6.3.30

The results of an experiment show that for a particular body falling in a resistive medium the distance that it has fallen when its velocity is 0.2 m/s, is given by:

$$\int_0^{0.2} \frac{v}{9.81 - 0.2v^2} dv$$

Evaluate this distance using the Trapezoidal rule and Simpson's rule, take  $d = 0.25$ . For this integral, we will set up a table.

| $v_i$        | $v$ initial<br>and final | $\frac{v}{9.81 - 0.2v^2}$<br>ODD | $\frac{v}{9.81 - 0.2v^2}$<br>EVEN |
|--------------|--------------------------|----------------------------------|-----------------------------------|
| $v_0 = 0$    | $v_0 = 0$                |                                  |                                   |
| $v_1 = 0.25$ |                          | 0.025 52                         |                                   |
| $v_2 = 0.5$  |                          |                                  | 0.051 23                          |
| $v_3 = 0.75$ |                          | 0.077 34                         |                                   |
| $v_4 = 1.0$  |                          |                                  | 0.104 06                          |
| $v_5 = 1.25$ |                          | 0.131 61                         |                                   |
| $v_6 = 1.5$  |                          |                                  | 0.160 26                          |
| $v_7 = 1.75$ |                          | 0.190 27                         |                                   |
| $v_8 = 2.0$  | $v_8 = 0.221\ 98$        |                                  |                                   |
| Total        | 0.221 98                 | 0.424 74                         | 0.315 55                          |

Then using the abbreviated versions of both rules, we get:

(i) From Trapezoidal rule:

$$\begin{aligned} \int_0^{0.2} \frac{v}{9.81 - 0.2v^2} dv &\approx d \left[ \frac{1}{2}(I + F) + ODD + EVEN \right] \\ &\approx 0.25[0.110\ 99 + 0.424\ 74 + 0.315\ 55] \\ &\approx 0.25(0.85128) \\ &\approx \mathbf{0.212\ 82\ m.} \end{aligned}$$

(ii) From Simpson's rule:

$$\begin{aligned} \int_0^{0.2} \frac{v}{9.81 - 0.2v^2} dv &\approx \frac{d}{3} [(I + F) + 4(ODD) + 2(EVEN)] \\ &\approx \frac{0.25}{3} [0.22198 + 4(0.424\ 74) \\ &\quad + 2(0.315\ 55)] \\ &\approx 0.083\ 33(0.22198 + 1.698\ 96 + 0.6311) \\ &\approx \mathbf{0.212\ 67\ m.} \end{aligned}$$

### Rates of change

The following simple examples illustrate the use of the differential calculus to determine the rate of change of variables with respect to time.

**Example 6.3.31**

The motion of a body is modelled by the relationship  $s = t^3 - 3t^2 + 3t + 8$ , where  $s$  is distance in metres and  $t$  time in seconds. Determine:

- (i) the velocity of the body at the end of 3 s;
- (ii) the time when the body has zero velocity;
- (iii) its acceleration at the end of 2 s;
- (iv) when its acceleration is zero.

In order to solve this problem you must be aware of the relationship between distance, velocity, and acceleration with respect to time.

The rate of change of distance with respect to time is given by  $ds/dt$  and is the *velocity*. Similarly the rate of change of velocity with respect to time is given by  $dv/dt$  and is the acceleration of the body.

$$\text{So } v = \frac{ds}{dt} \quad \text{and} \quad a = \frac{d^2s}{dt^2} = \frac{dv}{dt}.$$

Now the problem is simple:

$$s = t^3 - 3t^2 + 3t + 8 \quad \text{then:}$$

$$(i) \quad \frac{ds}{dt} = 3t^2 - 6t + 3 = \text{velocity}$$

so at 3 s the

$$\text{velocity} = 3(3)^2 - 6(3) + 3 = \mathbf{12 \text{ m/s}}.$$

- (ii) The body will have zero velocity when  $ds/dt = 0$ ; that is, when  $3t^2 - 6t + 3 = 0$ . So we need to solve this quadratic to determine the time  $t$  when the velocity is zero.

$$\text{Then } t^2 - 2t + 1 = 0$$

$$\text{so } (t - 1)(t - 1) = 0.$$

Then  $ds/dt = 0$  when  $t = 1$  or **body has velocity of zero at  $t = 1$  s.**

- (iii) The acceleration ( $a$ ) is obtained as stated previously, that is:

$$\frac{d^2s}{dt^2} = 6t - 6 \quad \text{so when } t = 2$$

$$\text{The body has acceleration } \frac{d^2s}{dt^2} = a = \mathbf{6 \text{ m/s}^2}.$$

- (iv) Acceleration is zero when  $d^2s/dt^2 = 6t - 6 = 0$  so **acceleration is zero at time  $t = 1$  s.**

**Example 6.3.32**

A particle is subject to harmonic motion given by the relationship  $x = A \sin \omega t$ . Show that the linear acceleration of the particle is given by  $a = \omega^2 x$ . Where  $\omega$  = angular velocity and  $x$  = the linear displacement from the centre of oscillation.

This again relates displacement, velocity, and acceleration, so all we need do is differentiate twice.

$$\text{Then } x = A \sin \omega t$$

$$\text{so } v = \frac{dx}{dt} = A\omega \cos \omega t$$

$$\text{and } \frac{d^2x}{dt^2} = -A\omega^2 \sin \omega t, \quad \text{but } x = A \sin \omega t$$

$$\text{so } \frac{d^2x}{dt^2} = -\omega^2 x.$$

We leave our study of the calculus with a number of problems which have been designed to give you practice in manipulating the calculus as well as applying it. The problems are of varying difficulty, in no particular order.

**Questions 6.3.1**

- (1) Differentiate the following using function of a function rule (i.e. by substitution) or chain rule.

$$\begin{array}{lll} \text{(a)} (3x + 1)^2, & \text{(b)} (2 - 5x)^{3/2}, & \text{(c)} \frac{1}{4x^2 + 3}, \\ \text{(d)} \sin^2 4x, & \text{(e)} \cos(2 - 5x), & \text{(f)} \log_e 9x, \end{array}$$

$$\text{(g) Find } \frac{d}{dt} \left( \frac{1}{\sqrt[3]{1 - 2t}} \right),$$

$$\text{(h) Find } f'(t) \text{ for the function } f(t) = Be^{Kt-b}.$$

- (2) Differentiate the following using the product rule as required.

$$\text{(a)} x \sin x, \quad \text{(b)} e^x \tan x, \quad \text{(c)} x \ln x, \quad \text{(d)} \frac{d}{dt} (\sin t \cos t),$$

$$\text{(e) Find } \frac{d}{dt} [6e^{3t}(t^2 - 1)], \quad \text{(f) Find } \frac{d}{ds} [(s - 3s^2) \log_e s].$$

- (3) Differentiate the following:

$$\text{(a)} \frac{x}{1 - x}, \quad \text{(b)} \frac{e^x}{3 \sin 2x}, \quad \text{(c)} \frac{(x^2 - 2)^2}{(3x + 4)^3},$$

$$\text{(d) Find } \frac{d}{dt} \left( \frac{\cos 2t}{e^{2t}} \right), \quad \text{(e) Find } \frac{d}{d\theta} (\cot \theta).$$

- (4) A curve is given in the form:  $y = 3 \cos 2\theta - 5 \tan \theta$  where  $\theta$  is in radians. Find the gradient of the curve at the point where  $\theta$  has a value equivalent to  $34^\circ$ .
- (5) Find the gradient of the curve  $\cos \theta / \theta$  at the point where  $\theta = 0.25$  (remember radians).
- (6) If  $y = 4 \log_e(1 - x)$ , find the value of  $\frac{dy}{dx}$  when  $x = 0.32$ .
- (7) If  $y = \frac{1 + x^2}{x - 2}$  find the value of  $\frac{dy}{dx}$  if  $x = -1.25$ .
- (8) If  $y = 3x^3 + 2x - 7$  find an expression for  $\frac{d^2y}{dx^2}$  and also its value when  $x = -3$ .
- (9) Given that  $y = \frac{3t^5 + 2t}{t^2}$  find the value of  $\frac{d^2y}{dx^2}$  when  $t = 0.6$ .
- (10) Find the value of  $\frac{d^2u}{dm^2}$  if  $u = \frac{1}{2}(e^{3m} - e^{-3m})$  given  $m = 1.3$ .
- (11) A body moves a distance  $s$  metres in a time  $t$  seconds so that  $s = 2t^3 - 9t^2 + 12t + 6$ . Find: (a) its velocity after 4 s; (b) its acceleration after 4 s; (c) when the velocity is zero.
- (12) The angular displacement  $\theta$  radians of the spoke of a wheel is given by the expression  $\theta = \frac{1}{2}t^4 - t^3$ , where  $t$  is the time in seconds. Find: (a) the angular velocity after 3 s; (b) the angular acceleration after 4 s; (c) when the angular acceleration is zero.
- (13) Find the maximum and minimum values of  $y$  given that:  $y = x^3 + 3x^2 - 9x + 6$ .
- (14) An open rectangular tank of height  $h$  metres with a square base of side  $x$  metres is to be constructed so that it has a capacity of  $500 \text{ m}^3$ . Prove that the surface area of the four walls and the base will be  $\frac{2000}{x} + x^2 \text{ m}^2$ . Find the value of  $x$  for this expression to be a minimum.
- (15) A mass of 5000 kg moves along a straight line so that the distance  $s$  metres travelled in a time  $t$  seconds is given by  $s = 3t^2 + 2t + 3$ . If  $v \text{ m/s}$  is its velocity and  $m \text{ kg}$  is its mass, then its kinetic energy is given by the formula  $\frac{1}{2}mv^2$ . Find its kinetic energy at a time  $t = 0.5 \text{ s}$ . Remember that the joule (J) is the unit of energy.
- (16) Find:
- $\int 4e^{-3x} - 2 \cos 3x - 5 \sin 2x \, dx$
  - $\int xe^{-x} \, dx$ ,
  - $\int \frac{1}{(5x + 1)(3 - 2x)} \, dx$ ,
  - $\int_a^b \frac{1}{x} \ln x \, dx$ ,
  - $\int \tan x \, dx$ ,
  - $\int \frac{1}{\sqrt{1 - x^2}} \, dx$ .

(17) Evaluate:

(a)  $\int_1^4 \sqrt[3]{2+3x} \, dx,$

(b)  $\int_0^{\pi/2} (\sin 2x)(\sin x) \, dx,$

(c)  $\int_0^1 x(x^2 + 2)^6 \, d\pi,$

(d)  $\int_1^3 \sin^3 x \, dx.$

(18) Evaluate  $\int_0^2 x^2 e^{-x^2} \, dx$  using Simpson's rule taking  $d = 0.25$ , to integrate the function numerically.

(19) The torque on a clutch plate is given by:

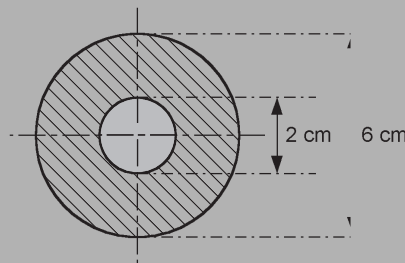
$$T = \int_{r_i}^{r_0} 2\pi\rho\mu r^2 \, dr.$$

If  $r_0 = 0.15 \, \text{m}$ ,  $r_i = 0.08 \, \text{m}$ ,  $\mu = .2$ , and  $\rho = 2000 \, \text{N/m}^2$ , find the value of the torque  $T$  in Nm.

(20) Find the mean and rms value of the function  $f(x) = \sin \theta + \sin 2\theta$ , between the ordinates  $\theta = 0$  and  $\frac{2\pi}{3}$  rad.

(21) Determine the position of the centroid of the area enclosed between the curve  $y = 3x^2 + 1$  and the  $x$ -axis, and values at  $x = 4$  and  $x = 2$ .

(22) Prove that the second moment of area  $J$  of a circle about its Polar axis ( $zz$ ) is equal to  $\frac{\pi D^4}{32}$ , hence calculate the second moment of area of the section shown below.



(23) An equation relating the moment  $M$ , modulus  $E$ , second moment of area  $I$ , and deflection  $y$  of a beam is:

$$\frac{d^2 y}{dx^2} = \frac{M}{EI}.$$

Find an expression for the deflection  $y$ .

(24) If  $EI \frac{d^2 y}{dx^2} = (Rx - w_1(x - a) - w_2(x - b)).$

Find an expression for  $y$ . Then given that when  $x = 0$ ,  $y = 0$  and that when  $x = L = 2$ ,  $y = 0$ . Find the constants of integration given that one constant

$$A = \frac{L^2}{6} - \frac{3(L-1)^3}{6L} - \frac{4(L-1.5)^3}{6L}.$$

(25) The e.m.f. induced in a secondary circuit is given by:

$$e = -M \frac{dI}{dt} \text{ V and the e.m.f. induced in an inductor is given}$$

by  $e = -N_2 \frac{d\Phi}{dt} \text{ V}$ . Show that the mutual inductance

$M$  is given by the expression  $M = N_2 \frac{d\Phi}{dI}$ . If there are 2000 turns on the secondary ( $N_2$ ) and the rate of change of flux with current  $\left(\frac{d\Phi}{dI}\right)$  is  $20 \times 10^{-4} \text{ Wb/amp}$ , find the mutual inductance.

## 6.4 STATISTICS AND PROBABILITY

Your view of statistics has probably been formed from what you read in the papers, or what you see on the television. Results of surveys to show which political party is going to win the election, why men grow moustaches, if smoking damages your health, the average cost of housing by area, and all sorts of other interesting data! Well, statistics is used to analyse the results of such surveys and when used correctly, it attempts to eliminate the bias which often appears when collecting data on controversial issues.

Statistics is concerned with collecting, sorting, and analysing numerical facts, which originate from several observations. Statistical inferences and predictions are then made upon which engineering decisions may be based. To assist in the simulation process, these facts are collated and summarised and then often presented in the form of tables, charts or diagrams, etc.

In this brief introduction to statistics, we look at five specific areas. First, we consider the *collection and presentation of data*, in its various forms. We then look at how we measure such data, concentrating on *measures of central tendency and dispersion*, in particular we look at mean values, standard deviation, and variance. *Regression and linear correlation* are then considered, with their application to the interpretation and validity of experimental data, being emphasised. Next *probability theory* is considered as a precursor to the study of *probability distributions* and their use in determining confidence intervals and the reliability and quality of engineering components and systems.

This section on statistics and probability departs slightly from the usual way in which we have presented the remainder of the mathematics. In that, statistical formulae and techniques are first covered together with the theory and techniques associated with probability distributions. Once these techniques have been mastered. The *major applications* of all these techniques are covered together at the end of the section.

## Data manipulation

### Methods

Since data manipulation is much more concerned with techniques, there are *no appropriate formulae* that need to be included for this topic. Therefore, we start by looking directly at some of the methods used to manipulate and collate data. The application of such techniques will be emphasised in the examples, you will find later concerned with experimental and reliability data.

In almost all scientific, engineering and business journals, newspapers and government reports, statistical information is presented in the form of charts, tables, and diagrams, as mentioned above. We now look at a small selection of these presentation methods, including the necessary manipulation of the data to produce them.

### Charts

Suppose, as the result of a survey, we are presented with the following statistical data.

| <i>Major category of employment</i> | <i>Number employed</i> |
|-------------------------------------|------------------------|
| Private business                    | 750                    |
| Public business                     | 900                    |
| Agriculture                         | 200                    |
| Engineering                         | 300                    |
| Transport                           | 425                    |
| Manufacture                         | 325                    |
| Leisure industry                    | 700                    |
| Education                           | 775                    |
| Health                              | 500                    |
| Other                               | 125                    |

Now ignoring for the moment, the accuracy of this data! Let us look at typical ways of presenting this information in the form of charts, in particular the *bar chart* and *pie chart*.

**The bar chart** In its simplest form, **the bar chart** may be used to represent data by drawing individual bars (Figure 6.4.1), using the figures from the raw data (the data in the table).

Now the scale for the vertical axis, the number employed, is easily decided by considering the highest and lowest values in the table, 900 and 125, respectively. Therefore, we use a scale from 0 to 1000 employees. Along the horizontal axis, we represent each category by a bar of even width. We could just as easily have chosen to represent the data using column widths instead of column heights.

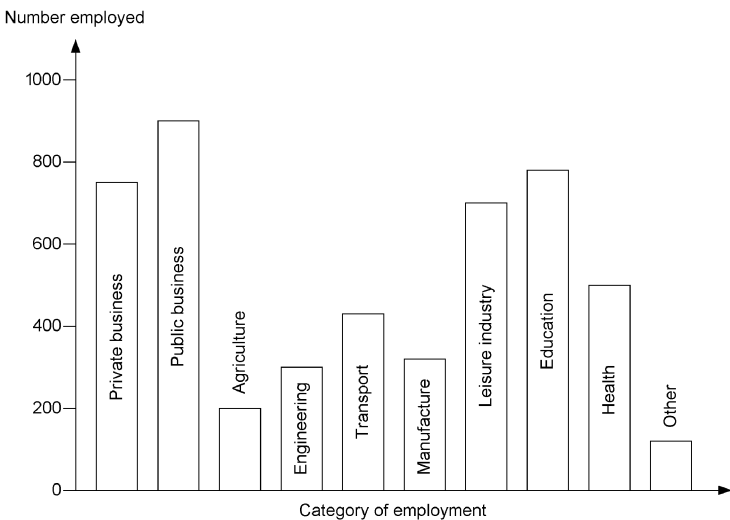


**Figure 6.4.1** Bar chart representing number employed by category

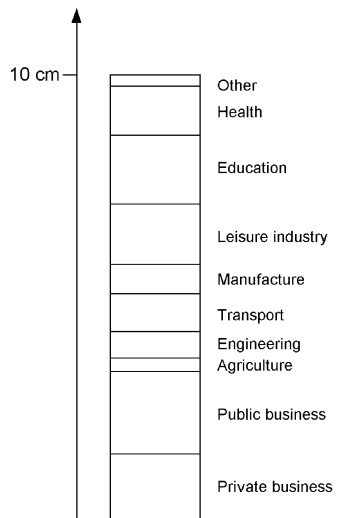


**Figure 6.4.2** A Proportionate bar chart graduated by height

Now the simple bar chart (Figure 6.4.1) tells us very little that we could not have determined from the table. So another type of bar chart, that enables us to make comparisons, *the proportionate bar chart*, may be used.

In this type of chart, we use *one bar*, with the same width throughout its height, with horizontal sections marked-off in proportion to the whole. In our example, each section would represent the number of people employed in each category, compared with the total number of people surveyed.

In order to draw a proportionate bar chart for our employment survey, we first need to total the number of people who took part in the survey, this total comes to 5000. Now, even with this type of chart we may represent the data either in proportion by height or in proportion by percentage. If we were to choose height, then we need to set our vertical scale at some convenient height say, 10 cm. Then we would need to carry out 10 simple calculations to determine the height of each individual column.

For example, given that the height of the total 10 cm represents 5000 people, then the height of the column for those employed in private business  $(750/5000)10 = 1.5$  cm, this type of calculation is then repeated for each category of employment. The resulting bar chart is shown in Figure 6.4.2.

### Example 6.4.1

Draw a proportionate bar chart for the employment survey, shown in the table, using the percentage method.

For this method all that is required is to find the appropriate percentage of the total (5000) for each category of employment. Then, choosing a suitable height of column to represent 100%, mark on the appropriate percentage for each of the 10 employment categories. To save space, only the first five categories of employment have been calculated in full.

$$(1) \text{ Private business} = \left( \frac{750}{5000} \right) \times 100 = 15\%$$

$$(2) \text{ Public business} = \left( \frac{900}{5000} \right) \times 100 = 18\%$$

$$(3) \text{ Agriculture} = \left( \frac{200}{5000} \right) \times 100\% = 4\%$$

$$(4) \text{ Engineering} = \left( \frac{300}{5000} \right) \times 100\% = 6\%$$

$$(5) \text{ Transport} = \left( \frac{425}{5000} \right) \times 100 = 8.5\%.$$

Similarly: manufacture = 6.5%, leisure industry = 14%, education = 15.5%, health = 10%, and other category = 2.5%.

Figure 6.4.3 shows the completed bar chart.

Other categories of bar chart include *horizontal bar charts*, where for instance Figure 6.4.1 is turned through 90° in a clockwise direction. One last type may be used to depict data given in chronological (time) order. Thus, for example, the horizontal x-axis is used to represent hours, days, years, etc. while the vertical axis shows the variation of the data with time.



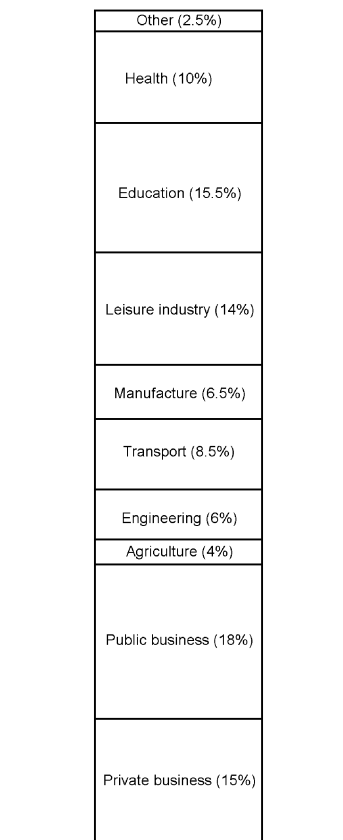


Figure 6.4.3 Proportionate percentage bar chart

**Example 6.4.2**

Represent the following data on a chronological bar chart.

| Year | Number employed in general engineering (thousands) |
|------|----------------------------------------------------|
| 1995 | 800                                                |
| 1996 | 785                                                |
| 1997 | 690                                                |
| 1998 | 670                                                |
| 1999 | 590                                                |

Since we have not been asked to represent the data on any *specific bar chart* we will use the simplest, involving only the raw data. Then, the only concern is the *scale* we should use for the vertical axis. To present a *true* representation, the scale should start from 0 and extend to say, 800 (Figure 6.4.4(a)). If we wish to emphasise a *trend*; that is the way the variable is rising or falling with time, we could use a very much exaggerated scale (Figure 6.4.4(b)). This immediately emphasises the *downward trend* since 1995. Note that this data is *fictitious* (made up) and used here merely for emphasis!

**Pie chart** In this type of chart the data is presented as a proportion of the total, using the angle or area of sectors. The method used to draw a pie chart is best illustrated by example.

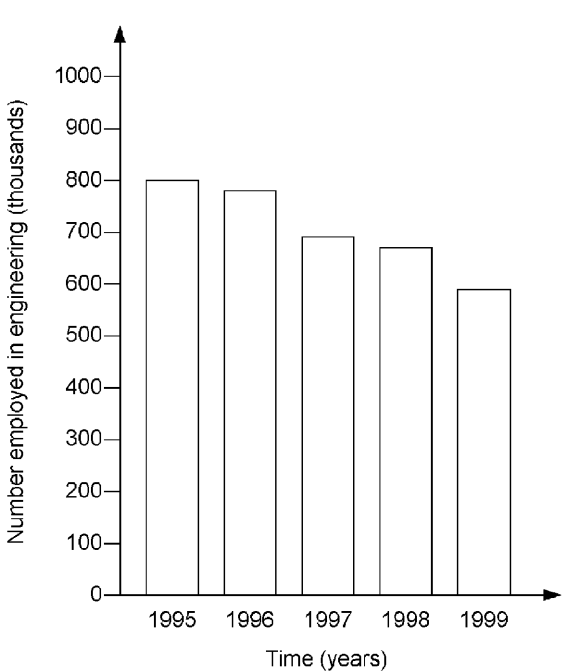


Figure 6.4.4a Chronological bar chart in correct proportion

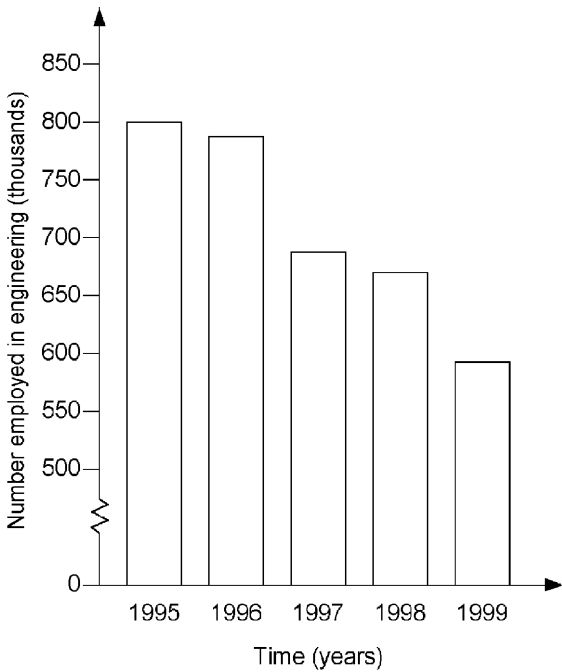


Figure 6.4.4b Chronological bar chart with graduated scale

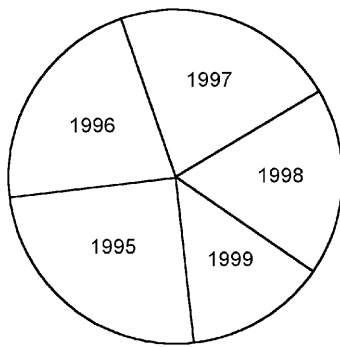


**Figure 6.4.5** Pie chart for Example 6.4.2 employment in engineering by year

### Example 6.4.3

Represent the data given in Example 6.4.2 on a pie chart.

Then, remembering that there are  $360^\circ$  in a circle and that the total number employed in general engineering (according to our figures) was  $800 + 785 + 690 + 670 + 590 = 3535$  (thousand).

Then, we manipulate the data as follows:

| Year  | Number employed in general engineering (thousands) | Sector angle (to nearest half degree)                   |
|-------|----------------------------------------------------|---------------------------------------------------------|
| 1995  | 800                                                | $\left(\frac{800}{3535}\right) \times 360 = 81.5^\circ$ |
| 1996  | 785                                                | $\left(\frac{785}{3535}\right) \times 360 = 80^\circ$   |
| 1997  | 690                                                | $\left(\frac{690}{3535}\right) \times 360 = 70.5^\circ$ |
| 1998  | 670                                                | $\left(\frac{670}{3535}\right) \times 360 = 68^\circ$   |
| 1999  | 590                                                | $\left(\frac{590}{3535}\right) \times 360 = 60^\circ$   |
| Total | 3535                                               | $= 360^\circ$                                           |

The resulting pie chart is shown in Figure 6.4.5.

Other methods of visual presentation include pictograms and ideographs. These are diagrams in pictorial form, used to present information to those who have a limited interest in the subject matter or who do not wish to deal with data presented in numerical form. They have little or no practical use when interpreting engineering or other scientific data and apart from acknowledging their existence, we will not be pursuing them further.

### Frequency distributions

One of the most common and most important ways of organising and presenting raw data is through use of *frequency distributions*.

Consider the data given below that shows the time in hours that it took 50 individual workers to complete a specific assembly line task.

*Data for assembly line task*

|     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1.1 | 1.0 | 0.6 | 1.1 | 0.9 | 1.1 | 0.8 | 0.9 | 1.2 | 0.7 |
| 1.0 | 1.5 | 0.9 | 1.4 | 1.0 | 0.9 | 1.1 | 1.0 | 1.0 | 1.1 |
| 0.8 | 0.9 | 1.2 | 0.7 | 0.6 | 1.2 | 0.9 | 0.8 | 0.7 | 1.0 |
| 1.0 | 1.2 | 1.0 | 1.0 | 1.1 | 1.4 | 0.7 | 1.1 | 0.9 | 0.9 |
| 0.8 | 1.1 | 1.0 | 1.0 | 1.3 | 0.5 | 0.8 | 1.3 | 1.3 | 0.8 |

From the data you should be able to see that the shortest time for completion of the task was 0.5 hours, the longest time was 1.5 hours. The *frequency* of appearance of these values is *once*. On the other hand

the number of times the job took 1 hour appears 11 times, or it has a *frequency of 11*. Trying to sort out the data in this ad hoc manner is time consuming and may lead to mistakes. To assist with the task we use a *tally chart*. This chart simply shows how many times the *event* of completing the task in a specific time takes place. To record the *frequency of events* we use the number 1 in a tally chart and when the frequency of the event reaches 5, we score through the existing four 1's to show a frequency of 5. The following example illustrates the procedure.

### Example 6.4.4

Use a tally chart to determine the frequency of events for the data given above on the assembly line task.

| <i>Time (hours)</i> | <i>Tally</i> | <i>Frequency</i> |
|---------------------|--------------|------------------|
| 0.5                 | 1            | 1                |
| 0.6                 | 11           | 2                |
| 0.7                 | 1111         | 4                |
| 0.8                 | 1111 1       | 6                |
| 0.9                 | 1111 111     | 8                |
| 1.0                 | 1111 1111 1  | 11               |
| 1.1                 | 1111 111     | 8                |
| 1.2                 | 1111         | 4                |
| 1.3                 | 111          | 3                |
| 1.4                 | 11           | 2                |
| 1.5                 | 1            | 1                |
|                     | Total        | 50               |

We now have a full numerical representation of the *frequency of events*. So, for example, eight people completed the assembly task in 1.1 hours or the time 1.1 hours has a frequency of 9. We will be using the above information later on, when we consider measures of central tendency.

Table 6.4.1

| <i>Height (cm)</i> | <i>Frequency</i> |
|--------------------|------------------|
| 150–154            | 4                |
| 155–159            | 9                |
| 160–164            | 15               |
| 165–169            | 21               |
| 170–174            | 32               |
| 175–179            | 45               |
| 180–184            | 41               |
| 185–189            | 22               |
| 190–194            | 9                |
| 195–199            | 2                |
| Total              | 200              |

The times in hours given in the above data are simply numbers. When data appears in a form where it can be *individually counted* we say that it is *discrete* data. It goes up or down in *countable* steps. Thus the numbers 1.2, 3.4, 8.6, 9, 11.1, 13.0 are said to be *discrete*. If, however, data is obtained by measurement, for example, the heights of a group of people, then we say that this data is *continuous*. When dealing with continuous data, we tend to quote its limits; that is, the limit of accuracy with which we take the measurements. So, for example, a person may be  $174 \pm 0.5$  cm in height. When dealing numerically with continuous data or a large amount of discrete data. It is often useful to *group* this data into *classes or categories*. We can then find out the numbers (frequency) of items within each group.

Table 6.4.1, shows the height of 200 adults, grouped into 10 classes.

The main advantage of grouping is that it produces a clear overall picture of the frequency distribution. In the table the first class interval is 150–154. The end number 150 is known as the *lower limit* of the class interval; the number 154 is the *upper limit*. The heights have been measured to the nearest centimetre. That means within  $\pm 0.5$  cm. Therefore, in effect, the first class interval includes all heights between the range

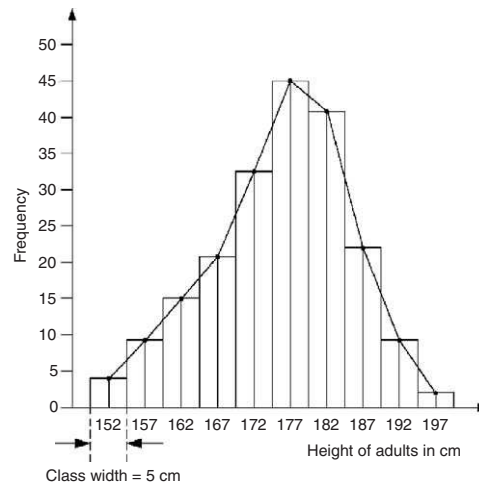


**Figure 6.4.6** Example 6.4.5, histogram showing frequency distribution

149.5–154.5 cm, these numbers are known as the lower and upper class *boundaries*, respectively. The **class width** is always taken as the **difference between the lower and upper class boundaries**, not the upper and lower limits of the class interval.

**The histogram** The *histogram* is a special diagram that is used to represent a frequency distribution, such as that for grouped heights, shown in Table 6.4.1. It consists of a set of rectangles, whose areas represent the frequencies of the various classes. Often when producing these diagrams, the class width is kept the same. So that the varying frequencies are represented by the height of each rectangle. When drawing histograms for grouped data, the *midpoints* of the rectangles represent the midpoints of the class intervals. So for our data, they will be 152, 157, 162, 167, etc.

An adaptation of the histogram, known as the *frequency polygon*, may also be used to represent a frequency distribution.

### Example 6.4.5

Represent the above data showing the frequency of the height of groups of adults on a histogram and draw in the frequency polygon for this distribution.

All that is required to produce the histogram is to plot frequency against the height intervals, where the intervals are drawn as class widths.

Then, as can be seen from Figure 6.4.6, the area of each part of the histogram is the *product of frequency  $\times$  class width*. The frequency polygon is drawn so that it connects the *midpoint* of the class widths.

Note that, the frequencies of a distribution may be added consecutively to produce a graph known as a **cumulative frequency distribution** or **ogive**. Cumulative frequency distributions and one or two of their related functions, will be studied later, when we consider probability distributions and their application to reliability engineering. In the mean time, an additional column may be added to the data shown in Table 6.4.1 and the cumulative frequency polygon can then be plotted.

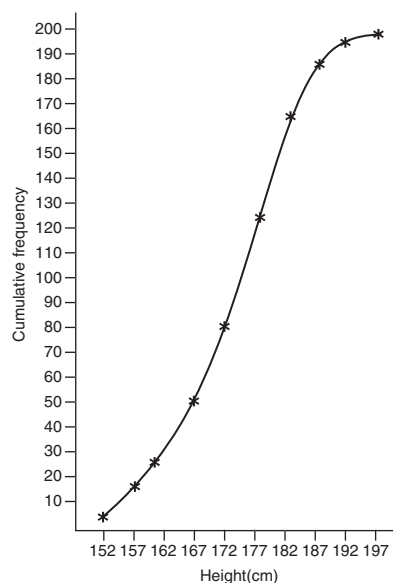


**Figure 6.4.7** Cumulative frequency distribution curve

### Example 6.4.6

Plot the cumulative frequency polygon for the data shown in Table 6.4.1.

All that is required is the addition of a column to the table as shown below, from which the associated cumulative frequency polygon is plotted (Figure 6.4.7).

| Height (cm) | Frequency | Cumulative frequency |
|-------------|-----------|----------------------|
| 150–154     | 4         | 4                    |
| 155–159     | 9         | 13                   |
| Height (cm) | Frequency | Cumulative frequency |
| 160–164     | 15        | 28                   |
| 165–169     | 21        | 49                   |
| 170–174     | 32        | 81                   |
| 175–179     | 45        | 126                  |
| 180–184     | 41        | 167                  |
| 185–189     | 22        | 189                  |
| 190–194     | 9         | 198                  |
| 195–199     | 2         | 200                  |
| Total       | 200       | 200                  |

### Test your knowledge 6.4.1

- (1) In a particular university, the number of students enrolled by faculty is given in the table below.

| Faculty                       | Number of students |
|-------------------------------|--------------------|
| Business and administration   | 1950               |
| Humanities and social science | 2820               |
| Physical and life sciences    | 1050               |
| Technology                    | 850                |
| Total                         | 6670               |

Illustrate this data on both a *bar chart* and *pie chart*.

- (2) For the group of numbers given below, produce a tally chart and determine their frequency of occurrence.

36 41 42 38 39 40 42 41 37 40  
 42 44 43 41 40 38 39 39 43 39  
 36 37 42 38 39 42 35 42 38 39  
 40 41 42 37 38 39 44 45 37 40

- (3) Given the following frequency distribution, produce a histogram and on it, draw the frequency polygon.

| Class interval | Frequency (f) |
|----------------|---------------|
| 60–64          | 4             |
| 65–69          | 11            |
| 70–74          | 18            |
| 75–79          | 16            |
| 80–84          | 7             |
| 85–90          | 4             |

## Statistical measurement of central tendency and dispersion

When considering statistical data it is often convenient to have one or two measured values which represent the data as a whole. Average values are often used. For example, we might talk about the average height of females in the UK being 170 cm, or that the average shoe size of British males is size 9. In statistics we may represent these average values using the mean, median, or mode of the data we are considering.

If we again consider the hypothetical data on the height of females. We may also wish to know how their individual heights vary or deviate from their average value. Thus, we need to consider measures of central tendency and dispersion, in particular, *mean deviation*, *standard deviation*, and *variance*, for the data concerned. These statistical measures of central tendency and the way in which they vary are considered next.

### Formulae

#### (1) Mean

The arithmetic mean or more simply the *mean* of the  $n$  values  $x_1, x_2, x_3, \dots, x_n$  is:

$$(i) \quad \bar{x} = \frac{1}{n} (x_1 + x_2 + x_3 + \dots + x_n) \quad \text{or} \quad \bar{x} = \frac{\sum x}{n}$$

If the  $n$  values have corresponding frequencies  $f_1, f_2, f_3, \dots, f_n$ , then the mean is:

$$(ii) \quad \frac{f_1x_1 + f_2x_2 + f_3x_3 + \dots + f_nx_n}{f_1 + f_2 + f_3 + \dots + f_n} \quad \text{or} \quad \bar{x} = \frac{\sum fx}{\sum f}$$

Best estimate of the arithmetic mean for *parent population* is given by:

$$(iii) \quad \mu \approx \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{or} \quad \mu \approx \bar{x} = \frac{\sum x}{n}$$

#### (2) Median

This is the middle or *median* value of a set of values placed in rank order. More formally the median may be defined as the 50th percentile of a data set.

If the number of items in a data set is *even*, then we add together the value of the *two middle terms* and divide by 2.

#### (3) Mode

The *mode* of a set of values containing discrete data is the value that occurs the most often.

When considering *frequency distributions* for grouped data, the *modal class* is that group which occurs most frequently.

#### (4) Variance and standard deviation for sample

The *variance* is the mean value of the square of the deviation, that is:

$$(i) \quad s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{or} \quad s^2 = \frac{\sum (x - \bar{x})^2}{n}$$

Also, if the  $n$  values have corresponding frequencies  $f_1, f_2, f_3, \dots, f_n$ , then the *variance* is:

$$(ii) \quad s^2 = \sum_{i=1}^n f_i (x_i - \bar{x})^2 \quad \text{or} \quad s^2 = \frac{\sum f(x - \bar{x})^2}{\sum f}.$$

The above formulae are for the variance of data obtained from a *random sample*, where  $s^2 = \text{variance}$  and  $\bar{x} = \text{mean value}$ , related to the random sample.

The measure of dispersion of the physical quantity the *standard deviation* is found by taking the square root of the variance.

(iii) *Standard deviation*:

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{or} \quad s = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}.$$

Also, if the  $n$  values have corresponding frequencies  $f_1, f_2, f_3, \dots, f_n$ , then:

(iv) *Standard deviation*:

$$s = \sqrt{\sum_{i=1}^n f_i (x_i - \bar{x})^2} \quad \text{or} \quad s = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}}.$$

(5) *Best estimate of variance and standard deviation (parent population)*

$$(i) \quad \text{Best estimate of variance} = \sigma^2 \approx \left( \frac{n}{n-1} \right) s^2.$$

$$(ii) \quad \text{Best estimate of standard deviation} \quad \sigma \approx \left( \sqrt{\frac{n}{n-1}} \right) s.$$

(6) *Error in the sample mean*

$$\sigma_m = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sum (x - \bar{x})^2}{n(n-1)}}.$$

## Methods

### The arithmetic mean

The arithmetic mean or simply the *mean* is probably the average with which you are most familiar. For example, to find the arithmetic mean of the numbers 8, 7, 9, 10, 5, 6, 12, 9, 6, 8. All we need do is to add them all up and divide by how many there are, or more formally:

$$\text{Arithmetic mean} = \frac{\text{Arithmetic total of all the individual values}}{\text{Number of values}} = \frac{\sum n}{n}.$$

where, the Greek symbol  $\Sigma$  = the sum of the individual values,  $x_1 + x_2 + x_3 + x_4 + \dots, x_n$  and  $n$  = the number of these values in the data.

So, for the *mean* of our 10 numbers, we have:

$$\text{Mean} = \frac{\sum n}{n} = \frac{8 + 7 + 9 + 10 + 5 + 6 + 12 + 9 + 6 + 8}{10} = \frac{80}{10} = 8.$$

Now, no matter how long or complex the data we are dealing with, *provided* that we are only dealing with individual values (discrete data), the above method will always produce the arithmetic mean. The mean of all the  $x$ -values is given the symbol  $\bar{x}$ , pronounced, *x-bar*.

**Example 6.4.7**

The height of 11 females were measured as follows: 165.6, 171.5, 159.4, 163, 167.5, 181.4, 172.5, 179.6, 162.3, 168.2, 157.3 cm. Find the mean height of these females. Then, for  $n = 12$ :

$$\bar{x} = \frac{165.6 + 171.5 + 159.4 + 163 + 167.5 + 181.4 + 172.5 + 179.6 + 162.3 + 168.2 + 157.3}{12}$$

$$\bar{x} = \frac{1848.3}{11} = \mathbf{168.03 \text{ cm.}}$$

**Mean for grouped data** What if we are required to find the mean for *grouped data*? Look back at Table 6.4.1 [page 520] showing the height of 200 adults, grouped into 10 classes. In this case, the *frequency* of the heights needs to be taken into account.

We select the *class midpoint*  $x$ , as being the average of that class and then multiply this value by the frequency ( $f$ ) of the class, so that a value for that particular class is obtained ( $fx$ ). Then, by adding up all class values in the frequency distribution, the total value for the distribution is obtained ( $\Sigma fx$ ). This total is then divided by the *sum of the frequencies* ( $\Sigma f$ ) in order to determine the mean. So, for grouped data:

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + f_3x_3 + \cdots + f_nx_n}{f_1 + f_2 + f_3 + \cdots + f_n} = \frac{\Sigma(f \times \text{midpoint})}{\Sigma f}.$$

This rather complicated looking procedure is best illustrated by example.

**Example 6.4.8**

Determine the mean value for the heights of the 200 adults, using the data in Table 6.4.1.

The values for each individual class are best found by producing a table, using the class *midpoints* and *frequencies*. Remembering that the *class midpoint* is found by dividing the *sum of the upper and lower class boundaries* by 2. So, for example, the mean value for the first class interval is  $(149.5 + 154.5)/2 = 152$ . The completed table is shown below.

| Midpoint (x) of height (cm) | Frequency (f)    | fx                    |
|-----------------------------|------------------|-----------------------|
| 152                         | 4                | 608                   |
| 157                         | 9                | 1413                  |
| 162                         | 15               | 2430                  |
| 167                         | 21               | 3507                  |
| 172                         | 32               | 5504                  |
| 177                         | 45               | 7965                  |
| 182                         | 41               | 7462                  |
| 187                         | 22               | 4114                  |
| 192                         | 9                | 1728                  |
| 197                         | 2                | 394                   |
| Total                       | $\Sigma f = 200$ | $\Sigma fx = 35\,125$ |



I hope you can see how each of the values was obtained. When dealing with relatively large numbers be careful with your arithmetic, especially when you are keying-in variables into your calculator!

Now that we have the required total the mean value of the distribution can be found.

$$\text{Mean value } \bar{x} = \frac{\sum fx}{\sum f} = \frac{35\,125}{200} = 175.625 \pm 0.5 \text{ cm.}$$

Notice that our mean value of heights has the same margin of error as the original measurements. The value of the mean cannot be any more accurate than the measured data from which it was found!

### Median

When some values within a set of data vary quite widely, the arithmetic mean gives a rather poor representative average of such data. Under these circumstances another more useful measure of the average is the *median*.

For example, the mean value of the numbers; 3, 2, 6, 5, 4, 93, 7 is 20, which is not representative of any of the numbers given. To find the median value of the same set of numbers, we simply place them in *rank order*; that is, 2, 3, 4, 5, 6, 7, 93. Then we select the middle (median) value. Since there are seven numbers (items) we choose the fourth item along, the number 5, as our *median value*.

If the number of items in the set of values is *even*, then we add together the value of the *two middle terms* and divide by 2.

### Example 6.4.9

Find the mean *and* median value for the set of numbers: 9, 7, 8, 7, 12, 70, 68, 6, 5, 8.

The arithmetic mean is found as:

$$\text{Mean } \bar{x} = \frac{9 + 7 + 8 + 7 + 12 + 70 + 68 + 6 + 5 + 8}{10} = \frac{200}{10} = 20.$$

This value is not really representative of any of the numbers in the set.

To find the *median* value, we first put the numbers in *rank order*, that is: 5, 6, 7, 7, 8, 8, 9, 12, 68, 70.

Then from the 10 numbers, the two middle values, the 5th and 6th values along are 8 and 8. So the

$$\text{median value} = \frac{(8 + 8)}{2} = 8.$$

### Mode

Yet another measure of central tendency for data containing extreme values is the *mode*. Now the *mode* of a set of values containing discrete data is the value that occurs most often. So for the set of values 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 7, 7, 7, the mode or *modal value* is 5, as this value occurs four

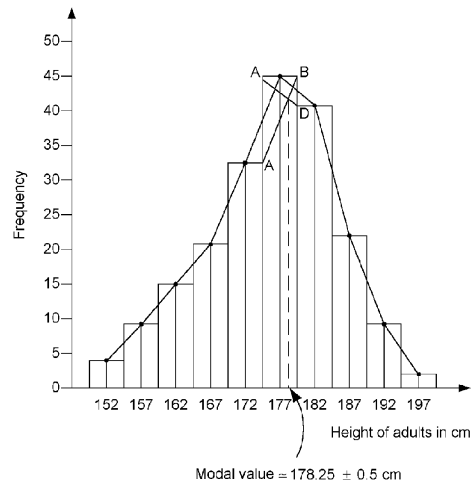


**Figure 6.4.8** Histogram for Example 6.4.10 showing modal value

times. Now it is possible for a set of data to have more than one mode, for example the data used in Example 6.4.8 above has two modes, 7 and 8, both of these numbers occurring twice and both occurring more than any of the others. A set of data may not have a modal value at all, for example the numbers: 2, 3, 4, 5, 6, 7, 8 all occur once and there is no mode.

A set of data that has one mode is called *unimodal*, data with two modes is *bimodal*, and data with more than two modes is known as *multimodal*.

When considering *frequency distributions* for grouped data. The *modal class*, is that group which occurs most frequently. If we wish to find the actual *modal value* of a frequency distribution, we need to draw a histogram.

### Example 6.4.10

Find the modal class and modal value for the frequency distribution on the height of adults given in Table 6.4.1 [on page 520].

Referring back to the table, it is easy to see that the class of heights, which occurs most frequently is 175–179 cm, which occurs 45 times.

Now, to find the modal value we need to produce a histogram for the data. We did this for Example 6.4.5. This histogram is shown again in Figure 6.4.8 with the modal value attached.

From Figure 6.4.8 it can be seen that the modal value =  $178.25 \pm 0.5$  cm.

This value is obtained from the intersection of the two construction lines, *AB* and *CD*. The line *AB* is drawn diagonally from the highest value of the preceding class, up to the top right-hand corner of the modal class. The line *CD* is drawn from the top left-hand corner of the modal group to the lowest value of the next class, immediately above the modal group. Then, as can be seen, the *modal value* is read-off where the projection line meets the *x-axis*.

### Mean deviation

We talked earlier of the need not only to consider statistical averages, which give us some idea of the position of a distribution, but also the

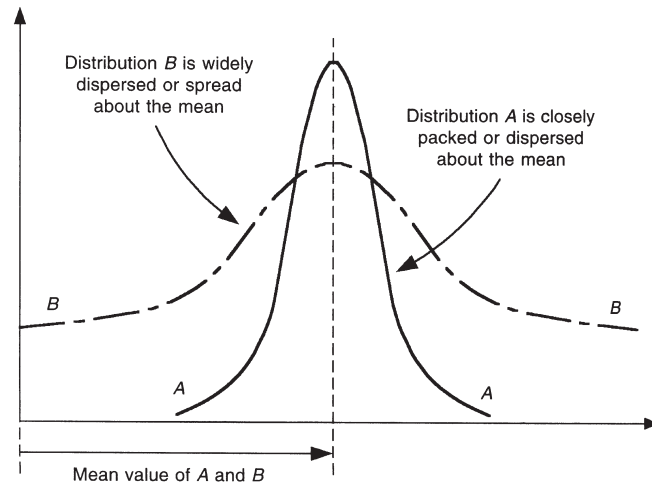


**Figure 6.4.9** Deviation from the mean value, for a distribution

need to consider how the data is *dispersed* or *spread* about this average value. Figure 6.4.9 illustrates this idea, showing how the data taken from two distributions is dispersed about the same mean value.

A measure of dispersion which is often used is the *mean deviation*. To determine the deviation from the statistical average (mean, median, or mode), we proceed in the following way.

We first find the statistical average for the distribution, the mean, median, or mode ( $\bar{x}$ ). We then find the difference between this average value and each of the individual values in the distribution. We then add up all these difference and divide by the number of individual values in the distribution. This all sounds rather complicated, but the mean deviation may be calculated quite easily using the formula:

$$\text{Mean deviation} = \frac{\sum |x - \bar{x}|}{n}$$

where  $x$  = a data value in the distribution,  $\bar{x}$  = the statistical average, mean, median or mode, as before, and  $n$  = the number of individual items in the distribution as before. The  $| |$  brackets tell us to use the *positive* value of the result contained within the brackets. So, for example, if  $x = 12$  and  $\bar{x} = 16$ , then  $|x - \bar{x}| = |12 - 16| = |-4| = +4$ , we use the *positive* value, even though, in this case, the result was negative.

For *frequency distributions*, using grouped data, we find the deviation from the mean, using a similar formula to that we used to find the arithmetic mean. Where the only addition is to multiply the individual differences from the mean by their frequency (Figure 6.4.9). Then, for a frequency distribution:

$$\text{Mean deviation} = \frac{\sum f|x - \bar{x}|}{\sum f}$$

### Example 6.4.11

Calculate the mean deviation from the arithmetic mean for the data shown in the table.

|                      |     |     |      |      |       |      |      |
|----------------------|-----|-----|------|------|-------|------|------|
| Length of rivet (mm) | 9.8 | 9.9 | 9.95 | 10.0 | 10.05 | 10.1 | 10.2 |
| Frequency            | 3   | 18  | 36   | 62   | 56    | 20   | 5    |

The easiest way to tackle this problem is to set up a *table of values* in a similar manner to the table we produced for Example 6.4.8. The headings for such a table being taken from the above formula for finding the mean deviation for a frequency distribution.

*Table of values*

| Rivet length (x) | f                | fx                   | $ x - \bar{x} $ | $f x - \bar{x} $                |
|------------------|------------------|----------------------|-----------------|---------------------------------|
| 9.8              | 3                | 29.4                 | 0.202           | 0.606                           |
| 9.9              | 18               | 178.2                | 0.108           | 1.944                           |
| 9.95             | 36               | 358.2                | 0.058           | 2.088                           |
| 10               | 62               | 620                  | 0.008           | 0.496                           |
| 10.05            | 56               | 562.8                | 0.03            | 1.68                            |
| 10.1             | 20               | 202                  | 0.092           | 1.84                            |
| 10.2             | 5                | 51                   | 0.192           | 0.92                            |
| Total            | $\Sigma f = 200$ | $\Sigma fx = 2001.6$ |                 | $\Sigma f x - \bar{x}  = 9.574$ |

The arithmetic mean  $\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{2001.6}{200} = 10.008$ .

$\bar{x} = 10.008$  was required to complete the last two columns in the table.

Then the *mean deviation* from the mean of the rivet lengths is:

$$= \frac{\Sigma f|x - \bar{x}|}{\Sigma f} = \frac{9.574}{200} = 0.04787 \text{ mm} \approx \mathbf{0.05 \text{ mm}}.$$

This small average deviation from the arithmetic mean for rivet length is what we would expect in this case. The deviation being due to very small manufacturing errors. This is, therefore, an example of a frequency distribution tightly packed around the average for the distribution.

### Standard deviation

The most important method in determining how a distribution is *dispersed* or *spread* around its average value is known as *standard deviation*. To find this measure of dispersion requires just one or two additional steps from those we used to find the mean deviation.

These additional mathematical steps involve further manipulation of the  $|x - \bar{x}|$  or  $f|x - \bar{x}|$  values, we needed to find when calculating the mean deviation for discrete or grouped data. The additional steps require us to first *square* these differences, then find *their* mean and finally take their *square root* to reverse the squaring process. This strange way of manipulating these differences is known as the *root mean square deviation* or *standard deviation*, which is identified using the Greek symbol sigma ( $\sigma$ ).

Thus, for frequency distributions with grouped data, we can represent these *three further processes* mathematically, as follows:

- (1) Square the differences and multiply by their frequency  $= f|x - \bar{x}|^2$

- (2) Sum all of these values and find their mean  $= \frac{\sum f|x - \bar{x}|^2}{\sum f}$ , this is a similar step to the way in which we found the mean deviation. The value of the deviation found at this stage is known as the *variance*.

- (3) Now take the square root of these mean squares to reverse the squaring

$$\text{process} = \sqrt{\frac{\sum f|x - \bar{x}|^2}{\sum f}}.$$

$$\text{Then the standard deviation } \sigma = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}}.$$

The  $||$  brackets have been replaced by ordinary brackets in this final version of the formula. This is because, when we square any quantity, whether positive or negative, the result is always positive, by the law of signs! It is therefore no longer necessary to use the special brackets.

This particular value of deviation is more representative than the mean deviation value, we found before, because it takes account of data that may have large differences between items. In a similar way to the use of the mode and median, when finding average values.

When considering discrete ungrouped data, we apply the same steps as above to the differences  $|x - \bar{x}|$  and obtain,  $\sqrt{\sum(x - \bar{x})^2/n}$ , therefore for *ungrouped* data:

$$\text{The standard deviation } \sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}.$$

Note that once again we have removed the special brackets, for the same reason as given above, for grouped data.

### Example 6.4.12

For the set of numbers 8, 12, 11, 9, 16, 14, 12, 13, 10, 9, find the arithmetic mean and the standard deviation.

Like most of the examples concerning central tendency and deviation measure, we will solve this problem by setting up a table of values. We will also need to find the arithmetic mean before we are able to complete the table, where in this case for non-grouped data  $n = 10$ .

$$\text{Then, } \bar{x} = \frac{\sum x}{n} = \frac{8 + 12 + 11 + 9 + 16 + 14 + 12 + 13 + 10 + 9}{10} = \frac{114}{10} = 11.4.$$

Table of values

| $x$ | $(x - \bar{x})$ | $(x - \bar{x})^2$ |
|-----|-----------------|-------------------|
| 8   | -3.4            | 11.56             |
| 12  | 0.6             | 0.36              |
| 11  | -0.4            | 0.16              |
| 9   | -2.4            | 5.76              |

| $x$              | $(x - \bar{x})$ | $(x - \bar{x})^2$               |
|------------------|-----------------|---------------------------------|
| 16               | 4.6             | 21.16                           |
| 14               | 2.6             | 6.76                            |
| 12               | 0.6             | 0.36                            |
| 13               | 1.6             | 2.56                            |
| 10               | -1.4            | 1.96                            |
| 9                | -2.4            | 5.76                            |
| $\Sigma x = 114$ |                 | $\Sigma (x - \bar{x})^2 = 56.4$ |

Then from the table of values:

$$\text{The standard deviation } \sigma = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n}} = \sqrt{\frac{56.4}{10}} = \sqrt{5.64} = 2.375.$$

Another measure of dispersion is the *variance* which is simply the value of the standard deviation before taking the square root, so in this example:

$$\text{The variance} = \frac{\Sigma (x - \bar{x})^2}{n} = \frac{56.4}{10} = 5.64.$$

So when finding the standard deviation, you can also find the variance.

Finally, make sure you can obtain the values given in the table!

Here is an example involving grouped data.

### Example 6.4.13

Calculate the standard deviation for the data on rivets given in Example 6.4.11.

For convenience the data from Example 6.4.11 is reproduced here.

|                      |     |     |      |      |       |      |      |
|----------------------|-----|-----|------|------|-------|------|------|
| Length of rivet (mm) | 9.8 | 9.9 | 9.95 | 10.0 | 10.05 | 10.1 | 10.2 |
| Frequency            | 3   | 18  | 36   | 62   | 56    | 20   | 5    |

Now in Example 6.4.11 we calculated the arithmetic mean and *mean deviation*, using a table of values, we obtained:

| Rivet length ( $x$ ) | $f$              | $fx$                 | $ x - \bar{x} $ | $f x - \bar{x} $                |
|----------------------|------------------|----------------------|-----------------|---------------------------------|
| 9.8                  | 3                | 29.4                 | 0.202           | 0.606                           |
| 9.9                  | 18               | 178.2                | 0.108           | 1.944                           |
| 9.95                 | 36               | 358.2                | 0.058           | 2.088                           |
| 10                   | 62               | 620                  | 0.008           | 0.496                           |
| 10.05                | 56               | 562.8                | 0.03            | 1.68                            |
| 10.1                 | 20               | 202                  | 0.092           | 1.84                            |
| 10.2                 | 5                | 51                   | 0.192           | 0.92                            |
| Total                | $\Sigma f = 200$ | $\Sigma fx = 2001.6$ |                 | $\Sigma f x - \bar{x}  = 9.574$ |

The arithmetic mean we found as,

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{2001.6}{200} = 10.008.$$

So, having found the mean, all we need do now, to find the *standard deviation* is modify the table by adding-in the extra steps. We then obtain:

| Rivet length ( $x$ ) | $f$ | $fx$   | $(x - \bar{x})$ | $(x - \bar{x})^2$ | $f(x - \bar{x})^2$ |
|----------------------|-----|--------|-----------------|-------------------|--------------------|
| 9.8                  | 3   | 29.4   | -0.202          | 0.040804          | 0.122412           |
| 9.9                  | 18  | 178.2  | -0.108          | 0.011664          | 0.209952           |
| 9.95                 | 36  | 358.2  | -0.058          | 0.003364          | 0.121104           |
| 10                   | 62  | 620    | -0.008          | 0.000064          | 0.003968           |
| 10.05                | 56  | 562.8  | 0.03            | 0.0009            | 0.0504             |
| 10.1                 | 20  | 202    | 0.092           | 0.008464          | 0.16928            |
| 10.2                 | 5   | 51     | 0.192           | 0.036864          | 0.18432            |
| Total                | 200 | 2001.6 |                 |                   | 0.861436           |

Then from the table  $\sum f = 200$ ,  $\sum f(x - \bar{x})^2 = 0.861436$  and

$$\text{Standard deviation } \sigma = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}} = \sqrt{\frac{0.861436}{200}} = \mathbf{0.066 \text{ mm.}}$$

This value is slightly more accurate than the value we found in Example 6.4.10 for the mean deviation = 0.05 mm, but as you can see, there is also a lot more arithmetic manipulation! Again, you should make sure that you are able to obtain the additional values shown in the table.

### **Mean value and variance in a random sample and parent population**

You will appreciate the need to sample, for engineering quality control purposes. It would simply be too prohibitively expensive to sample all the parent population. For example, in the batch production of say 10 000 close tolerance bolts, to check each one for dimensional accuracy is ludicrous, so we need to consider a random sample. We will be looking at this process in more detail when we consider the application of statistics to quality control. In the mean time let us define what we mean by sample and parent population.

A *random sample* contains  $n$  measurements of one quantity. We consider the measurements as a random selection out of the set of all possible measurements. These measurements could be from an experiment or from a batch product, such as the bolts mentioned earlier. All possible measurements represent the *parent population*, which is always larger than the sample. The parent population like the random sample is characterised by a mean value and a variance. The mean value for the parent population is the hypothetical *true* value. The values which relate to the parent population can be *estimated* from the random sample data. The larger the random sample is, then the more reliable the estimate becomes.

If you look back at formulae (1), (4), and (5) at the beginning of this section you will see that quantities that relate to the *random sample* are represented by Latin letters, that is:

$\bar{x}$  = mean value and  $s^2$  = variance.

and those which relate to the *parent population* are denoted by Greek symbols, that is:

$\mu$  = mean value and  $\sigma^2$  = variance.

The estimates for the true values for the parent population are really the ones we are interested in.

### Example 6.4.14

The length of a number of bolts has been measured a number of times, as shown in the table set out below. Calculate the mean value, variance, and standard deviation for the *parent* population.

| Length $x_i$<br>(cm)                                   |
|--------------------------------------------------------|
| 14.1 13.8 14.3 14.2 14.5 14.1 14.2 14.4 14.3 13.9 14.4 |

Then we first determine the mean, standard deviation, and variance of the sample using the formulae:

$$\bar{x} = \frac{\sum x}{n} \quad \text{and} \quad s = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}.$$

The mean value of the sample from above formula is:

$$\bar{x} = \frac{\sum x}{n} = \frac{156.2}{11} = 14.2 \text{ cm.}$$

and to determine the variance and standard deviation of the sample, we use a tabular method, as shown below.

| Length $x_i$ (cm) | $(x_i - \bar{x})$ (cm) | $(x_i - \bar{x})^2$ (cm) <sup>2</sup> |
|-------------------|------------------------|---------------------------------------|
| 14.1              | −0.1                   | 0.01                                  |
| 13.8              | −0.4                   | 0.16                                  |
| 14.3              | 0.1                    | 0.01                                  |
| 14.2              | 0                      | 0                                     |
| 14.5              | 0.3                    | 0.09                                  |
| 14.1              | −0.1                   | 0.01                                  |
| 14.2              | 0                      | 0                                     |
| 14.4              | 0.2                    | 0.04                                  |
| 14.3              | 0.1                    | 0.01                                  |
| 13.9              | −0.3                   | 0.09                                  |
| 14.4              | 0.2                    | 0.04                                  |
| Sum 156.2         | 0                      | 0.46                                  |

From which:

$$s = \sqrt{\frac{0.46}{11}} = 0.205 \text{ cm} \quad \text{where} \quad s^2 = 0.0418.$$



Now from formulae 1(iii), 5(i), and 5(ii) the mean value, variance, and standard deviation for the *parent population* are as follows:

**Mean**  $\mu \approx \bar{x} = \frac{\sum x}{n}$ , that is  $\mu = \mathbf{0.142 \text{ cm}}$ , as for sample

**Variance**  $\sigma^2 \approx s^2 = \frac{n}{n-1} = (0.0418) \frac{11}{10} = \mathbf{0.046 \text{ cm}}$  and

**Standard deviation** =  $\mathbf{0.215 \text{ cm}}$ .

**Error in the sample mean** We have so far, considered the mean value of the sample to be the best estimate for the mean value of the parent population. What we have not considered is how much error there may be in our sample mean when compared to the *true* mean value of the parent population.

In order to find our best estimate, we would need to take several random samples from the parent population and find the *mean* of *each* set of these samples. The mean values of these samples will be scattered around the *true* value for the parent population. Then what we need to know is whether the dispersion of the mean values is smaller than that of the individual values. This dispersion of the mean values is important, because it helps us to determine the reliability of the result of a series of readings (data).

So, if the variance of the mean values of the random samples is represented as  $\sigma_m^2$  and the variance of the individual values of the parent population is represented as before, that is as  $\sigma^2$ .

Then the variance of the mean values is given by  $\sigma_m^2 = \frac{\sigma^2}{n}$  so the

standard deviation of the mean values of the sample is  $\sigma_m = \frac{\sigma}{\sqrt{n}}$ .

This standard deviation of the mean values is a measure of the accuracy of the mean value of a sample of measurements or readings. It is referred to as the *sample error*.

Thus the standard deviation of the mean value (the sampling error) is given by formula (6), that is:

$$\sigma_m = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sum (x - \bar{x})^2}{n(n-1)}}.$$

It should be fairly obvious that the accuracy of the mean value of the sample can be improved by increasing the number of individual measurements.

### Example 6.4.15

Consider the arithmetic mean we found for the length of the bolts in Example 6.4.15.

Arithmetic mean  $\bar{x} = 0.142 \text{ cm}$ . What are the error limits of this mean value of length?

All we need do is find the mean error of this mean value, in other words find the standard deviation of this mean value.

This is given by formula (6), as:

$$\sigma_m = \frac{\sigma}{\sqrt{n}} = \frac{0.205}{\sqrt{11}} = 0.0618 \text{ cm}.$$

Therefore the error in our mean value is:

$$\bar{x} = 0.142 \pm 0.0618 \text{ cm}, \text{ quite a large error!}$$

From what has been said earlier, if we wish to reduce this error in our mean value by say a half, we would need to take 44 readings.

As mentioned at the beginning of this section, the application of the above methods will be considered later, after we have studied some further statistical methods and probability theory. So we end this topic, *for the moment*, with a few simple test your knowledge questions.

### Test your knowledge 6.4.2

- (1) Calculate the mean of the numbers 176.5, 98.6, 112.4, 189.8, 95.9, and 88.8.
- (2) Determine the mean, median, and mode for the set of numbers, 9, 8, 7, 27, 16, 3, 1, 9, 4, and 116.
- (3) Estimates for the length of wood required for a shelf were as follows:

|             |    |    |    |    |    |    |    |    |
|-------------|----|----|----|----|----|----|----|----|
| Length (cm) | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 |
| Frequency   | 1  | 3  | 4  | 8  | 6  | 5  | 3  | 2  |

Calculate the arithmetic mean and mean deviation.

- (4) Calculate the arithmetic mean and the mean deviation for the data shown in the table.

|             |     |     |     |     |     |
|-------------|-----|-----|-----|-----|-----|
| Length (mm) | 167 | 168 | 169 | 170 | 171 |
| Frequency   | 2   | 7   | 20  | 8   | 3   |

- (5) Calculate the standard deviation from the median value, for the numbers given in question 2.
- (6) Tests were carried out on 50 occasions, to determine the percentage of greenhouse gases in the emissions from an internal combustion engine. The results from the tests showing the percentage of greenhouse gas recorded were as follows:

|                          |     |     |     |     |     |     |
|--------------------------|-----|-----|-----|-----|-----|-----|
| % greenhouse gas present | 3.2 | 3.3 | 3.4 | 3.5 | 3.6 | 3.7 |
| Frequency                | 2   | 12  | 20  | 8   | 6   | 2   |

Determine the arithmetic mean and the standard deviation for the greenhouse gases present.

- (7) Nine different samples of an aluminium alloy were taken and the percent of copper by volume was determined, with the following results:

3.6, 3.3, 3.2, 3.0, 3.2, 3.1, 3.0, 3.1, 3.3

Determine:

- (i) the mean value and standard deviation of the parent population,
- (ii) the error of the mean.

## Statistical probability and probability distributions

The concepts and methods of probability have been used increasingly for many years in both science and engineering. These concepts are the basis for an understanding of a large part of theoretical physics, in particular with respect to statistical and quantum mechanics.

Using the statistical approach the behaviour of solids, liquids, and gases can be modelled, by considering what happens at the microscopic (atomic) level the macroscopic behaviour of the parent material can be determined. In the study of quantum mechanics the behaviour of atomic and sub-atomic particles can be studied. In this respect probability theory plays a significant part because only probability statements can be made about the properties and behaviour of sub-atomic particles.

Probability theory is also used in the modelling of physical systems that contain a large number of elements. In other words those that are too complex to be modelled using conventional mathematics, for example a cities traffic system or the probability of failure of a complex instrument system on an aircraft. The *reliability* of such systems is primarily determined using statistical and probability theory.

A further field of application is in the theory of errors, which you have looked at briefly in the previous section, where we considered the error in our estimates of mean values. All physical measurements are in principle liable to errors and we will look again at this topic when we consider the analysis of experimental data and the confidence we can have in the results.

One final important application is in quality control in industrial production of engineering artefacts. Manufacturers need the assurance of quality control which statistics can provide.

### Formulae

- (1) *Rules of probability*

$$\begin{aligned}
 \text{(i) } P(A) &= \frac{a}{N} \\
 &= \frac{\text{Number of elementary events contained in event } a}{\text{Total number of elementary events}}
 \end{aligned}$$

- (ii) Success or failure:  $p$  = success,  $q$  = failure  
 or  $P(A)$  probability of event occurring  
 $P(\bar{A})$  probability of event *not* occurring.

- (iii) Addition rule (mutually exclusive events):

$$\begin{aligned} P(A \text{ or } B) &= P(A) + P(B), \\ \text{or } P(A \cup B) &= P(A) + P(B), \\ \text{or } P(A + B) &= P(A) + P(B). \end{aligned}$$

- (iv) Addition rule (*not* mutually exclusive events):

$$\begin{aligned} P(A \text{ or } B) &= P(A) + P(B) - P(AB), \\ \text{or } P(A \cup B) &= P(A) + P(B) - P(A \text{ and } B), \\ \text{or } P(A + B) &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

- (v) Independent events

$$P(AB) = P(A)P(B) \quad \text{or} \quad P(A \cap B) = P(A)P(B)$$

and for  $n$  independent events  $A_1, A_2, \dots, A_n$

$$\begin{aligned} P(A_1, A_2, \dots, A_n) &= P(A_1)P(A_2) \dots P(A_n) \\ \text{or } P(A_1 \cap A_2 \cap \dots \cap A_n) &= P(A_1)P(A_2) \dots P(A_n). \end{aligned}$$

- (vi) Conditional probabilities

$$\begin{aligned} P(AB) &= P(A)P(B/A) = P(B)P(A/B) \\ \text{or } P(A \cap B) &= P(A)P(B/A) = P(B)P(A/B). \end{aligned}$$

- (vii) Posteriori probability – Bayes' theorem:

Let  $A_1, A_2, \dots, A_n$  be a mutually exclusive and exhaustive set of outcomes of a random process, and  $B$  be the chance event such that  $P(B) \neq 0$ , then

$$P(A_r/B) = \frac{P(A_r)P(B/A_r)}{\sum_{r=1}^n P(A_r)P(B/A_r)}.$$

- (2) *Normal distribution*

- (i) The equation of the normal distribution curve is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2/\sigma^2} \quad \text{or} \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}.$$

- (ii) The area under the graph to left of a given value of  $x$  is:

$$A(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}z^2} dz.$$

- (iii) Standardised variable  $z = \frac{x - \mu}{\sigma}$ .

- (3) *Binomial distribution*

If the probability of success in one trial is  $p$ , then the probability of  $r$  successes out of  $n$  trials is:

$$\frac{n!}{r!(n-r)!} p^r (q)^{n-r} \quad \text{where } q = (1 - p).$$

The terms of the general binomial expansion are given by:

$$(q + p)^n = q^n + \binom{n}{1} q^{n-1} p + \binom{n}{2} q^{n-2} p^2 + \dots + p^n$$

Mean of the binomial is  $np$  and variance of the binomial is  $npq$ .

(4) *Poisson distribution*

The probability of 0, 1, 2, 3, ...  $n$  successes are:

$$e^{-a}, ae^{-a}, \frac{a^2}{2!} e^{-a}, \dots, \frac{a^n}{n!} e^{-a} \text{ where } a = np, \text{ when } n \text{ is large and}$$

$p$  is small.

The mean of the Poisson distribution = the variance =  $a$ .

## Methods

### *The laws of probability*

We first consider the use of the laws of probability, given above. The explanation of these laws is particularly wordy therefore, wherever possible, we will try to illustrate them by example. Unfortunately the examples needed to demonstrate the unambiguous use of these laws need to be rather trivial. For this reason we will adopt the conventional approach, by considering examples that involve the games of chance, such as the throwing of die, or the probabilities involved when using playing cards.

After you have mastered simple probability and the nature and use of probability distributions, you will be in a position to attempt the engineering applications given in the final section.

#### **Example 6.4.7**

What is the probability that an odd number will appear when a die is thrown?

Here all we need to do is to use formula 1(i), that is:

$$P(A) = \frac{a}{N} = \frac{\text{Number of elementary events contained in event } a}{\text{Total number of elementary events}}$$

There are six possible ways a die can drop; so  $N = 6$ . Out of these six ways, there are three that will show an odd number (1, 3, 5) and three that will show an even number (2, 4, 6), therefore  $a = 3$  and the probability of throwing an odd number is  $\frac{a}{N} = \frac{3}{6} = \frac{1}{2}$ , so the probability of throwing an

$$\text{odd number } P(\text{odd}) = \frac{1}{2}.$$

Applying similar logic we can determine the probability of selecting an ace from a pack of cards. Assume first we are using an ordinary pack of playing cards without jokers, then the total number of cards involved is 52.

Now there are four aces in the pack, therefore we have four possibilities to realise the event of 'selecting one ace' from the pack. Therefore:

$$P(\text{ace}) = \frac{a}{N} = \frac{4}{52} = \frac{1}{13}.$$

Suppose now that in order to win, you are required to throw a six or a two, with a die. Then the probability of you winning is:

$$P(A \text{ or } B) = P(A) + P(B) \quad \text{or} \quad P(6) + P(2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

This is an example of the use of the *addition rule* for *mutually exclusive* events. What we mean by mutually exclusive is that the outcome of the individual events are independent, providing of course we are using a *fair* die!

### Example 6.4.8

What is the probability of (i) drawing a heart or an ace and (ii) drawing a heart or an ace or both.

For (i) the events are mutually exclusive so:

$$P(S \text{ or } A) = P(S) + P(A) = \frac{13}{52} + \frac{4}{52} = \frac{17}{52}.$$

However for situation (ii) the probability of drawing a heart and an ace, i.e. both is conditional on success with the first draw. So these events are *not* mutually exclusive, therefore we need to use rule 1(iv). Then,

$$\begin{aligned} P(H \cup A) &= P(H) + P(A) - P(H \cap A) \\ &= \frac{13}{52} + \frac{4}{52} - \left( \frac{13}{52} \right) \left( \frac{4}{52} \right) \\ &= \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52}. \end{aligned}$$

Note the use of  $\cup$  (read as *or*) and  $\cap$  (read as *and*).

In the next example we consider again, a combination of independent events and events that are conditional.

### Example 6.4.9

If three dice are thrown simultaneously, what is the probability of obtaining:

- (a) 3 sixes
- (b) 1 and only one six
- (c) at least 1 six.

(a) The probability of 3 sixes is given by the multiplication rule for independent events (rule 1(v)). The probability of throwing a six from any one die is  $\frac{1}{6}$ , therefore for three dice it is  $\left(\frac{1}{6}\right)^3 = \frac{1}{216}$ .

(b) 1 and only six. The logic here is that the first die must be a six, with the second and third not or, the second die must be a six with the first and third not or, the third die must be a six with the first and second not. Thus the probability of throwing 1 and only one six is:

$$3\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{25}{72}.$$

(c) The probability of obtaining at least 1 six, is best solved remembering the rule that at least 1 = 1 – the probability of obtaining no sixes.

$$\text{So at least 1 six} = 1 - \left(\frac{5}{6}\right)^3 = \frac{91}{216}.$$

This last variant of the rules is particularly useful, when trying to determine the probability of 'at least 1, 2, etc'.

The logic to Example 6.4.9(c), may be applied to many situations. For example, if 5% of the transistors manufactured by a company are defective. What is the probability of getting at least one defective transistor in a batch of four?

Since 5% are considered to be defective, then we may assume that any transistor selected at random has a 5% chance of being defective or a 95% chance of being serviceable. So the probability of all four of our sample being sound is  $(0.95)^4 = 0.8145$ . So applying the logic of the last example, then the probability of at least one transistor being defective is  $1 - 0.8145 = 0.185$  or approximately  $\frac{2}{11}$ .

We finish illustrating the laws of probability, by considering *posteriori probabilities*. So far we have with all previous laws only considered the probability of an event that *will* occur, that is those that are calculated prior to observing the results. These are known as *priori* or *prior probabilities*. What about the probability that is calculated after the outcome of an experiment has been observed? Well, these are known as *posteriori probabilities*. For the later type of probability we may use *Bayes theorem* to solve them.

Consider for example two machining shops (A and B), within a factory. Where machine shop A produces 45% of total production and machine shop B produces the remaining 55%. In addition suppose that from machine shop A there are 12 defective components out of every thousand and from workshop B there are six defective items out of every thousand. After a thorough mix of the components from both workshops, a single component is drawn at random and found to be defective. What is the probability that it came from workshop A?

Then the probability that the component comes from A is  $P(A) = 0.45$  and the probability that if it comes from A, that it is defective is  $P(D/A) = 0.012$ .

So the probability that it comes from A and is defective, is

$$P(A \cap D) = P(A)P(D/A) = (0.45)(0.012) = 0.0054.$$

By a similar argument, the probability that the component came from B and was defective, is:

$$P(B \cap D) = P(B)P(D/B) = (0.55)(0.006) = 0.0033.$$

The ratio of the above probabilities is 54:33. Therefore the probability that it came from A, is given by:

$$P(A/D) = \frac{54}{54 + 33} = \frac{54}{87}.$$

The above process can be generalised so that:

$$P(A/D) = \frac{P(D/A)P(A)}{P(D/A)P(A) + P(D/B)P(B)} \quad \text{and similarly}$$

$$P(B/D) = \frac{P(D/B)P(B)}{P(D/A)P(A) + P(D/B)P(B)}.$$

This law can then be extended for any number of workshops and their defectives. So for example if there were four workshops and we needed the probability of drawing a defective item from workshop C, we would find it from the relationship

$$P(C/D) = \frac{P(D/C)P(C)}{P(D/A)P(A) + P(D/B)P(B) + P(D/C)P(C) + P(D/D)P(D)}.$$

So finally, for  $n$  outcomes of a random process  $A_1, A_2, \dots, A_n$  be the mutually exhaustive set of outcomes of the random process, and  $C$  be a chance event with probability  $P(C) \neq 0$ , then we may find such an event using

$$\text{Bayes Theorem, } P(A_r/B) = \frac{P(A_r)P(B/A_r)}{\sum_{r=1}^n P(A_r)P(B/A_r)}.$$

### Example 6.4.10

In a factory manufacturing engine components, three machines produce cylinder heads. Machine A produces 40% of the total, machine B 25% and machine C the remainder. On average 6% of the cylinder heads produced by A do not conform to tolerance requirements. The corresponding percentages for B and C are 4 and 3, respectively. One head selected at random for sampling from the whole output, is found to be out of tolerance. Calculate the probability that it was produced by machine A.

The solution to this problem is very straight forward, all that is required is to use Bayes theorem for the probability that the defective will come from machine A, rather than from machine B and C. So we solve using:

$$P(A/D) = \frac{P(D/A)P(A)}{P(D/A)P(A) + P(D/B)P(B) + P(D/C)P(C)}.$$

Therefore we consider the three probabilities that it comes from A, B, or C and it is defective.



So in full the probability that the component comes from  $A = 0.4$ . If it comes from  $A$ , the probability that it is defective  $= 0.06$ . Thus the probability that it comes from  $A$  and is defective i.e.  $P(A)P(D/A) = (0.4)(0.06) = 0.024$ .

Similarly  $P(B)P(B/D) = (0.25)(0.04) = 0.01$  and

$$P(C)P(C/D) = (0.35)(0.03) = 0.0105.$$

So on substitution of these values:

$$P(A/D) = \frac{0.024}{0.024 + 0.01 + 0.0105} = \frac{0.024}{0.0445} = \frac{24}{44.5} = \frac{48}{89}.$$

### Probability distributions

Before we consider the important probability distributions, it is worth revising the concept of permutations and combinations. If you look back at page 448, you will see our simple definitions, for permutations and combinations.

In other words: A *permutation* is the selection of  $r$  objects from  $n$  objects *in a specified order*, in symbols we represent a permutation as:

$${}^n P_r = \frac{n!}{(n-r)!}$$

Thus three objects selected from six objects, irrespective of order is:

$${}^6 P_3 = \frac{6!}{(6-3)!} = \frac{6!}{3!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1} = 120.$$

Remembering, factorial notation, from your earlier study of the binomial theorem.

Permutations, although useful, are not as useful as combinations, because it is *combinations* that are used to determine the *coefficients in the Binomial theorem*.

We define combinations as the selection of  $r$  objects from  $n$  objects, irrespective of order, or in symbols:

$${}^n C_r = \frac{n!}{r!(n-r)!} = \binom{n}{r}.$$

Thus if four transistors are selected at random from a box of 10. Then there are:

$$\begin{aligned} {}^n C_r &= \frac{10!}{4!(10-6)!} = \binom{10}{4} \\ &= \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{(4 \times 3 \times 2 \times 1)(6 \times 5 \times 4 \times 3 \times 2 \times 1)} \\ &= 210 \text{ ways.} \end{aligned}$$

Note also the short cut way of writing this combination, that is

$$\binom{10}{4} = \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2 \times 1} = 210.$$

The above is all very straight forward, difficulties usually occur when it comes to using the above formula for problems related to probability. The techniques involved are illustrated in the next couple of examples.

### Example 6.4.11

Two cards are selected at random from a pack of 52 playing cards. What is the probability that (i) they are both kings, (ii) neither are kings?

(i) The probability that they are both kings, requires us to compare the number of ways of selecting two kings from the four in the pack, with the total number of ways of selecting any two cards from the 52 in the pack. In other words:

$$P_{2k} = \frac{\begin{array}{c} \text{the number of ways of selecting two kings} \\ \text{from the four in the pack} \end{array}}{\begin{array}{c} \text{the number of ways of selecting any two} \\ \text{cards from the 52 in the pack} \end{array}}$$

$$P_{2k} = \frac{\binom{4}{2}}{\binom{52}{2}} = \frac{\frac{4 \times 3}{1 \times 2}}{\frac{52 \times 51}{1 \times 2}} = \frac{6}{1326} = \frac{1}{221}.$$

(ii) In this case we are interested in establishing the number of ways of selecting two cards that are not kings! That means we have only 48 cards in the pack from which to make our

choice, i.e.  $\binom{48}{2}$ . This probability must be compared with the total number of ways of selecting any two cards from the 52 in the pack, as we did in part (i). Therefore:

$$P_{no\ kings} = \frac{\binom{48}{2}}{\binom{52}{2}} = \frac{\frac{48 \times 47}{2 \times 1}}{\frac{52 \times 51}{2 \times 1}} = \frac{1128}{1326} = \frac{188}{221}.$$

### Example 6.4.12

A jobbing factory has 10 machines which may require adjustment, as necessary. Three of the machines are old lathes, each having a probability of  $\frac{1}{12}$  of needing adjust during the working day, the remaining seven machines are new having a corresponding probability of needing adjustment of  $\frac{1}{22}$ .

Assuming that no machine needs adjust twice on the same day. Determine the probability that two and only two old

machines need adjustment and no new machines need adjustment, on the same day.

This may appear to be a rather wordy example, but it is useful to illustrate the technique of combining the laws of probability with the rule for selecting combinations.

We first need to determine the probability that two of the old machines need adjustment but that the third does not.

Thus using the probability law for independent events then the, probability that two old machines need adjustment =  $\left(\frac{1}{12}\right)\left(\frac{1}{12}\right)$  and combined with the probability that the third does not,  $\left(\frac{11}{12}\right)$  we get:  $\left(\frac{1}{12}\right)^2\left(\frac{11}{12}\right)$ .

Now we need to consider the ways of selecting any two old machines at random from the three possible, that is  $\binom{3}{2}$  ways.

Thus the total probability that two and only two of the three old machines need adjustment is:  $\binom{3}{2}\left(\frac{1}{12}\right)^2\left(\frac{11}{12}\right)$ . Also, the probability that *no new* machines need adjustment is,  $\left(\frac{21}{22}\right)^7$ .

Therefore, the probability that two and only two old machines and no new machines need adjustment is:

$$\binom{3}{2}\left(\frac{1}{12}\right)^2\left(\frac{11}{12}\right)\left(\frac{21}{22}\right)^7 = \mathbf{0.0138}.$$

Think about, the probability of just two new machines and no old machines, needing adjustment. You will be asked to find this probability later!

### ***The binomial distribution***

You have already met the binomial expansion, when you studied series earlier (see page 448). Then we wrote the expansion in general form as:

$$(a + b)^n = {}^nC_0a^n + {}^nC_1a^{n-1}b + {}^nC_2a^{n-2}b^2 + \dots {}^nC_nb^n.$$

Now in this section we write the same expansion in terms of the *probability* of success ( $p$ ) and the probability of failure ( $q$ ), where  $q = 1 - p$ . Then we represent the binomial distribution as:

$$(q + p)^n = q^n + \binom{n}{1}q^{n-1}p + \binom{n}{2}q^{n-2}p^2 + \dots + p^n \quad \text{where,}$$

$${}^nC_r = \binom{n}{r}.$$

The use and power of the binomial probability distribution is best emphasised by example.

**Example 6.4.13**

Use the binomial expansion to determine:

- the probabilities of obtaining 0, 1, 2, 3 sixes, when *three* dice are thrown together;
- the probability of obtaining two or more sixes, when the three dice are thrown together;
- estimates for the expected frequencies, if the three dice are thrown together 1000 times.

(a) All that needs to be recognised is that the terms of the binomial expansion, give the probabilities of obtaining a 0, 1, 2, 3, ...,  $n$  sixes, when  $n$  dice are thrown. In this case the probability of success  $p = \frac{1}{6}$  and the probability of failure, i.e. not throwing a six  $= \frac{5}{6}$ . So for the above case, the required probabilities can be found by calculating the terms of the expansion,

$$\begin{aligned}(q + p)^3 &= q^3 + \binom{3}{1}q^{3-1}p + \binom{3}{2}q^2p^2 + p^3 \\&= \left(\frac{5}{6}\right)^3 + \binom{3}{1}\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right) + \binom{3}{2}\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^2 + \left(\frac{1}{6}\right)^3 \\&= \frac{125}{216} + \frac{75}{216} + \frac{15}{216} + \frac{1}{216}.\end{aligned}$$

Then the probability of throwing, 0, 1, 2 or 3 sixes is:

$$P(0) = \frac{125}{216}, P(1) = \frac{75}{216}, P(2) = \frac{15}{216}, P(3) = \frac{1}{216}.$$

(b) Two or more sixes. This is easily resolved by remembering that two or more is given by  $1 - [P(0) + P(1)]$  in this case we get,

$$1 - \left[ \frac{125}{216} + \frac{75}{216} \right] = 1 - \frac{200}{216} = \frac{16}{216}.$$

Of course, all we needed to do with this example was to add  $P(2)$  and  $P(3)$  together, to get the same result. Imagine how tedious this latter method would be if  $n = 6$  or more!

(c) The probable frequency of occurrence for each event are tabulated below. Should you be so inclined, you might like to carry out an experiment with three dice, throwing them simultaneously 1000 times and recording and comparing the results, with the theoretical probabilities given below!

| Number of sixes | Probability $P(x)$       | Probable frequency |
|-----------------|--------------------------|--------------------|
| 0               | $P(0) = \frac{125}{216}$ | 578.7              |
| 1               | $P(1) = \frac{75}{216}$  | 347.3              |

| Number of sixes | Probability $P(x)$      | Probable frequency |
|-----------------|-------------------------|--------------------|
| 2               | $P(2) = \frac{15}{216}$ | 69.4               |
| 3               | $P(3) = \frac{1}{216}$  | 4.6                |
| Totals          | $P(T) = 1$              | 1000               |

### Example 6.4.14

In a large batch of electric-light bulbs 5% of the bulbs are defective. What is the probability that in a random sample of 20 light bulbs, there will be at most two defectives?

Here again we use the first three terms of the binomial expansion to yield the values in the table, where  $p = 0.05$  and  $q = 0.95$ .

| Number of defectives<br>(x) in sample of 20 | Probability $P(x)$                               |
|---------------------------------------------|--------------------------------------------------|
| 0                                           | $P(0) = (0.95)^{20} = 0.3581$                    |
| 1                                           | $P(1) = \binom{20}{1}(0.95)^{19}(0.05) = 0.3770$ |
| 2                                           | $P(2) = \binom{20}{2}(0.9)^{18}(0.1)^2 = 0.1885$ |
| At most 2                                   | $P(\leq 2) = P(0) + P(1) + P(2) = 0.9236$        |

If we wished to find the probability of more than two defectives then we use the rule, as before i.e.

$$P(>2) = 1 - P(0) - P(1) - P(2) = 1 - 0.9236 = \mathbf{0.0764}.$$

### The Poisson distribution

So far with the binomial distribution we have been dealing with relatively small samples, where  $n$  is small and the probability  $P$  of an event has also been relatively large. When we use large samples and the probability of an event is small, then we may approximate the binomial distribution with the *Poisson distribution*. The formal proof for the Poisson approximation is not given here but in practice the results for large samples are very closely to those produced by the binomial, without the need for the tedious amount of arithmetic required when using the binomial expansion!

In formula 4 above, we replace  $np$  by  $a$ , then the set of Poisson probabilities is given by:

$$e^{-a}, ae^{-a}, \frac{a^2}{2!}e^{-a}, \dots, \frac{a^n}{n!}e^{-a}.$$

**Example 6.4.15**

One percent of the fittings produced in a workshop are defective. What is the probability that there will be four or more defectives in a random sample of 200?

This problem could be solved by substituting  $p = 0.01$ ,  $q = 0.99$ , and  $n = 200$  in the binomial distribution. However, we will find the solution using the Poisson distribution, where  $a = pn = (0.01)(200) = 2$ .

Now the probability of four or more defectives is the equivalent to  $1 - \{P(0) + P(1) + P(2) + P(3)\}$ . Then using formula 4, in tabular form, we get:

| Number of defectives<br>(x) in sample of 200 | Probability P(x)                                      |
|----------------------------------------------|-------------------------------------------------------|
| 0                                            | $P(0) = e^{-2} = 0.1353$                              |
| 1                                            | $P(1) = 2e^{-2} = 0.2076$                             |
| 2                                            | $P(2) = \frac{2^2}{(1)(2)} e^{-2} = 0.2706$           |
| 3                                            | $P(3) = \frac{2^3}{(1)(2)(3)} e^{-2} = 0.1804$        |
| 4 or more                                    | $P(>3) = 1 - P(0) + P(1)$<br>$+ P(2) + P(3) = 0.2151$ |

That is a more than 21% change of four or more defectives in the sample of 200!

We now move on to consider *continuous probability distributions*, in particular the normal distribution.

**Continuous distributions**

When  $X$  say, is a continuous random variable and we wish to find the probabilities concerning  $X$ , we use a continuous function known as a *density function*. Essentially we produce a graph to show how the continuous function  $f(x)$  varies against parameters such as time, distance, etc. Then to find the value of the random variable  $X$ , we find the area under the curve of the density function  $f(x)$ , between the required values. Of course this pre-supposes that the density function for the variable can be found.

This may sound rather complicated, but it is easily understood by example.

**Example 6.4.16**

Given that the random variable  $X$  has a density function

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{else where} \end{cases} \quad \text{Find } P(0.5 < x < 0.75).$$

Make sure that you are familiar with this kind of terminology. In words we are saying that the function  $f(x) = 2x$  when  $x$  is

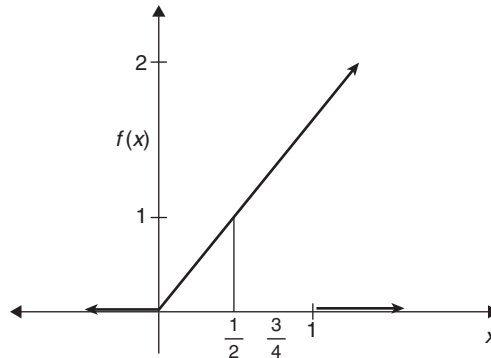


**Figure 6.4.10** Graph of density function  $f(x)$ .

greater than 0 and less than 1.0 and for all other values of the random variable outside of these limits, the function  $f(x) = 0$ .

Now the graph of this function is shown in Figure 6.4.10.

Now the total area under  $f(x)$  between the limits of the function 0 to 1, is given by the triangle with base  $x$  varying between 0 to 1 and corresponding values of  $f(x)$  from 0 to 2.

Thus the total area of the density function,  $f(x) = (0.5)(1)(2) = 1.0$  (remembering that the area of a right-angled triangle is given by  $A = \text{half the base multiplied by the perpendicular height}$ ). This shows that  $f(x)$  is a *proper probability density function* where, the sum of all the probabilities for the random variable  $X$ , must add up to 1.0.

Now to find the probability of that  $P(0.5 < x < 0.75)$ , we find the area of the shaded region in the diagram. This is the *difference* between the triangles with vertices  $\{(0, 0), (0.5, 0), (0.5, f(0.5))\}$  and  $\{(0, 0), (0.75, 0), (0.75, f(0.75))\}$  where  $f(0.5) = 1.0$  and  $f(0.75) = 1.5$ .

$$\begin{aligned} \text{Thus } P(0.5 < x < 0.75) &= (0.5)(0.75)(1.5) - (0.5)(0.5)(1.0) \\ &= 0.5625 - 0.25 \\ &= 0.3125 \quad \text{or} \quad \frac{5}{16}. \end{aligned}$$

### The normal distribution

The normal distribution is recognised as a *bell-shaped curve* (Figure 6.4.11). It is used frequently to describe the distribution of dimensions, mass, and other important parameters.

The *normal distribution* is a continuous probability distribution and as such is described by the *density function*:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} \quad \text{or} \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

where the variable  $-\infty < x < +\infty$ .

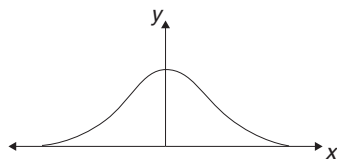
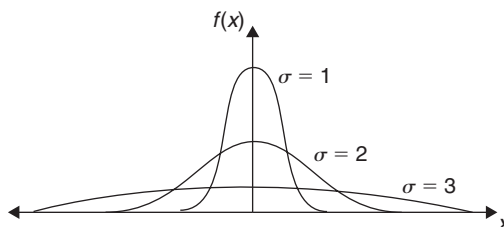


**Figure 6.4.11** The normal distribution probability density function curve

This distribution is symmetrical with respect to the origin of the coordinates and because of this symmetry it has a maximum value at  $x = 0$ . Also, the *mean value* of the random variable ( $\bar{x}$ ) is found at 0, i.e.  $(\bar{x}) = 0$ .

**Figure 6.4.12** Variation in shape of normal distribution, with respect to the value of the standard deviation



The standard deviation of the normal distribution, defines the shape of the probability density function, in that it defines its width (Figure 6.4.12).

It should be noted that as  $\sigma$  increases the curve becomes longer and flatter and with small values of standard deviation  $\sigma$ , the distribution is much more closely bunched around the mean. However, whatever the value of  $\sigma$ , the area under the curve remains constant. Since the curve represents a *probability density function*, then the sum of all the probabilities must be equal to 1.

That is the area under the curve, given by the integral

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad \text{or} \quad \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} dx = 1.$$

This function is far more complex than the one described in Example 6.4.16 and as such requires us to integrate between limits in order to find the individual probabilities of the random variable  $x$ .

A new normal distribution is obtained if we shift it along the  $x$ -axis by an amount  $\mu$ . Then this new distribution may be represented by replacing

$x$  by  $(x - \mu)$ , so that now:  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}$ . This function,

now has its maximum value at  $x = \mu$  and the mean value of the random variable, in this case is  $\bar{x} = \mu$ .

You will also note in formulae 2, for the normal distribution, that *the area under the curve, to the left of a given value of the random variable  $x$* , is given by:

$$A(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}z^2} dz. \quad \text{Where} \quad z = \frac{x - \mu}{\sigma}$$

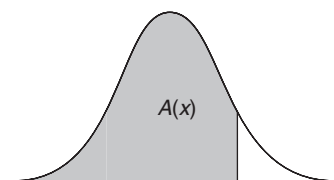


**Figure 6.4.13** The area  $A(x)$  to the left of a given value of  $x$

(see standardising the variable, next). This value is of great practical importance, **gives the area  $A(x)$  under the curve to the left of a given value of  $x$** . This idea is illustrated in Figure 6.4.13.

Thus, from the table on the next page,  $A(0) = 0.5000$  indicates that half the area under the curve is to the left of  $x = 0$ . Similarly for  $A(2) = 0.97725$  indicates that 0.97725 of the area or 97.725% of the area is to the left of  $x = 2$ . So that, for example,  $A(2) - A(0) = 0.47725$  indicates that 44.725% of the area lies, between the ordinates of the random variable  $x = 0$  and  $x = 2$ .

Suppose we were required to know the area under the normal curve (the probability of occurrence) between the variable  $x = -1.65$  to  $x = 1.65$ . Then from the table  $A(1.65) = 0.9505$ . Thus approximately 95% of the area is to the left of  $x = 1.65$  and approximately 5% is to the right. As the normal curve is symmetrical about the  $f(x)$ -axis, it follows



Table of values for area  $A(x)$  under normal probability distribution, to the left of a given value of  $x$ , where  $A(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}z^2} dz$

| $x$  | $A(x)$               | $x$  | $A(x)$               | $x$  | $A(x)$               | $x$  | $A(x)$               | $x$  | $A(x)$                | $x$  | $A(x)$                | $x$  | $A(x)$                |
|------|----------------------|------|----------------------|------|----------------------|------|----------------------|------|-----------------------|------|-----------------------|------|-----------------------|
| 0.00 | 0.5000 <sub>40</sub> | 0.45 | 0.6736 <sub>36</sub> | 0.90 | 0.8159 <sub>27</sub> | 1.35 | 0.9115 <sub>16</sub> | 1.80 | 0.9641 <sub>8</sub>   | 2.25 | 0.98778 <sub>31</sub> | 2.70 | 0.99653 <sub>11</sub> |
| 0.01 | 0.5040 <sub>40</sub> | 0.46 | 0.6772 <sub>36</sub> | 0.91 | 0.8186 <sub>26</sub> | 1.36 | 0.9131 <sub>16</sub> | 1.81 | 0.9649 <sub>7</sub>   | 2.26 | 0.98809 <sub>31</sub> | 2.71 | 0.99664 <sub>10</sub> |
| 0.02 | 0.5080 <sub>40</sub> | 0.47 | 0.6808 <sub>36</sub> | 0.92 | 0.8212 <sub>26</sub> | 1.37 | 0.9147 <sub>15</sub> | 1.82 | 0.9656 <sub>8</sub>   | 2.27 | 0.98840 <sub>30</sub> | 2.72 | 0.99674 <sub>9</sub>  |
| 0.03 | 0.5120 <sub>40</sub> | 0.48 | 0.6844 <sub>35</sub> | 0.93 | 0.8238 <sub>26</sub> | 1.38 | 0.9162 <sub>15</sub> | 1.83 | 0.9664 <sub>7</sub>   | 2.28 | 0.98870 <sub>29</sub> | 2.73 | 0.99683 <sub>10</sub> |
| 0.04 | 0.5160 <sub>39</sub> | 0.49 | 0.6879 <sub>36</sub> | 0.94 | 0.8264 <sub>25</sub> | 1.39 | 0.9177 <sub>15</sub> | 1.84 | 0.9671 <sub>7</sub>   | 2.29 | 0.98899 <sub>29</sub> | 2.74 | 0.99693 <sub>9</sub>  |
| 0.05 | 0.5199 <sub>40</sub> | 0.50 | 0.6915 <sub>35</sub> | 0.95 | 0.8289 <sub>26</sub> | 1.40 | 0.9192 <sub>15</sub> | 1.85 | 0.9678 <sub>8</sub>   | 2.30 | 0.98928 <sub>28</sub> | 2.75 | 0.99702 <sub>9</sub>  |
| 0.06 | 0.5239 <sub>40</sub> | 0.51 | 0.6950 <sub>35</sub> | 0.96 | 0.8315 <sub>25</sub> | 1.41 | 0.9207 <sub>15</sub> | 1.86 | 0.9686 <sub>7</sub>   | 2.31 | 0.98956 <sub>27</sub> | 2.76 | 0.99711 <sub>9</sub>  |
| 0.07 | 0.5279 <sub>40</sub> | 0.52 | 0.6985 <sub>34</sub> | 0.97 | 0.8340 <sub>25</sub> | 1.42 | 0.9222 <sub>14</sub> | 1.87 | 0.9693 <sub>6</sub>   | 2.32 | 0.98983 <sub>27</sub> | 2.77 | 0.99720 <sub>8</sub>  |
| 0.08 | 0.5319 <sub>40</sub> | 0.53 | 0.7019 <sub>35</sub> | 0.98 | 0.8365 <sub>24</sub> | 1.43 | 0.9236 <sub>15</sub> | 1.88 | 0.9699 <sub>7</sub>   | 2.33 | 0.99010 <sub>26</sub> | 2.78 | 0.99728 <sub>8</sub>  |
| 0.09 | 0.5359 <sub>39</sub> | 0.54 | 0.7054 <sub>34</sub> | 0.99 | 0.8389 <sub>24</sub> | 1.44 | 0.9251 <sub>14</sub> | 1.89 | 0.9706 <sub>7</sub>   | 2.34 | 0.99036 <sub>25</sub> | 2.79 | 0.99736 <sub>8</sub>  |
| 0.10 | 0.5398 <sub>40</sub> | 0.55 | 0.7088 <sub>35</sub> | 1.00 | 0.8413 <sub>25</sub> | 1.45 | 0.9265 <sub>14</sub> | 1.90 | 0.9713 <sub>6</sub>   | 2.35 | 0.99061 <sub>25</sub> | 2.80 | 0.99744 <sub>8</sub>  |
| 0.11 | 0.5438 <sub>40</sub> | 0.56 | 0.7123 <sub>34</sub> | 1.01 | 0.8438 <sub>23</sub> | 1.46 | 0.9279 <sub>13</sub> | 1.91 | 0.9719 <sub>7</sub>   | 2.36 | 0.99086 <sub>25</sub> | 2.81 | 0.99752 <sub>8</sub>  |
| 0.12 | 0.5478 <sub>39</sub> | 0.57 | 0.7157 <sub>33</sub> | 1.02 | 0.8461 <sub>24</sub> | 1.47 | 0.9292 <sub>14</sub> | 1.92 | 0.9726 <sub>6</sub>   | 2.37 | 0.99111 <sub>23</sub> | 2.82 | 0.99760 <sub>7</sub>  |
| 0.13 | 0.5517 <sub>40</sub> | 0.58 | 0.7190 <sub>34</sub> | 1.03 | 0.8485 <sub>23</sub> | 1.48 | 0.9306 <sub>13</sub> | 1.93 | 0.9732 <sub>6</sub>   | 2.38 | 0.99134 <sub>24</sub> | 2.83 | 0.99767 <sub>7</sub>  |
| 0.14 | 0.5557 <sub>39</sub> | 0.59 | 0.7224 <sub>33</sub> | 1.04 | 0.8508 <sub>23</sub> | 1.49 | 0.9319 <sub>13</sub> | 1.94 | 0.9738 <sub>6</sub>   | 2.39 | 0.99158 <sub>22</sub> | 2.84 | 0.99774 <sub>7</sub>  |
| 0.15 | 0.5596 <sub>40</sub> | 0.60 | 0.7257 <sub>34</sub> | 1.05 | 0.8531 <sub>23</sub> | 1.50 | 0.9332 <sub>13</sub> | 1.95 | 0.9744 <sub>6</sub>   | 2.40 | 0.99180 <sub>22</sub> | 2.85 | 0.99781 <sub>7</sub>  |
| 0.16 | 0.5636 <sub>39</sub> | 0.61 | 0.7291 <sub>33</sub> | 1.06 | 0.8554 <sub>23</sub> | 1.51 | 0.9345 <sub>12</sub> | 1.96 | 0.9750 <sub>6</sub>   | 2.41 | 0.99202 <sub>22</sub> | 2.86 | 0.99788 <sub>7</sub>  |
| 0.17 | 0.5675 <sub>39</sub> | 0.62 | 0.7324 <sub>33</sub> | 1.07 | 0.8577 <sub>22</sub> | 1.52 | 0.9357 <sub>13</sub> | 1.97 | 0.9756 <sub>5</sub>   | 2.42 | 0.99224 <sub>21</sub> | 2.87 | 0.99795 <sub>6</sub>  |
| 0.18 | 0.5714 <sub>39</sub> | 0.63 | 0.7357 <sub>32</sub> | 1.08 | 0.8599 <sub>22</sub> | 1.53 | 0.9370 <sub>12</sub> | 1.98 | 0.9761 <sub>6</sub>   | 2.43 | 0.99245 <sub>21</sub> | 2.88 | 0.99801 <sub>6</sub>  |
| 0.19 | 0.5753 <sub>40</sub> | 0.64 | 0.7389 <sub>33</sub> | 1.09 | 0.8621 <sub>22</sub> | 1.54 | 0.9382 <sub>12</sub> | 1.99 | 0.9767 <sub>5</sub>   | 2.44 | 0.99266 <sub>20</sub> | 2.89 | 0.99807 <sub>6</sub>  |
| 0.20 | 0.5793 <sub>39</sub> | 0.65 | 0.7422 <sub>32</sub> | 1.10 | 0.8643 <sub>22</sub> | 1.55 | 0.9394 <sub>12</sub> | 2.00 | 0.97725 <sub>53</sub> | 2.45 | 0.99286 <sub>19</sub> | 2.90 | 0.99813 <sub>6</sub>  |
| 0.21 | 0.5832 <sub>39</sub> | 0.66 | 0.7454 <sub>32</sub> | 1.11 | 0.8665 <sub>21</sub> | 1.56 | 0.9406 <sub>12</sub> | 2.01 | 0.97778 <sub>53</sub> | 2.46 | 0.99305 <sub>19</sub> | 2.91 | 0.99819 <sub>6</sub>  |
| 0.22 | 0.5871 <sub>39</sub> | 0.67 | 0.7486 <sub>31</sub> | 1.12 | 0.8686 <sub>22</sub> | 1.57 | 0.9418 <sub>11</sub> | 2.02 | 0.97831 <sub>51</sub> | 2.47 | 0.99324 <sub>19</sub> | 2.92 | 0.99825 <sub>6</sub>  |
| 0.23 | 0.5910 <sub>38</sub> | 0.68 | 0.7517 <sub>32</sub> | 1.13 | 0.8708 <sub>21</sub> | 1.58 | 0.9429 <sub>12</sub> | 2.03 | 0.97882 <sub>10</sub> | 2.48 | 0.99343 <sub>18</sub> | 2.93 | 0.99831 <sub>5</sub>  |
| 0.24 | 0.5948 <sub>39</sub> | 0.69 | 0.7549 <sub>31</sub> | 1.14 | 0.8729 <sub>20</sub> | 1.59 | 0.9441 <sub>11</sub> | 2.04 | 0.97932 <sub>50</sub> | 2.49 | 0.99361 <sub>18</sub> | 2.94 | 0.99836 <sub>5</sub>  |
| 0.25 | 0.5987 <sub>39</sub> | 0.70 | 0.7580 <sub>31</sub> | 1.15 | 0.8749 <sub>21</sub> | 1.60 | 0.9452 <sub>11</sub> | 2.05 | 0.97982 <sub>48</sub> | 2.50 | 0.99379 <sub>17</sub> | 2.95 | 0.99841 <sub>5</sub>  |
| 0.26 | 0.6026 <sub>38</sub> | 0.71 | 0.7611 <sub>31</sub> | 1.16 | 0.8770 <sub>20</sub> | 1.61 | 0.9463 <sub>11</sub> | 2.06 | 0.98030 <sub>47</sub> | 2.51 | 0.99396 <sub>17</sub> | 2.96 | 0.99846 <sub>5</sub>  |
| 0.27 | 0.6064 <sub>39</sub> | 0.72 | 0.7642 <sub>31</sub> | 1.17 | 0.8790 <sub>20</sub> | 1.62 | 0.9474 <sub>10</sub> | 2.07 | 0.98077 <sub>47</sub> | 2.52 | 0.99413 <sub>17</sub> | 2.97 | 0.99851 <sub>5</sub>  |
| 0.28 | 0.6103 <sub>38</sub> | 0.73 | 0.7673 <sub>31</sub> | 1.18 | 0.8810 <sub>20</sub> | 1.63 | 0.9484 <sub>11</sub> | 2.08 | 0.98124 <sub>45</sub> | 2.53 | 0.99430 <sub>16</sub> | 2.98 | 0.99856 <sub>5</sub>  |
| 0.29 | 0.6141 <sub>38</sub> | 0.74 | 0.7704 <sub>30</sub> | 1.19 | 0.8830 <sub>19</sub> | 1.64 | 0.9495 <sub>10</sub> | 2.09 | 0.98169 <sub>75</sub> | 2.54 | 0.99446 <sub>15</sub> | 2.99 | 0.99861 <sub>4</sub>  |
| 0.30 | 0.6179 <sub>38</sub> | 0.75 | 0.7734 <sub>30</sub> | 1.20 | 0.8849 <sub>20</sub> | 1.65 | 0.9505 <sub>10</sub> | 2.10 | 0.98214 <sub>43</sub> | 2.55 | 0.99461 <sub>16</sub> | 3.0  | 0.99865 <sub>38</sub> |
| 0.31 | 0.6217 <sub>38</sub> | 0.76 | 0.7764 <sub>30</sub> | 1.21 | 0.8869 <sub>19</sub> | 1.66 | 0.9515 <sub>10</sub> | 2.11 | 0.98257 <sub>43</sub> | 2.56 | 0.99477 <sub>15</sub> | 3.1  | 0.99903 <sub>28</sub> |
| 0.32 | 0.6255 <sub>38</sub> | 0.77 | 0.7794 <sub>29</sub> | 1.22 | 0.8888 <sub>19</sub> | 1.67 | 0.9525 <sub>10</sub> | 2.12 | 0.98300 <sub>41</sub> | 2.57 | 0.99492 <sub>14</sub> | 3.2  | 0.99931 <sub>21</sub> |
| 0.33 | 0.6293 <sub>38</sub> | 0.78 | 0.7823 <sub>29</sub> | 1.23 | 0.8907 <sub>18</sub> | 1.68 | 0.9535 <sub>10</sub> | 2.13 | 0.98341 <sub>41</sub> | 2.58 | 0.99506 <sub>14</sub> | 3.3  | 0.99952 <sub>14</sub> |
| 0.34 | 0.6331 <sub>37</sub> | 0.79 | 0.7852 <sub>29</sub> | 1.24 | 0.8925 <sub>19</sub> | 1.69 | 0.9545 <sub>9</sub>  | 2.14 | 0.98392 <sub>40</sub> | 2.59 | 0.99520 <sub>14</sub> | 3.4  | 0.99966 <sub>11</sub> |
| 0.35 | 0.6368 <sub>38</sub> | 0.80 | 0.7881 <sub>29</sub> | 1.25 | 0.8944 <sub>18</sub> | 1.70 | 0.9554 <sub>10</sub> | 2.15 | 0.98422 <sub>39</sub> | 2.60 | 0.99534 <sub>13</sub> | 3.5  | 0.99977 <sub>7</sub>  |
| 0.36 | 0.6406 <sub>37</sub> | 0.81 | 0.7910 <sub>29</sub> | 1.26 | 0.8962 <sub>18</sub> | 1.71 | 0.9564 <sub>9</sub>  | 2.16 | 0.98461 <sub>39</sub> | 2.61 | 0.99547 <sub>13</sub> | 3.6  | 0.99984 <sub>5</sub>  |
| 0.37 | 0.6443 <sub>37</sub> | 0.82 | 0.7939 <sub>28</sub> | 1.27 | 0.8980 <sub>17</sub> | 1.72 | 0.9573 <sub>9</sub>  | 2.17 | 0.98500 <sub>37</sub> | 2.62 | 0.99560 <sub>13</sub> | 3.7  | 0.99989 <sub>4</sub>  |
| 0.38 | 0.6480 <sub>37</sub> | 0.83 | 0.7967 <sub>28</sub> | 1.28 | 0.8997 <sub>18</sub> | 1.73 | 0.9582 <sub>9</sub>  | 2.18 | 0.98537 <sub>37</sub> | 2.63 | 0.99573 <sub>12</sub> | 3.8  | 0.99993 <sub>2</sub>  |
| 0.39 | 0.6517 <sub>37</sub> | 0.84 | 0.7995 <sub>28</sub> | 1.29 | 0.9015 <sub>17</sub> | 1.74 | 0.9591 <sub>9</sub>  | 2.19 | 0.98574 <sub>36</sub> | 2.64 | 0.99585 <sub>13</sub> | 3.9  | 0.99995 <sub>2</sub>  |
| 0.40 | 0.6554 <sub>37</sub> | 0.85 | 0.8023 <sub>28</sub> | 1.30 | 0.9032 <sub>17</sub> | 1.75 | 0.9599 <sub>9</sub>  | 2.20 | 0.98610 <sub>35</sub> | 2.65 | 0.99598 <sub>11</sub> | 4.0  | 0.99997               |
| 0.41 | 0.6591 <sub>37</sub> | 0.86 | 0.8051 <sub>27</sub> | 1.31 | 0.9049 <sub>17</sub> | 1.76 | 0.9608 <sub>8</sub>  | 2.21 | 0.98645 <sub>34</sub> | 2.66 | 0.99609 <sub>12</sub> |      |                       |
| 0.42 | 0.6628 <sub>36</sub> | 0.87 | 0.8078 <sub>28</sub> | 1.32 | 0.9066 <sub>16</sub> | 1.77 | 0.9616 <sub>9</sub>  | 2.22 | 0.98679 <sub>34</sub> | 2.67 | 0.99621 <sub>11</sub> |      |                       |
| 0.43 | 0.6664 <sub>36</sub> | 0.88 | 0.8106 <sub>27</sub> | 1.33 | 0.9082 <sub>17</sub> | 1.78 | 0.9625 <sub>8</sub>  | 2.23 | 0.98713 <sub>32</sub> | 2.68 | 0.99632 <sub>11</sub> |      |                       |
| 0.44 | 0.6700 <sub>36</sub> | 0.89 | 0.8133 <sub>26</sub> | 1.34 | 0.9099 <sub>16</sub> | 1.79 | 0.9633 <sub>8</sub>  | 2.24 | 0.98745 <sub>33</sub> | 2.69 | 0.99643 <sub>10</sub> |      |                       |

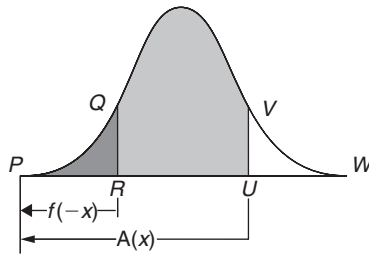


**Figure 6.4.14**  $A(-x) = 1 - A(x)$ , since area PQR = UVW

that approximately 5% of the area lies to the left of  $x = -1.65$ . Therefore, we can write,

$$A(-1.65) = 1 - A(1.65) \text{ or in general } A(-x) = 1 - A(x)$$

This relationship is illustrated in Figure 6.4.14.

**Note:** The use of the standard normal accumulative distribution function, in the form of a table is becoming unnecessary as electronic spreadsheets, such as Microsoft Excel have built-in statistical functions.

**Standardising the variable** In order to use the table for normal distributions other than those with a mean  $\bar{x} = 0$  and  $\sigma = 1.0$ , it is necessary to manipulate the data in a special way. This is known as *standardising the variable*. This standardisation is achieved by use of the formula:

$$z = \frac{x - \mu}{\sigma}, \text{ where } \mu = \bar{x}$$

that is the mean of the given data and  $\sigma$  is the standard deviation for this data.

### Example 6.4.17

Given that a batch of components have a mean of 50 and standard deviation of 10, for a set of measurements that are normally distributed. Find the probability that a randomly selected component lies between the measurements 50 and 55.

This rather trivial example illustrate clearly the process we adopt to find the probabilities, after normalising the variable.

Thus we wish to find  $P(50 \leq x \leq 55)$  where  $x$  is normally distributed with mean  $\mu = 50$  and standard deviation  $\sigma = 10$ . To standardise our  $x$  values to say  $z$  values we use,

$$z = \frac{x - \mu}{\sigma}, \text{ so the,}$$

$$\begin{aligned} P(50 \leq x \leq 55) &= P\left(\frac{50 - 50}{10} \leq \frac{x - \mu}{\sigma} \leq \frac{55 - 50}{10}\right) \\ &= P(0 \leq z \leq 0.5) \end{aligned}$$

$$\text{and } P(z \leq 0.5) - P(z \leq 0) = P(0 \leq z \leq 0.5)$$

So, from the table

$$P(z \leq 0.5) = 0.691 \text{ and } P(z \leq 0) = 0.5 \text{ or on substitution, } P(50 \leq x \leq 55) = 0.691 - 0.5 = 0.191.$$

You should check these values for yourself, making sure you can read the table.

Let us now consider the significance of these  $z$  values, where  $z = 1, 2, 3$ , etc means that for these values the variable is 1, 2, 3, etc. *standard deviations* from the mean value of the data. Looking at the area under the normal distribution curve, you will note that for  $\pm 2\sigma$  (standard deviations) then the probability that any variable falling outside of this area belongs to the parent population, is small. In fact, from the table it is  $(1 - 0.97725) = 0.02275 = 2.275\%$ . So if it is even further out the chances of it belonging to the parent population is highly improbable!

**Example 6.4.18**

The melting point of the polymer polytetrafluoroethylene (PTFE) is known to be 327°C. This average value has been found experimentally from a large number of tests. Suppose that the standard deviation in the melting point of PTFE was found to be 3°C. What is the probability that an unknown polymer that was found to have a melting temperature of 339°C is PTFE?

Let  $X$  be the random variable denoting the observed melting point of PTFE.  $X$  is assumed to be normally distributed with mean 327°C and standard deviation 3°C.

To determine this probability we standardise the variable i.e.

$$z = \frac{x - \mu}{\sigma} = \frac{339 - 327}{3} = 4.$$

That is, the observation is 4 standard deviations from the mean. From the argument before, and the value from the table of  $1 - 0.99997 = 0.00003$ , it can be seen that the chance of this material being PTFE is extremely unlikely!

**Test your knowledge 6.4.2**

- (1) A bag contains four white balls, six black balls, three red balls, and eight yellow balls. If one ball is drawn from the bag find the probability that it is either a white or a yellow ball. (4/7)
- (2) Find the probability of scoring a 5 or 6 or 7 on one throw of a pair of dice. (5/12)
- (3) A pound coin is to be tossed three times. What is the probability there will be two heads and one tail? (3/8)
- (4) A hydraulic system has two easily removed pumps A and B, that are both active under normal circumstances. If either pump fails, the system is deemed to be inoperative. The probability that it will be necessary to replace pump A is 0.50. The failure of pump A has a detrimental effect on pump B. If pump A has to be replaced the probability that pump B will also have to be replaced is 0.70. If pump A remains serviceable the probability of failure of pump B is reduced to 0.10. What percentage of all failures requires both pump A and pump B to be replaced? (35%)
- (5) In a factory four machines produce the same product. Machine A produces 10% of the output; machine B, 20%; machine C, 30%; and machine D, the remainder. The proportion of defective items produced by each machine is as follows: machine A, 0.001; machine B, 0.0005; machine C, 0.005; and machine D, 0.002. An item selected at random is found to be defective. What is the

- probability that the item was produced by (i) machine A? (ii) machine D? (1/25), (8/25)
- (6) Use the binomial distribution to expand  $(a + 2b)^5$   
 $(a^5 + 10a^4b + 40a^3b^2 + 80a^2b^3 + 80ab^4 + 32b^5)$
- (7) Use the binomial expansion to determine the probability of getting exactly three heads from five tosses of a balanced coin (5/16).
- (8) A machine produces valve bodies of which 1% are defective. If a random sample of 50 valve bodies is drawn from a large batch. What is the probability that the sample contains no defective items? (0.605)
- (9) If the random variable  $x$  is normally distributed with mean = 9 and standard deviation = 3. Find  $P(5 < x < 11)$  (0.6536)
- (10) A company manufactures transistors where the average lifespan is 1000 h and the standard deviation is 100 h. Assuming that transistor lifespan is normally distributed. Find the probability that a transistor selected at random, will have a lifespan between 850 h and 1075 h. (0.7066)

## Regression and correlation

The final set of techniques we require concern the methods we use to assess the reliability of data drawn from experiments that involve the relationship between two physical quantities.

We will use two methods to assess the reliability of experimental data. The first, involves the determination of *the line of best fit* for the data and to establish the law connecting the data. The second method involves the *degree of correlation* that exists between the two physical quantities.

### Formulae

- Least squares regression line

$$\text{For the line, } y = ax + b, \quad a = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} \quad \text{and} \quad b = \bar{y} - a\bar{x}$$

- Correlation and correlation coefficient

$$\text{Correlation } r^2 = \frac{(\sum x_i y_i - n\bar{x}\bar{y})^2}{(\sum x_i^2 - n\bar{x}^2)(\sum y_i^2 - n\bar{y}^2)}$$

$$\text{Correlation coefficient } r = \frac{(\sum x_i y_i - n\bar{x}\bar{y})}{\sqrt{(\sum x_i^2 - n\bar{x}^2)(\sum y_i^2 - n\bar{y}^2)}}$$

## Methods

Before we consider the methods, through examples, we need to briefly explain the formulae which are given here without proof.

The line of best fit uses a method known as *least squares* in order to determine the coefficients  $a$  and  $b$  for the line of best fit, which we try and match to the data. This method takes into account the error margins that occur in the experimental data and enables us to draw a line through the average values of these errors.

In order to see if correlation exists between the two experimental variables ( $x, y$ ) we use the correlation formulae given above, the degree of this correlation is measured using the correlation coefficient.

You should note that the correlation coefficient can take on values between  $-1$  and  $+1$ . We will get  $r = 1$  if all the points ( $x_i, y_i$ ) lie on the straight regression line, with a positive slope. We get  $r = 0$ , if  $x_i$  and  $y_i$  are statistically independent, that is *no correlation exists* between them and they are spread in a circular cluster. The correlation coefficient is *negative*, if the regression line has a negative slope.

For all the formulae given above,  $x_i, y_i$ , are the measured experimental values and  $\bar{x}, \bar{y}$  are the mean values. All the sums are formed from  $i = 1$  to  $i = n$ , where  $n$ , the number of measurements.

### Example 6.4.19

In an experiment to determine the relationship between frequency and inductive reactance within an electrical circuit, the following readings were obtained.

| Frequency<br>(Hz) | Inductive reactance<br>(Ohms) |
|-------------------|-------------------------------|
| 30                | 20                            |
| 60                | 40                            |
| 90                | 65                            |
| 120               | 80                            |
| 150               | 105                           |
| 180               | 120                           |
| 210               | 140                           |
| 240               | 165                           |
| 270               | 185                           |
| 300               | 200                           |

- (i) Assuming a linear relationship exists, determine the equation of the regression line for inductive reactance *against* frequency.
  - (ii) Determine the frequency for which the inductive reactance is  $135 \Omega$ .
- (i) In this case, the dependent variable  $y$ , inductive reactance, while the independent variable  $x$ , frequency. Therefore all that we are required to do is calculate the values of the coefficients  $a$  and  $b$ , using the linear regression formulae. The variables for the formulae are best found using a tabular

method, as shown below.

| Frequency<br>x     | Inductive<br>reactance y | $x^2$                   | xy                     | $y^2$                   |
|--------------------|--------------------------|-------------------------|------------------------|-------------------------|
| 30                 | 20                       | 900                     | 600                    | 400                     |
| 60                 | 40                       | 3600                    | 2400                   | 1600                    |
| 90                 | 65                       | 8100                    | 5850                   | 4225                    |
| 120                | 80                       | 14 400                  | 9600                   | 6400                    |
| 150                | 105                      | 22 500                  | 15 750                 | 11 025                  |
| 180                | 120                      | 32 400                  | 21 600                 | 14 400                  |
| 210                | 140                      | 44 100                  | 29 400                 | 19 600                  |
| 240                | 165                      | 57 600                  | 39 600                 | 27 225                  |
| 270                | 185                      | 72 900                  | 49 950                 | 34 225                  |
| 300                | 200                      | 90 000                  | 60 000                 | 40 000                  |
| $\sum x =$<br>1650 | $\sum y =$<br>1120       | $\sum x^2 =$<br>346 500 | $\sum xy =$<br>234 750 | $\sum y^2 =$<br>159 100 |

Also from the above table,  $\bar{x} = 165$  and  $\bar{y} = 112$ .

$$\text{Therefore coefficient } a = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$a = \frac{234\,750 - [(10)(165)(112)]}{346\,500 - [(10)(165)^2]} = \frac{49\,950}{74\,250} \\ = 0.672727 = 0.673 \text{ to 3 d.p.}$$

similarly,  $b = \bar{y} - a\bar{x}$  or  $b = 112 - (0.672727)(165) = 1$ .

(ii) Therefore the relationship is  $y = 0.673x + 1$  and the frequency  $x$  for which the inductive reactance  $y = 135 \Omega$ , is given by:

$$x = \frac{y - 1}{0.673} = \frac{135 - 1}{0.673} = \mathbf{199 \text{ Hz.}}$$

*Note:* If we had wanted to find the regression line of best fit for frequency against inductive reactance, then frequency would become the dependent variable  $y$  and inductive reactance the independent variable  $x$  in the above formula.

For the above example we have assumed a *linear* relationship. How safe is this assumption? We can now use the correlation coefficient, to make a prediction about the accuracy of our assumption.

### Example 6.4.20

Determine the correlation coefficient for the data given in example 6.4.19 and comment on its value.

$$\text{Then using } r = \frac{(\sum x_i y_i - n\bar{x}\bar{y})}{\sqrt{(\sum x_i^2 - n\bar{x}^2)(\sum y_i^2 - n\bar{y}^2)}} \quad \text{and the}$$

calculated values for the data found in Example 6.6.19 we get:

$$r = \frac{234\,750 - [(10)(165)(112)]}{\sqrt{[346\,500 - (10)(165)^2][159\,100 - (10)(112)^2]}}$$

$$= \mathbf{0.999}.$$

As expected, this figure suggests that nearly all, if not all points lie on the regression line, so the assumption about a linear relationship between frequency and inductive reactance, has very much been verified!

### Example 6.4.21

The data shown below, relates the number of aircraft from a fleet taken out of service over eight equal monthly periods and the corresponding number of maintenance hours spent on these unserviceable aircraft during each month. Calculate the correlation coefficient for this data and comment on its value.

|                         |     |     |     |     |     |     |     |     |
|-------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| Aircraft out of service | 10  | 13  | 4   | 2   | 9   | 8   | 7   | 16  |
| Maintenance hours       | 460 | 515 | 400 | 360 | 440 | 380 | 415 | 570 |

Again the easiest way to tackle this problem is to set up a table of values, as in Example 6.4.19. This table is shown below.

| Aircraft out of service, $x$ | Maintenance hours, $y$ | $x^2$        | $xy$        | $y^2$        |
|------------------------------|------------------------|--------------|-------------|--------------|
| 10                           | 460                    | 100          | 4600        | 211 600      |
| 13                           | 515                    | 169          | 6695        | 265 225      |
| 4                            | 400                    | 16           | 1600        | 160 000      |
| 2                            | 360                    | 4            | 720         | 129 600      |
| 9                            | 440                    | 81           | 3960        | 193 600      |
| 8                            | 380                    | 64           | 3040        | 144 400      |
| 7                            | 415                    | 49           | 2905        | 172 225      |
| 16                           | 570                    | 256          | 9120        | 324 900      |
| $\sum x =$                   | $\sum y =$             | $\sum x^2 =$ | $\sum xy =$ | $\sum y^2 =$ |
| 69                           | 3540                   | 739          | 32 640      | 160 1550     |

Then,  $n = 8$ ,  $\bar{x} = 8.625$  and  $\bar{y} = 442.5$

$$\text{So } r = \frac{32\,640 - [(8)(8.625)(442.5)]}{\sqrt{[739 - (8)(8.625)^2][160\,1550 - (8)(442.5)^2]}}$$

$$= \frac{2\,107.5}{2\,247.22} = \mathbf{0.938}.$$

Thus, there is clearly a very high degree of correlation between the number of unserviceable aircraft and the maintenance hours required to bring them back into service.

**Test your knowledge 6.4.3**

- (1) The sensitivity of a spring balance is to be determined. This is achieved by placing different masses on the balance and recording the spring extension. The results of such a test are given below.

|                 |     |     |     |     |     |
|-----------------|-----|-----|-----|-----|-----|
| Mass (kg)       | 2   | 3   | 4   | 5   | 6   |
| Deflection (cm) | 1.6 | 2.7 | 3.2 | 3.5 | 4.0 |

Assuming a linear relationship, determine the equation of the regression line for deflection *against* mass ( $s = 0.56m + 0.76$ ).

- (2) The relationship between the applied voltage and the electric current, in a circuit, is shown in the table below.

|              |   |    |    |    |    |    |    |
|--------------|---|----|----|----|----|----|----|
| Voltage (V)  | 5 | 11 | 15 | 19 | 24 | 28 | 33 |
| Current (mA) | 2 | 4  | 6  | 8  | 10 | 12 | 14 |

Assuming a linear relationship, determine the equation of the regression line for current *against* voltage. ( $I = -0.48 + 2.27V$ ).

- (3) The following values were obtained between two variables  $X$ ,  $Y$ , in an experiment.

|     |    |    |   |   |     |
|-----|----|----|---|---|-----|
| $X$ | -1 | -2 | 0 | 1 | 1.5 |
| $Y$ | 1  | 3  | 1 | 2 | 4   |

- Find the value of the constants  $a$  and  $b$  in the equation  $Y = aX + b$  ( $Y = 0.256X + 2.226$ ).
- Determine the correlation  $r^2$  for the relationship in (i) and comment on your results (0.079 low).
- Now use linear regression to determine the relationship between  $Y$  and  $X^2$  ( $Y = 0.619X^2 + 3.22$ ).
- Determine the correlation  $r^2$  for the relationship found in part (iii) and again comment on your answer. (0.53 more appropriate)

**Engineering and scientific applications**

In this final section, you will find examples relating to the application of probability and statistical techniques, in particular, to the analysis of data, quality control, and engineering reliability. It is hoped that these applications will assist you with your future studies. We start with the analysis of errors in experimental data.

**Additional formulae for applications**

The following formulae extend those you have already met for measures of central location and dispersion. In particular the formulae for random



sampling, quality control, and reliability theory are given here *without proof*.

(1) *Sampling and quality control*

(a) Then for *large samples*

Single mean,  $\frac{|\mu_0 - m|}{s/\sqrt{n}} > 1.96$ , significant at the 5% level

Single mean,  $\frac{|\mu_0 - m|}{s/\sqrt{n}} > 3.09$ , significant at the 0.2% level

$\frac{|m_1 - m_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} > 1.96$ , significant at the 5% level

$\frac{|m_1 - m_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} > 3.09$ , significant at the 0.2% level

(b) Then for *small samples*

Apply the *t*-test, then:

Single mean,  $t = \frac{|\mu_0 - m|}{s/\sqrt{n}}$ , and  $\nu = n - 1$

$$t = \frac{|m_1 - m_2|}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where dependent on availability of individual observations or standard deviations, we have:

$$s^2 = \frac{\sum (x_1 - m_1)^2 + \sum (x_2 - m_2)^2}{(n_1 - 1) + (n_2 - 1)} \quad (\text{if individual observations are available})$$

$$\text{or } s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)} \quad (\text{if the standard deviations are available but not the individual observations}).$$

(2) *Error arithmetic*

Errors combined by *addition or subtraction*,

$$(\Delta Z)^2 = (\Delta X)^2 + (\Delta Y)^2$$

Errors combined by *multiplication or division*,

$$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta X}{X}\right)^2 + \left(\frac{\Delta Y}{Y}\right)^2$$

Error involving *powers*,

$$\left(\frac{\Delta Z}{\bar{Z}}\right)^2 = n \left(\frac{\Delta X}{\bar{X}}\right)^2$$

*Gauss law of error propagation*

$$\sigma_{mz} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 (\sigma_{mx})^2 + \left(\frac{\partial f}{\partial y}\right)^2 (\sigma_{my})^2}$$

where  $\frac{\partial f}{\partial x}$  and  $\frac{\partial f}{\partial y}$  are the *partial derivatives* of the function  $f$  calculated at the mean values  $\bar{x}$  and  $\bar{y}$ .

### (3) *Reliability theory*

#### (a) *The exponential distribution*

The exponential probability density function is,

$$f(x) = \frac{1}{\theta} e^{-x/\theta} = \lambda e^{-\lambda x}, \quad x \geq 0$$

The exponential reliability function is,

$$R(x) = e^{-x/\theta} = e^{-\lambda x}, \quad x \geq 0$$

The exponential hazard function or failure rate is,

$$h(x) = \frac{f(x)}{R(x)} = \frac{1}{\theta} = \lambda$$

#### (b) *The Poisson probability density function*

$$P(x, \lambda t) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}, \quad x = 0, 1, 2, 3, \dots, n$$

#### (c) *The Weibull distribution*

The Weibull probability density function is,

$$f(x) = \frac{\beta x^{\beta-1}}{\theta^\beta} \exp\left[-\left(\frac{x}{\theta}\right)^\beta\right], \quad x \geq 0$$

Weibull reliability function

$$R(x) = \exp\left[-\left(\frac{x}{\theta}\right)^\beta\right], \quad x > 0$$

Weibull hazard function

$$h(x) = \frac{\beta x^{(\beta-1)}}{\theta^\beta}, \quad x > 0$$

### Experimental error analysis

Experimental errors are those that occur as a result of differences in readings between the true value and the measured value of a variable. Errors occur as a result of human imperfections when taking measurements and also from errors in the instruments that are used to take such measurements.

The term *random error* is used for those errors that vary in a random way, between successive readings of the same quantity. It is random errors that can be analysed and improved through statistical analysis and it is this type of error that we concentrate on here.

*Systematic errors*, on the other hand, are those that do not vary from one reading to another but result from an incorrect initial setting of say, an instrument. An example of this type of error is an incorrectly zeroed micrometer, that is subsequently used to check a measurement repeatedly, that does not vary. Systematic errors can be eliminated by ensuring the accuracy of instruments being used and adopting common practice when taking measurements.

Most measurements are given a tolerance, that is, a dimension that may fall either side of the exact value required. So for example a batch of 10 mm diameter bolts may be given a tolerance of  $\pm 5 \mu\text{m}$  or  $10 \pm 0.005 \text{ mm}$ . This tolerance or error will be dependent on the fit required from the bolt and the accuracy with which the measurement can be taken.

#### Example 6.4.22

A factory manufactures close tolerance bolts for use in the motor industry because the manufacturing process is not perfect not all bolts are identical. Measurement of the diameter of 10 bolts selected at random from a large batch, gave the following results:

|                    |       |       |      |      |      |      |      |       |       |       |
|--------------------|-------|-------|------|------|------|------|------|-------|-------|-------|
| Bolt diameter (mm) | 10.02 | 10.05 | 9.98 | 9.96 | 10.1 | 9.95 | 9.94 | 10.11 | 10.04 | 10.01 |
|--------------------|-------|-------|------|------|------|------|------|-------|-------|-------|

Determine the mean diameter of the bolts, their standard deviation and the error from the mean.

The sample mean is easily determined from:

$$\bar{x} = \frac{\sum x}{n}$$

and is:

$$\bar{x} = \frac{10.02 + 10.05 + 9.98 + 9.96 + 10.1 + 9.95 + 9.94 + 10.11 + 10.04 + 10.01}{10} = 10.016.$$

The best estimate for the standard deviation may be calculated using the formulae shown below:

$$\sigma^2 \approx \left( \frac{n}{n-1} \right) s^2 \quad \text{and} \quad s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{or}$$

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}.$$

Then

| $x - \bar{x}$                       | $(x - \bar{x})^2$ |
|-------------------------------------|-------------------|
| 0.004                               | 0.000 016         |
| 0.034                               | 0.001 156         |
| -0.036                              | 0.001 296         |
| -0.056                              | 0.003 136         |
| 0.084                               | 0.007 056         |
| -0.066                              | 0.004 356         |
| -0.076                              | 0.005 776         |
| 0.094                               | 0.008 836         |
| 0.024                               | 0.00 576          |
| -0.006                              | 0.000 036         |
| $\Sigma(x - \bar{x})^2 = 0.037 424$ |                   |

$$\text{So } s = \sqrt{\frac{0.037 424}{10}} = 0.061 175 \quad \text{and} \quad s^2 = 0.0037 424.$$

$$\text{Therefore } \sigma = \sqrt{\left(\frac{10}{9}\right) 0.0037 424} = 0.06 448 \text{ mm.}$$

$$\text{Then the standard error is } \sigma_m = \frac{\sigma}{\sqrt{n}} = \frac{0.06 448}{\sqrt{10}} = 0.0204 \text{ mm.}$$

Thus the bolts have nominal diameter of **10.016 ± 0.0204 mm**.

### Example 6.4.23

In an experiment to verify Hooke's law the following readings were recorded.

|                   |   |       |       |       |       |       |       |       |       |       |
|-------------------|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Load<br>(kN)      | 0 | 10    | 20    | 30    | 40    | 45    | 47.5  | 50    | 51.25 | 52.5  |
| Extension<br>(mm) | 0 | 0.070 | 0.140 | 0.210 | 0.280 | 0.315 | 0.333 | 0.350 | 0.361 | 0.373 |

For the above data:

- (i) Find the line of best fit using linear regression.
  - (ii) Find the correlation coefficient for load against effort and comment on the result.
- (i) As we are trying to verify Hooke's law, it is reasonable to assume that a linear relationship exists between the load and the extension. Therefore, our regression line will take the form,  $L = aE + b$ . So setting up a table of values in the normal way yields the following results:

| Extension<br>(mm) | Load<br>(kN) | $E^2$  | EL  | $L^2$ |
|-------------------|--------------|--------|-----|-------|
| 0                 | 0            | 0      | 0   | 0     |
| 0.070             | 10           | 0.0049 | 0.7 | 100   |

| Extension<br>(mm)  | Load<br>(kN)           | $E^2$                       | EL                          | $L^2$                        |
|--------------------|------------------------|-----------------------------|-----------------------------|------------------------------|
| 0.140              | 20                     | 0.0196                      | 2.8                         | 400                          |
| 0.210              | 30                     | 0.0441                      | 6.3                         | 900                          |
| 0.280              | 40                     | 0.0784                      | 11.2                        | 1600                         |
| 0.315              | 45                     | 0.099 225                   | 14.175                      | 2025                         |
| 0.333              | 47.5                   | 0.110 889                   | 15.8175                     | 2256.25                      |
| 0.350              | 50                     | 0.1225                      | 17.5                        | 2500                         |
| 0.361              | 51.25                  | 0.130 321                   | 18.50 125                   | 2626.5625                    |
| 0.373              | 52.5                   | 0.139 129                   | 19.5825                     | 2756.25                      |
| $\Sigma E = 2.432$ | $\Sigma L =$<br>346.25 | $\Sigma E^2 =$<br>0.749 064 | $\Sigma EL =$<br>106.57 625 | $\Sigma L^2 =$<br>15164.0625 |

Now the coefficients of the regression line are given by:

$$a = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2} \quad \text{and} \quad b = \bar{y} - a \bar{x},$$

where in our case

$$a = \frac{\sum E_i L_i - n \bar{E} \bar{L}}{\sum E_i^2 - n \bar{E}^2}, \quad b = \bar{L} - a \bar{E}, \quad \bar{E} = 0.2432 \text{ and } \bar{L} = 34.625.$$

$$\begin{aligned} \text{Then } a &= \frac{106.57625 - [(10)(0.2432)(34.625)]}{0.749064 - [(10)(0.2432)^2]} \\ &= \frac{22.36825}{0.1576016} = 141.93 \end{aligned}$$

$$\text{and } b = 34.625 - (141.93)(0.2432) = 0.1.$$

Therefore, the line of best fit is  **$L = 141.93E + 0.1$** .

(ii) The correlation coefficient is given by:

$$r = \frac{(\sum x_i y_i - n \bar{x} \bar{y})}{\sqrt{(\sum x_i^2 - n \bar{x}^2)(\sum y_i^2 - n \bar{y}^2)}} \quad \text{or in our case,}$$

$$\begin{aligned} r &= \frac{(\sum E_i L_i - n \bar{E} \bar{L})}{\sqrt{(\sum E_i^2 - n \bar{E}^2)(\sum L_i^2 - n \bar{L}^2)}} \\ &= \frac{[(106.57625) - (10)(0.2432)(34.625)]}{\sqrt{[(0.749064) - (10)(0.2432)^2] [(15164.0625) - (10)(34.625)^2]}} \\ &= \frac{22.368}{22.369} = 1.0. \end{aligned}$$

Thus the coefficient of correlation  **$r = 1.0$**  meaning that all the points lie on the line of best fit, which is a straight line. This of course, is what we would expect!

### Combined errors

When we carry out an experiment that involves taking the measurement of *two or more quantities* and then combining these quantities, it is important to be able to find a *combined error* that relates to the errors in all the individual quantities, that go to make up the combination. For example, if we are measuring *mass* and *acceleration*, in order to determine *force*. Then we need to be able to combine the errors in the measurement of the mass ( $\Delta m$ ) and the acceleration ( $\Delta a$ ) in order to find the combined error ( $\Delta F$ ) in the force. How do we combine the errors? Since  $F = ma$ , is the combined error for the force  $\Delta F = \Delta m \Delta a$ ? The answer is *no*! We must consider first, the fact that the error is measured from the *mean* value. Therefore, we would be looking at the following situation:  $\bar{F} + \Delta F = (\bar{m} + \Delta m)(\bar{a} + \Delta a)$ . Now on expansion of the RHS we get:

$$\bar{F} + \Delta F = (\bar{m}\bar{a} + \bar{a}\Delta m + \bar{m}\Delta a + \Delta a\Delta m)$$

Now, if the errors in the mass and in the acceleration are small *their product would be very small*, so on this basis as an approximation we will ignore  $\Delta a\Delta m$ . So we are left with,

$$\bar{F} + \Delta F = (\bar{m}\bar{a} + \bar{a}\Delta m + \bar{m}\Delta a).$$

Now without error it is also true that

$$\bar{F} = \bar{m}\bar{a}$$

therefore, we may eliminate this relationship from our argument to leave:

$$\Delta F = \bar{a}\Delta m + \bar{m}\Delta a,$$

as the relationship *between the errors*. We can now manipulate this formulae to obtain:

$$\frac{\Delta F}{\bar{F}} = \frac{\bar{a}\Delta m + \bar{m}\Delta a}{\bar{a}\bar{m}} = \frac{\Delta m}{\bar{m}} + \frac{\Delta a}{\bar{a}} \quad (1)$$

In other words the fractional error in the force is equal to the fractional error in the mass plus the fractional error in the acceleration.

Now Equation (1) tells us about the *standard error from the mean*. That is, the chance of being *one standard deviation* about the mean. If we assume that the errors are *normally distributed* then, from the table for  $A(x)$ , at  $x = 1.0$  we see that  $A(x) = 0.8413$  and for  $x = 0$ ,  $A(x) = 0.5$ . Therefore, for one standard deviation (standard error)  $A(x) = 0.8413 - 0.5 = 0.3413$  and since we have with or without errors, then by symmetry, the standard error about the mean has  $A(x) = \pm (0.3413 + 0.3413) = 0.6826$  or about 68% chance that the mean value for the mass and acceleration will be within that amount of the true mean. Therefore by considering all the measurements and their standard errors used to obtain the mean values of the mass and acceleration, then from previous theory we can write:

$$\left(\frac{\Delta F}{\bar{F}}\right)^2 = \left(\frac{\Delta m}{\bar{m}} + \frac{\Delta a}{\bar{a}}\right)^2 = \left(\frac{\Delta m}{\bar{m}}\right)^2 + \left(\frac{\Delta a}{\bar{a}}\right)^2 + \left(\frac{\Delta m}{\bar{m}}\right)\left(\frac{\Delta a}{\bar{a}}\right)$$

Now when we consider all possible measurements, some errors will be positive with respect to the mean and some will be negative.

Therefore the non-squared term in the above expression, that is:

$$\left(\frac{\Delta m}{\bar{m}}\right)\left(\frac{\Delta a}{\bar{a}}\right)$$

can take  $\pm$  values, and over a large set of measurements these will tend to cancel one another out, so this term can be ignored and we are finally left with the equation we need to use to find the error in the force. Which is,

$$\left(\frac{\Delta F}{\bar{F}}\right)^2 = \left(\frac{\Delta m}{\bar{m}}\right)^2 + \left(\frac{\Delta a}{\bar{a}}\right)^2.$$

This equation can be generalised for the *errors in any two sets of measurements that need to be combined by multiplication*, to give:

$$\left(\frac{\Delta Z}{\bar{Z}}\right)^2 = \left(\frac{\Delta X}{\bar{X}}\right)^2 + \left(\frac{\Delta Y}{\bar{Y}}\right)^2$$

Now you will be pleased to know that we will not go through the proof of the formulae needed to combine the errors in two set of measurements, combined by addition/subtraction, which yield the same formula. Nor do we need to go through the proof for errors combined by division because they also yield the same result as for multiplication. So we will just list the equations for these two types of combination! You will also be pleased to know that these equations are much easier to use, than to verify!

Then, *to combine errors from two independent variables*, we may use the following formulae.

When combined by *addition or subtraction*,

$$(\Delta Z)^2 = (\Delta X)^2 + (\Delta Y)^2.$$

When combined by *multiplication or division*,

$$\left(\frac{\Delta Z}{\bar{Z}}\right)^2 = \left(\frac{\Delta X}{\bar{X}}\right)^2 + \left(\frac{\Delta Y}{\bar{Y}}\right)^2.$$

Finally, before we consider an example, if the combined variable  $Z$  is connected to an independent variable that is raised to a power, that is, by the law,  $Z = X^n$ , then the error involving a power, may be derived from the above rule for multiplication, as:

$$\left(\frac{\Delta Z}{\bar{Z}}\right)^2 = n \left(\frac{\Delta X}{\bar{X}}\right)^2.$$

### Example 6.4.24

Determine the mean value and error for the variable  $Z$  when:

- (i)  $Z = X - Y$ ; (ii)  $Z = XY$ ; (iii)  $Z = Y^3$   
given that  $X = 80 \pm 2.5$  and  $Y = 30 \pm 1.5$ .

Then for

$$\begin{aligned} \text{(i) } (\Delta Z)^2 &= (\Delta X)^2 + (\Delta Y)^2 = (\pm 2.5)^2 + (\pm 1.5)^2 \\ &= 6.25 + 2.25 = 8.5 \end{aligned}$$

Therefore the mean value is  $\bar{Z} = 50$  and the error  $\Delta Z = \pm 2.91$ .

- (ii) The mean value is  $\bar{Z} = 2400$  and the fractional error is given by

$$\left(\frac{\Delta Z}{\bar{Z}}\right)^2 = \left(\frac{\Delta X}{\bar{X}}\right)^2 + \left(\frac{\Delta Y}{\bar{Y}}\right)^2$$

$$\left(\frac{\Delta Z}{2400}\right)^2 = \left(\frac{2.5}{80}\right)^2 + \left(\frac{1.5}{30}\right)^2 = (0.03125)^2 + (0.05)^2$$

$$\left(\frac{\Delta Z}{2400}\right)^2 = (0.03125)^2 + (0.05)^2 = 3.4765 \times 10^{-3}$$

$$(\Delta Z)^2 = 2400^2(3.4765 \times 10^{-3}) = 20025 \quad \text{and}$$

the error  $\Delta Z = \pm 141.5$

- (iii) Then the mean value is  $\bar{Z} = 27000$  and the fractional error is given by

$$\left(\frac{\Delta Z}{\bar{Z}}\right)^2 = n\left(\frac{\Delta Y}{\bar{Y}}\right)^2 \quad \text{or} \quad \left(\frac{\Delta Z}{27000}\right)^2 = \left(\frac{0.15}{30}\right)^2 = 2.5 \times 10^{-5}$$

so,  $(\Delta Z)^2 = 27000^2(2.5 \times 10^{-5})$  and so the error  $\Delta Z = 135$ .

### Generalised Gaussian law

Now the value of the variable  $Z$  above, is dependent on the values of the variables  $X$  and  $Y$ , and their errors. To *generalise even further*, if we assume that the value  $Z$  is some physical quantity that cannot be measured directly but is *any function*  $f$  of the two measurable quantities  $X$  and  $Y$ , then we can measure  $X$  and  $Y$  and obtain their mean values  $(\bar{x}, \bar{y})$  and standard deviations from the mean  $(\sigma_{mx}, \sigma_{my})$ , then use these values to obtain the standard deviation  $(\sigma_{mz})$  for the quantity  $Z$ , where  $z = f(x, y)$ . To do this we use the *Gaussian law of error propagation*, which may be derived from the error formulae given above, but is given here without proof as:

$$\sigma_{mz} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 (\sigma_{mx})^2 + \left(\frac{\partial f}{\partial y}\right)^2 (\sigma_{my})^2}$$

where  $\partial f / \partial x$  and  $\partial f / \partial y$  are the *partial derivatives* of the function  $f$  calculated at the mean values  $\bar{x}$  and  $\bar{y}$ . You may not have met partial derivatives before, we use them where the function involves two or more variables. So to find  $\partial f / \partial x$  we differentiate the function  $f$ , with respect to  $x$ , treating the variable  $y$  as a constant. Similarly to find  $\partial f / \partial y$  we differentiate the function  $f$  with respect to  $y$ , treating the variable  $x$  as a constant. An example of the use of this law, should help you understand this process.



**Example 6.4.25**

A current of  $5.0 \pm 0.02$  A, flows through a resistor of  $10 \pm 0.05 \Omega$ , for exactly 10 min, What is the *standard deviation* of the heat energy dissipated? As far as the science is concerned, all that you need to know is that power,  $P = I^2 R$  in watts and that the heat energy generated from the resistor is given by  $Ps$ , where  $s$  is the time in seconds. Therefore, we need to find the mean value of the heat energy evolved and the standard deviation of the power multiplied by the exact time.

Mean value of power =  $5^2 \cdot 10 = 250$  W, therefore, the mean value of the heat energy =  $(250)(10)(60) = 150\,000$  J or **150 kJ**.

Now we need to determine the standard deviation of the power, where  $\sigma_{mI} = 0.02$  A and  $\sigma_{mR} = 0.05 \Omega$ . Then we need to find from,  $P = I^2 R$  the partial derivatives  $\frac{\partial P}{\partial I}$  and  $\frac{\partial P}{\partial R}$ .

That is:

$$\frac{\partial P}{\partial I} = 2IR = (2)(5)(10) = 100 \quad \text{and} \quad \frac{\partial P}{\partial R} = I^2 = 5^2 = 25.$$

Then from the law by Gauss,

$$\begin{aligned} \sigma_{mP} &= \sqrt{(100)^2(0.02)^2 + (25)^2(0.05)^2} \\ &= \sqrt{4 + 1.5625} = \pm 2.36 \text{ W.} \end{aligned}$$

Therefore the standard deviation of the heat energy generated =  $\pm(2.36)(10)(60) = \pm 1416$  J, because we are multiplying by an *exact time*.

So finally the heat energy generated =  **$150 \pm 1.416$  kJ**.

**Sampling and quality control**

Statistics quality control, rather than quality assurance has been used as a tool in mass production, for many years. Most people will be aware of the fact that when mass producing engineering artefacts a close-eye has to be kept on the control of dimensions. In particular, ensuring that the product is manufactured within certain dimensional limits, known as *tolerance limits*. Thus a dimension stated as  $10 \pm 0.05$  cm, has tolerance limits of 9.95 and 10.05 cm. It is often possible, by random sampling, to ensure that total output is within the tolerance limits. This application of statistics to industrial and manufacturing processes is known as *quality control*.

A second type of quality control is an application of the *Poisson distribution*. It is used when articles being manufactured are not classified by dimensions but deemed either serviceable or defective (acceptable or unacceptable) according to whether they pass or fail a certain quality control test. This type of quality test is known as *control of the fraction defective*.

Before we consider the importance of *random sampling*. We will first use the formulae, given at the beginning of this section, to estimate the *standard error of the difference between the means* for both *small* and *large* samples.

The use of these formula being best illustrated by example.

**Example 6.4.26**

A company which manufactures silver-coated jewellery, suspects that the silver coating is being applied more liberally by the night shift than by the day shift. To keep down cost, the silver must be as thin as possible but must be nowhere less than 0.2 mm, in order to guarantee the quality and aesthetics of the jewellery. Part of the results of an investigation are tabulated below, where 100 measurements for thickness of silver coating, were taken for both shifts.

|                         | <i>Day shift (100 measurements)</i> | <i>Night shift (100 measurements)</i> |
|-------------------------|-------------------------------------|---------------------------------------|
| Mean thickness (mm)     | 0.292                               | 0.298                                 |
| Standard deviation (mm) | 0.021                               | 0.019                                 |

Determine whether the difference between the means is significant or not and make comments on the implications of the given figures and results.

Thus, for the estimate of the standard error of the difference between the means of *two large samples*, we use:

$$\frac{|m_1 - m_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} > 1.96, \quad \text{for significance at the 5\% level.}$$

Then using the above,

$$\frac{|0.292 - 0.298|}{\sqrt{\frac{0.021^2}{100} + \frac{0.019^2}{100}}} = \frac{|-0.006|}{0.00283} = 2.12.$$

Thus there is a significant difference between the means at the 5% level in other words there is less than 1 in 20 chance of such a large difference in the depth of the silver coating. Also, from the standard deviations we know that everywhere, the depth of the silver coating was never less than the minimum 0.2 mm.

Let us consider an example where we are required to use the *t-test*, for significance between the means. That is, we find a value for *t* based on the data and then we find the percentage probability values, from the *t-distribution*, which is used for a *small* number of samples. We use,

$$t = \frac{|m_1 - m_2|}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

The percentage probability values ( $P/100$ ) are often tabulated for  $|t| \geq t_p$ , as shown below. Where  $t$  is the ratio of the random variable normally distributed with zero mean, to the independent estimate of its standard deviation based on  $\nu$  degrees of freedom (where  $\nu = n_1 + n_2 - 2$ ).

Table showing some percentage points of the  $t$ -distribution

| $P$       | 25   | 10   | 5     | 2     | 1     | 0.2   | 0.1   |
|-----------|------|------|-------|-------|-------|-------|-------|
| $\nu = 1$ | 2.41 | 6.31 | 12.71 | 31.82 | 63.66 | 318.3 | 636.6 |
| 2         | 1.60 | 2.92 | 4.30  | 6.96  | 9.92  | 22.33 | 31.60 |
| 3         | 1.42 | 2.35 | 3.18  | 4.54  | 5.84  | 10.21 | 12.92 |
| 4         | 1.34 | 2.13 | 2.78  | 3.75  | 4.60  | 7.17  | 8.61  |
| 5         | 1.30 | 2.02 | 2.57  | 3.36  | 4.03  | 5.89  | 6.87  |
| 6         | 1.27 | 1.94 | 2.45  | 3.14  | 3.71  | 5.21  | 5.96  |
| 7         | 1.25 | 1.89 | 2.36  | 3.00  | 3.50  | 4.79  | 5.41  |
| 8         | 1.24 | 1.86 | 2.31  | 2.90  | 3.36  | 4.50  | 5.04  |
| 9         | 1.23 | 1.83 | 2.26  | 2.82  | 3.25  | 4.30  | 4.78  |
| 10        | 1.22 | 1.81 | 2.23  | 2.76  | 3.17  | 4.14  | 4.59  |
| 12        | 1.21 | 1.78 | 2.18  | 2.68  | 3.05  | 3.93  | 4.32  |
| 15        | 1.18 | 1.75 | 2.13  | 2.60  | 2.95  | 3.73  | 4.07  |
| 20        | 1.18 | 1.72 | 2.09  | 2.53  | 2.85  | 3.55  | 3.85  |
| 24        | 1.17 | 1.71 | 2.06  | 2.49  | 2.80  | 3.47  | 3.75  |
| 30        | 1.17 | 1.70 | 2.04  | 2.46  | 2.75  | 3.39  | 3.65  |
| 40        | 1.17 | 1.68 | 2.02  | 2.42  | 2.70  | 3.31  | 3.55  |
| 60        | 1.16 | 1.67 | 2.00  | 2.39  | 2.66  | 3.23  | 3.46  |
| 120       | 1.16 | 1.66 | 1.98  | 2.36  | 2.62  | 3.16  | 3.37  |
| $\infty$  | 1.15 | 1.64 | 1.96  | 2.33  | 2.58  | 3.09  | 3.29  |

### Example 6.4.27

An aircraft towing arm shear pin, is intended to have a fracture stress of 130 kPa, with standard deviation of 0.4 kPa. The mean shear stress of nine shear pins was found to be 129.5 kPa. Is a complaint, that the towing arm shear pins are not up to specification, statistically justified?

This example is for the significance of a *single mean*, using a small number of samples. Therefore we use,

$$t = \frac{|\mu_0 - m|}{s/\sqrt{n}}$$

and the  $t$ -distribution tables given above to determine whether or not a significant difference exists between the expected mean and that of the sample. So:

$$t = \frac{|130 - 129.5|}{0.4/\sqrt{9}} = 3.75$$

and from the  $t$ -distribution tables above, where  $\nu = n - 1 = 8$  for a *single mean*, the mean of the sample differs significantly (i.e. at the 1% level) from the specified average. So the complaint that the shear pins are not up to standard is statistically justified!

The *value of random sampling*, is that it can be used to judge the quality of the total output/production run. It therefore saves both money and time. It is also used where, for example, it is necessary to carry out some form of destructive test, in order to assess the product. A fire works manufacturer would need to let-off sample fire works, in order to assess their reliability, quality, and safety characteristics. Under these circumstances sampling is the only method of assessing quality!

### Quality control charts and limits

Quality control of given dimensions is maintained by setting-up two charts, one for the *means of the random samples* and the other for their *ranges*, or *standard deviations*. You are already aware that if the estimated standard deviation of the individual observations is  $\sigma$ , the estimated standard deviation of the mean of the samples size  $n$  is  $\sigma/\sqrt{n}$ . Now, even though the individual observations are not normally distributed, it is more than likely that the means of the samples will be, therefore, we are able to estimate the probability of a sample having a mean outside the limits, with either 95% or 99.8% confidence, using:

$$\text{Mean} \pm \frac{1.96\sigma}{\sqrt{n}} \quad (95\% \text{ confidence})$$

$$\text{Mean} \pm \frac{3.09\sigma}{\sqrt{n}} \quad (99.8\% \text{ confidence}).$$

Thus in the case of the 95% confidence limits the chance of a sample having a mean outside these limits is 1 in 20 and in the later case (99.8% confidence), the chance of a mean outside these limits is 1 in 500.

### Example 6.4.28

The measurements tabulated below were taken by randomly sampling the production run of high tensile steel compression springs.

Table showing spring diameter in (mm), for the five samples.

| Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|----------|----------|----------|----------|----------|
| 10.010   | 10.002   | 10.009   | 10.005   | 10.006   |
| 10.002   | 10.007   | 10.007   | 10.011   | 10.002   |
| 10.007   | 10.008   | 10.003   | 10.002   | 10.001   |
| 10.007   | 10.005   | 10.000   | 10.008   | 10.008   |

Using this data and given that the best estimate of the standard deviation for the five samples is, determine the:

- mean of the sample means;
- best estimate of the standard deviation;
- 95% confidence limits;
- 99.8% confidence limits;
- sketch the quality control chart for the means.

(i) Then the mean of the sample means, is easily found. Using 10 mm as a datum, we get for each sample:

| Sample 1           | Sample 2           | Sample 3            | Sample 4           | Sample 5            |
|--------------------|--------------------|---------------------|--------------------|---------------------|
| 0.010              | 0.002              | 0.009               | 0.005              | 0.006               |
| 0.002              | 0.007              | 0.007               | 0.011              | 0.002               |
| 0.007              | 0.008              | 0.003               | 0.002              | 0.001               |
| 0.007              | 0.005              | 0.000               | 0.008              | 0.008               |
| $\Sigma x = 0.026$ | $\Sigma x = 0.022$ | $\Sigma x = 0.019$  | $\Sigma x = 0.026$ | $\Sigma x = 0.0017$ |
| $\bar{x} = 0.0065$ | $\bar{x} = 0.0055$ | $\bar{x} = 0.00475$ | $\bar{x} = 0.0065$ | $\bar{x} = 0.00425$ |

Then

$$\mu = 10 \pm \frac{0.0065 + 0.0055 + 0.00475 + 0.0065 + 0.00425}{5}$$

$$= 10 \pm 0.0055.$$

(ii) For our best estimate of the standard deviation from the mean of the sample means, we will use the formula

$$s^2 = \frac{1}{n} \sum (x - \bar{x})^2$$

and by considering all the readings of the samples *collectively*, we will then adjust for our best estimate (see Example 6.4.15).

We first produce the necessary table of values, in thousandths of a millimetre. Where  $\bar{x} = 5.5$ .

| $x$                           | $x - \bar{x}$ | $(x - \bar{x})^2$ | $x$ | $x - \bar{x}$ | $(x - \bar{x})^2$ |
|-------------------------------|---------------|-------------------|-----|---------------|-------------------|
| 10                            | 4.5           | 20.25             | 5   | -0.5          | 0.25              |
| 2                             | -3.5          | 12.25             | 11  | 5.5           | 30.25             |
| 7                             | 1.5           | 2.25              | 2   | -3.5          | 12.25             |
| 7                             | 1.5           | 2.25              | 8   | 2.5           | 6.25              |
| 2                             | -3.5          | 12.25             | 6   | 0.5           | 0.25              |
| 7                             | 1.5           | 2.25              | 2   | -3.5          | 12.25             |
| 8                             | 2.5           | 6.25              | 1   | -4.5          | 20.25             |
| 5                             | -0.5          | 0.25              | 8   | 2.5           | 6.25              |
| 9                             | 3.5           | 12.25             |     |               |                   |
| 7                             | 1.5           | 2.25              |     |               |                   |
| 3                             | -2.5          | 6.25              |     |               |                   |
| 0                             | -5.5          | 30.25             |     |               |                   |
| $\Sigma(x - \bar{x})^2 = 197$ |               |                   |     |               |                   |

Then

$$s^2 = \frac{1}{n} \sum (x - \bar{x})^2 \quad \text{or} \quad s^2 = 197/20 = 9.85 \quad \text{and the best}$$

$$\text{estimate of the variance } \sigma^2 = s^2 \frac{n}{n-1} = 9.85 \left( \frac{20}{19} \right) = 10.3684$$

thousandths, and so the best estimate for the *standard deviation* from the mean of the sample means is  $\sigma_m = 0.00322 \text{ mm}$ .

You will note that the *calculation method* using the sum of the squares is quite long winded! Normally in industrial quality control the standard deviation from the mean of the sample means is estimated from the *average range* of the samples and then *conversion tables\** are consulted to find the required standard deviation.

(iii) In order to find the 95% confidence limits for a sample having a mean outside these limits, we use:

$$\text{mean} \pm \frac{1.96\sigma}{\sqrt{n}},$$

where  $n$  refers to the number of observations in the sample.

$$\begin{aligned}\text{Then the 95\% zone} &= 10.005 \pm \frac{1.96(0.00322)}{\sqrt{4}} \\ &= 10.005 \pm 0.00315 \text{ mm.}\end{aligned}$$

That is between **10.002** and **10.008 mm** to the nearest thousandth of a millimetre.

(iv) We adopt the same procedure as in part (iii), except we use the 99.8% confidence limits. So using,

$$\text{mean} \pm \frac{3.09\sigma}{\sqrt{n}},$$

we get:

$$10.005 \pm \frac{3.09(0.00322)}{\sqrt{4}} = 10.005 \pm 0.0049749 \text{ mm.}$$

That is between **10.000** and **10.01 mm** to the nearest thousandth of a millimetre.

(v) The quality control chart is shown in Figure 6.4.15. Providing only 1 mean out of every 20, falls outside the inner

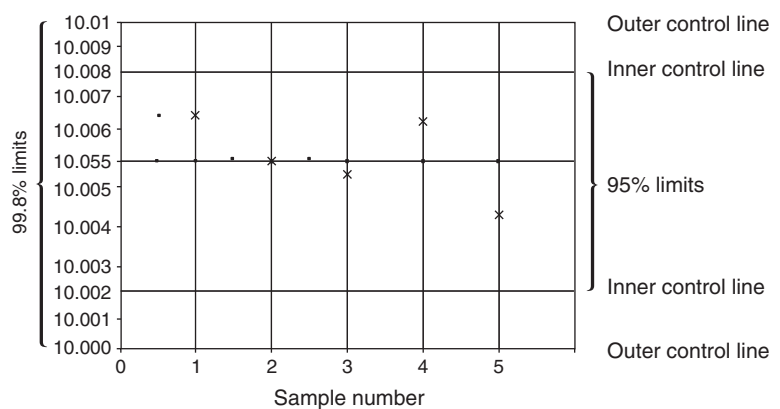


**Figure 6.4.15** Figure example 6.4.28 – A quality control chart for the means, showing production under control

control line (95% confidence interval) and only 1 mean in 500 outside the outer control line (99.8% confidence interval), then production is under control. *Immediately* there is any increase in the proportion of means falling outside of these limits, then production is out of control.

Only five small samples were considered here, therefore *it would be advisable to make further estimates of the mean and standard deviation* using more samples and adjust the control limits accordingly.

*\*Note:* In answering part (ii) of the above example, to find the *range*. All that is required is to look at each sample and find the difference between the highest and lowest observation. Thus in sample 1 the observations range from 12 thousandths to +10 thousandths, thus the range is 8 thousandths of a mm. Similarly for the other samples we get ranges of 6, 9, 9, and 7. Thus the mean of the ranges is:  $\frac{1}{5}(8 + 6 + 9 + 9 + 7) = 7.8$  thousandths.

Conversion tables, such as in the Cambridge Elementary Statistics Tables, or others, are then used to find the conversion for the sample deviations when  $n = 4$ . In this case the conversion factor is 0.4856, which is multiplied by 0.0078 to give an estimated standard deviation of 0.00378 mm. This is fractionally larger than our calculated value.

Thus it is often the case that a *quality control chart for ranges* is set up and through use of conversion tables the standard deviations from the mean of the sample means can be monitored and controlled. We will stick with our *calculated values* for determining the standard deviation.

If we continue our argument for our spring diameters and wish to keep an eye on the quality of the production process. We would take samples at varying times after the start of production and check that the *mean and standard deviation between samples* is not too large.

### Example 6.4.29

On the fourth day of the large production run for the compression springs (Example 6.4.28), samples are again taken. The observed results are tabulated below.

Day 4 of production run. Table showing observed measurements from nominal 10 mm diameter of compression springs

| Sample 6 | Sample 7 | Sample 8 | Sample 9 | Sample 10 |
|----------|----------|----------|----------|-----------|
| 0.014    | 0.012    | 0.009    | 0.015    | 0.016     |
| 0.001    | 0.017    | 0.007    | 0.011    | 0.002     |
| 0.005    | 0.008    | 0.013    | 0.002    | 0.001     |
| 0.012    | 0.003    | 0.003    | 0.008    | 0.009     |

Determine whether or not production is under control, (i) at the 95% limits; (ii) at the 99.8% limits.

All that we are required to do is to determine the sample means and compare them with the mean range shown on the control chart for (Example 6.4.28).

The table of observed values is shown again with the calculated sample means added.

| Sample 6           | Sample 7          | Sample 8          | Sample 9           | Sample 10          |
|--------------------|-------------------|-------------------|--------------------|--------------------|
| 0.014              | 0.012             | 0.009             | 0.015              | 0.016              |
| 0.001              | 0.017             | 0.007             | 0.011              | 0.002              |
| 0.005              | 0.008             | 0.013             | 0.002              | 0.001              |
| 0.012              | 0.003             | 0.003             | 0.008              | 0.009              |
| $\Sigma x = 0.032$ | $\Sigma x = 0.04$ | 0.032             | $\Sigma x = 0.036$ | $\Sigma x = 0.028$ |
| $\bar{x} = 0.008$  | $\bar{x} = 0.01$  | $\bar{x} = 0.008$ | $\bar{x} = 0.009$  | $\bar{x} = 0.007$  |

- (i) From the calculated mean values, then at the 95% limits if we look at the original control chart we see that the limits for the mean were 10.002–10.008 mm, this shows that sample 9 above has a mean value that falls outside these limits, therefore, at the 95% limits, *production is out of control*.
- (ii) Again at the 99.8% limits, then the mean zone on our control chart was 10.000–10.008 mm, so once more one sample in five is outside these limits and *production is out of control*.

Now, our original control chart was set up with just five samples being taken and it would have been prudent to carry out more sampling at the beginning of the process, in order to obtain a more accurate and realistic mean range and standard deviation. Under these circumstances, our second set of samples may have been within limits.

### Example 6.4.30

Suppose the criteria for meeting the customer specification was that the dimensional accuracy of the diameter of the compression springs had to be within  $10 \pm 0.01$  mm. Would the production run carried out on day 4, meet the customers quality criteria?

To answer this question, we need to determine new values for the mean of the sample means and the corresponding mean range. Then, from Example 4.6.29, the mean of the sample means is:

$$\begin{aligned}\mu &= 10 \pm \frac{0.008 + 0.01 + 0.008 + 0.009 + 0.007}{5} \\ &= 10 \pm 0.0084.\end{aligned}$$

Then setting up an appropriate table of values, as before, we get:

| x  | $x - \bar{x}$ | $(x - \bar{x})^2$ | x  | $x - \bar{x}$ | $(x - \bar{x})^2$ |
|----|---------------|-------------------|----|---------------|-------------------|
| 14 | 5.6           | 31.36             | 13 | 4.6           | 21.16             |
| 1  | -7.4          | 54.76             | 3  | -5.4          | 29.16             |
| 5  | -3.4          | 11.56             | 15 | 6.6           | 43.56             |



| $x$ | $x - \bar{x}$ | $(x - \bar{x})^2$ | $x$ | $x - \bar{x}$ | $(x - \bar{x})^2$               |
|-----|---------------|-------------------|-----|---------------|---------------------------------|
| 12  | 3.6           | 12.96             | 11  | 2.6           | 6.76                            |
| 12  | 3.6           | 12.96             | 2   | -6.4          | 40.96                           |
| 17  | 8.6           | 73.96             | 8   | -0.4          | 0.16                            |
| 8   | -0.4          | 0.16              | 16  | 7.6           | 57.76                           |
| 3   | -5.4          | 29.16             | 2   | -6.4          | 40.96                           |
| 9   | 0.6           | 0.36              | 1   | -7.4          | 54.76                           |
| 7   | -1.4          | 1.96              | 9   | 0.6           | 0.36                            |
|     |               |                   |     |               | $\Sigma(x - \bar{x})^2 = 524.8$ |

Therefore

$$s^2 = \frac{1}{n} \Sigma (x - \bar{x})^2 \text{ or } s^2 = \frac{524.8}{20} = 26.24 \text{ and as before:}$$

$$\sigma^2 = s^2 \frac{n}{n-1} = 26.24 \left( \frac{20}{19} \right) = 27.62 \text{ and}$$

$$\sigma = 0.005256 \text{ mm.}$$

Then at the 95% zone, the *mean range* is given by

$$\begin{aligned} \text{mean} \pm \frac{1.96\sigma}{\sqrt{n}} \text{ or} \\ 10.0084 \pm \frac{(1.96)(0.005256)}{\sqrt{4}} = 10.0084 \pm 0.0051510. \end{aligned}$$

Therefore the mean range lies between **10.003** and **10.014 mm**, to the nearest thousandth of a millimetre. Now this is the more stringent of the criteria in that no more than 1 in 20 sample means should appear outside this band. The individual sample means *for production on that day*, are all within the range 10.007–10.01, well within the mean range limits, *so production is under control and meets customer requirements*.

However, production would have to be sampled on a regular basis, to ensure that customer requirements continued to be met. It would also be advisable to consider machine adjustment at an early stage, to tighten the upper control limit of the mean range.

We finish this short study of quality control by considering the use of the Poisson distribution for fractional defectives.

### **The Poisson distribution and fractional defectives**

As mentioned earlier the Poisson distribution may be used to determine the control limit for fractional defectives. What this means is controlling the number of defective items, that may appear in a production run, by considering the number of defectives there are in samples and seeing whether or not such samples are within the limits of the control chart.

**Example 6.4.31**

The table given below shows the number of missing rivets noted at final inspection on 30 new aircraft.

|    |    |    |    |    |    |
|----|----|----|----|----|----|
| 10 | 14 | 21 | 15 | 12 | 16 |
| 17 | 7  | 13 | 11 | 18 | 7  |
| 16 | 14 | 10 | 12 | 6  | 8  |
| 20 | 19 | 11 | 8  | 10 | 10 |
| 10 | 16 | 23 | 26 | 8  | 12 |

Establish the outer control limits for the quality chart in order that the riveting process for the aircraft is in statistical quality control.

The data which gives the number of defectives per sample in a finite population, follows the *Poisson distribution*.

In your previous study of probability distributions, you will have noted that,  $np = a = \text{mean} = \text{variance}$ , so that the standard deviation for data that follows the Poisson distribution  $= \sqrt{a}$ .

Then from the above data, the mean

$$a = \frac{\sum x}{n} = \frac{400}{30} = 13.33$$

and the standard deviation

$$\sqrt{a} = 3.65.$$

Now the outer control limits, that is, the limits for the 99.8% zone are given by

$$\begin{aligned} \text{Mean} \pm \frac{3.09\sigma}{\sqrt{n}} & \text{ or } 13.33 \pm \frac{(3.09)(3.65)}{\sqrt{1}} \\ & = 13.33 \pm 11.2785. \end{aligned}$$

Then the upper and lower control limits are **24.6** and **2.05**, respectively. Note that the *number of observations per sample*, is 1.0, that is, only on one occasion, final inspection, was the sample taken, hence  $\sqrt{1.0}$  is used in the above calculation, to determine the outer limits.

Now for these limits the sample value of 26 in the table falls outside these control limits, and it would be best to base future calculations on a value of  $a$ , that does not involve this particular aircraft. So mean

$$a = \frac{\sum x}{n} = \frac{374}{29} = 12.897 \quad \text{and} \quad \sqrt{a} = 3.59.$$

Once again the upper and lower control limits are given by

$$\begin{aligned} \text{Mean} \pm \frac{3.09\sigma}{\sqrt{n}} & \text{ or } 12.897 \pm \frac{(3.09)(3.59)}{\sqrt{1}} \\ & = 12.897 \pm 11.097. \end{aligned}$$

Then the upper and lower control limits are **23.99** and **1.8**, respectively. All mean values lie within these limits, *so these are appropriate control limits* for guiding inspectors with future riveting quality control.

Here is one final example that uses the Poisson distribution to determine the probabilities of defectives.

### Example 6.4.32

A random sample of 20 articles was drawn from every 100 produced by a machining process. The number of defective articles per sample for the first 30 samples drawn were: 0, 1, 0, 1, 1, 2, 3, 0, 0, 1, 2, 0, 0, 0, 2, 1, 1, 2, 1, 0, 2, 3, 2, 0, 0, 2, 1, 0, 2, 0.

- (i) Calculate the mean number of defectives per sample.
- (ii) Obtain an upper control limit ( $c_{lim}$ ), such that the probability of there being  $c_{lim}$  or more defectives in a sample is less than 10%.
- (i) This is simply 30 (the number of samples) divided by 30 (the total number of defectives in these samples, that is,  $30/30 = 1.0$ ). Then the *mean number of defectives per sample is 1.0*.
- (ii) Now using the Poisson distribution with  $a = 1.0$ , we can find the probability of 0, 1, 2, 3, ... defectives per sample, from:

$$e^{-a}, ae^{-a}, \frac{a^2}{2!}e^{-a}, \frac{a^3}{3!}e^{-a}, \dots$$

These probabilities are tabulated below:

| Number of defectives<br>per sample | Probability<br>$P(x)$                      |
|------------------------------------|--------------------------------------------|
| 0                                  | $P(0) = e^{-1} = 0.3679$                   |
| 1                                  | $P(1) = 1e^{-1} = 0.3679$                  |
| 2                                  | $P(2) = \frac{1^2}{(1)(2)}e^{-1} = 0.1839$ |
| 3 or more                          | $P(>2) = 1 - P(0) - P(1) - P(2) = 0.0803$  |

Then from the table, because the sum of the first three probabilities,  $P(0)$ ,  $P(1)$ , and  $P(2)$  is 0.9197, the probabilities of three or more defectives per sample is less than 0.1 or 10%. *Thus the required control upper limit is 3.*

For our final look at statistical applications, we consider a few examples, taken from engineering reliability.

### Engineering reliability

The reliability of engineering structures, components and systems, is of vital importance throughout the design, manufacture, maintenance, and

disposal of engineering artefacts. Reliability prediction involves both *probability theory* and the use of *probability density functions*, for continuous distributions.

The use of these statistical techniques is illustrated, by the examples that follow.

### Example 6.4.33

A missile is test fired from an aircraft. It is known that some events may prevent the missile from reaching its target. These events are: ( $A_1$ ) cloud reflection that has a probability of occurrence (within any one flight) of 1 in 10 000, ( $A_2$ ) rain with a probability of occurrence of 6 in a 1000, ( $A_3$ ) evasion by target with a probability of occurrence of 2 in 1000, and ( $A_4$ ) electronic counter measures with a probability of occurrence of 5 in 100. The corresponding probabilities of failure, if these events occur are:  $P(F/A_1) = 0.25$ ,  $P(F/A_2) = 0.02$ ,  $P(F/A_3) = 0.005$ ,  $P(F/A_4) = 0.0002$ .

Use Bayes' theorem to calculate the probability of each of these events being the cause of the missile not reaching its target.

Now this simply involves the application of Bayes' theorem to the probability that events,  $A_1$ ,  $A_2$ ,  $A_3$ , and  $A_4$ , respectively, cause the missile to fail, that is, it fails to reach its target. So we use:

$$P(A_r/F) = \frac{P(A_r)P(F/A_r)}{\sum_{r=1}^n P(A_r)P(F/A_r)}, \text{ then the denominator is}$$

$$\begin{aligned} \sum_{r=1}^n P(A_r)P(F/A_r) &= (0.0001)(0.25) + (0.006)(0.02) \\ &+ (0.002)(0.005) + (0.05)(0.0002) \\ &= (2 \times 10^{-5}) + (1.2 \times 10^{-4}) + (1 \times 10^{-5}) + (1 \times 10^{-5}) \\ &= (1.65 \times 10^{-4}). \end{aligned}$$

Then the separate probabilities are given as:

$$P(A_1/F) = \frac{2.5 \times 10^{-5}}{1.65 \times 10^{-4}} = \mathbf{0.1515}, \text{ similarly}$$

$$P(A_2/F) = \frac{1.2 \times 10^{-4}}{1.65 \times 10^{-4}} = \mathbf{0.7272}$$

$$P(A_3/F) = \frac{1 \times 10^{-5}}{1.65 \times 10^{-4}} = \mathbf{0.0606}$$

$$P(A_4/F) = \frac{1 \times 10^{-5}}{1.65 \times 10^{-4}} = \mathbf{0.0606}.$$

Quite clearly the biggest probability that the missile will fail ( $A_2/F$ ), is due to the prevailing weather conditions, interesting!

In the next example we use the binomial distribution.

### Example 6.4.34

A motor boat has twin engines that are known to be in identical condition. If the motor boat runs on both engines for 100 h, it is known from experience that the probability of failure of any one engine is 0.01. Assuming engine failures are independent, determine:

- (i) the probability of both engines failing;
- (ii) the probability of one engine surviving and one engine failing;
- (iii) the probability of both engines surviving;
- (iv) the probability of at least one engine surviving.

To solve this problem we may use the general terms of the binomial, where  $p = 0.99$  (success no engine failure) and  $q = 0.01$  (probability of engine failure), and for two engines  $n = 2$ .

$$(q + p)^n = q^n + \binom{n}{1} q^{n-1}p + \binom{n}{2} q^{n-2}p^2 + \dots + p^n$$

So, for the given values, the above relationship may be written as:

$$(q + p)^2 = (0.01)^2 + \binom{2}{1} (0.99)^{2-1}(0.01) + (0.99)^2.$$

Then tabulating these results we get:

| Engine 1 | Engine 2 | Probability    |
|----------|----------|----------------|
| Fail     | Fail     | $(0.01)^2$     |
| Fail     | Survive  | $(0.01)(0.99)$ |
| Survive  | Fail     | $(0.99)(0.01)$ |
| Survive  | Survive  | $(0.99)^2$     |

Then the probability of:

- (i) both engines failing is,  $(0.01)^2 = \mathbf{0.0001}$
- (ii) one engine surviving and one failing is,  $2(0.99)(0.01) = \mathbf{0.0198}$
- (iii) both engines surviving is,  $(0.99)^2 = \mathbf{0.9801}$
- (iv) at least one engine surviving is  $1 - \text{probability of both engines failing, that is } 1 - (0.01)^2 = \mathbf{0.9999}$ .

We now, once again, use probability theory to determine the reliability of an electrical circuit, having been given information on the reliability of its components.

**Example 6.4.35**

In the circuit shown in Figure 6.4.16, the probability of switch *A* and switch *B* being closed is 0.85, the probability of switch *C* or switch *D* being closed is 0.8.

- (i) What is the probability that a closed circuit will exist?
  - (ii) Given that a closed-circuit exists, what is the probability that switches *A* and *B*, will be closed?
- (i) Let a closed circuit = *X* and the probability that a closed circuit exists be  $P(X)$ , then for a closed circuit either *A* and *B* must be closed or *C* or *D* must be closed, minus the probability that all are closed. That is:

$$P(X) = P(AB) + P(C + D) - P(AB)P(C + D) \dots \quad (1)$$

where from our rules of probability, the probability of *A* and *B* being closed =  $P(AB) = P(A)P(B)$  and the probability of *C* or *D* =  $P(C) + P(D) - P(C \cap D)$ . Where,

$$P(AB) = (0.85)(0.85) = 0.7725 \text{ and}$$

$$P(C + D) = (0.8) + (0.8) - (0.8)(0.8) = 0.96 \text{ and}$$

$$P(AB)P(C + D) = (0.7725)(0.96) = 0.7416.$$

So the total probability for a closed circuit (from Equation (1)) is:

$$P(X) = 0.7725 + 0.96 - 0.7416 = \mathbf{0.991}.$$

$$(ii) \quad P\left(\frac{AB}{X}\right) = \frac{P\left(\frac{AB}{X}\right)}{P(X)}.$$

Now because both switch *A* and *B* must be closed to give *X*, then the

$$P\left(\frac{AB}{X}\right) = \frac{P(AB)}{P(X)} = \frac{P(A)P(B)}{P(X)}.$$

In words what we are saying is that the probability that *A* and *B* are closed giving us the completed circuit, that is:

$$\frac{\text{Probability that switch } A \text{ and switch } B \text{ are closed}}{\text{Probability that the circuit is closed}}$$

$$= \frac{P(A)P(B)}{P(X)} = \frac{0.7725}{0.991} = \mathbf{0.7795}.$$

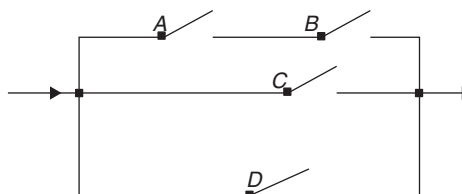


**Figure 6.4.16** Figure example 6.4.35 – Switch circuit

**The exponential distribution**

The *exponential distribution* is a *continuous distribution* that has evolved from the Poisson distribution that is discrete. Thus, if the number of failures per unit time is Poisson distributed, then the *mean time between failures* is exponentially distributed. The exponential function has a *constant failure rate*, and is used to model the *useful life* of many items. The useful life is that period between *burn-in* and *wear-out*, when failure rate is considered constant.

For example, if you buy a new washing machine, and a fault is to occur due to faulty manufacture or assembly, it is likely to happen in the first few weeks of ownership, we often refer to such faults as *teething troubles* or *burn-in*. Then there follows a long period during which white goods and other machines give trouble free service at a constant low failure rate, finally as components start to wear out we get problems associated with age and condition, these are known as *wear-out* problems.

The exponential distribution is a one parameter continuous distribution commonly expressed in terms of its mean,  $\theta$  and the inverse of its mean,  $\lambda$ . The exponential probability density function is given at the beginning of this section and is,

$$f(x) = \frac{1}{\theta} e^{-1/\theta} = \lambda e^{-\lambda x}, \quad x \geq 0$$

where  $\theta$  = the mean of the distribution, and  $\lambda = 1/\theta$  = the failure rate.

Also:

The exponential reliability function is

$$R(x) = e^{-x/\theta} = e^{-\lambda x}, \quad x \geq 0$$

and the, exponential hazard function or failure rate is,

$$h(x) = \frac{f(x)}{R(x)} = \frac{1}{\theta} = \lambda.$$

**Example 6.4.36**

- (i) What is the probability of an electrical component surviving until (time)  $t = 120$  h, if the item is exponentially distributed with a mean time between failure of 80 h?
- (ii) Given that the item survives for 240 h, what is the probability of survival until  $t = 360$  h?
- (iii) What is the value of the failure rate at 240 and 360 h?

The only problem in answering this question is to fully understand what is being asked, then the use of the distribution in all its forms, is easy to apply.

- (i) The probability of survival, is the reliability of the component, thus we may use:

$$R(x) = e^{-x/\theta} = e^{-\lambda x}$$

where  $x = 120$  h and the mean  $\theta = 85$  h.

$$\text{So, } R(100) = e^{-\left(\frac{120}{85}\right)} = 0.2437.$$

- (ii) Now the probability of survival until 360 h given that the component has survived up to 240 h, is simply given by the ratio of the reliabilities, that is:

$$\text{The } R(360/240) = \frac{R(360)}{R(240)} = \frac{e^{-360/85}}{e^{-240/85}} = 0.2437.$$

Note that this is equal to the probability of failure in the interval from  $t = 0$  to  $t = 120$ . This will obviously be the case since we are dealing with *constant failure rates*, through the components useful life.

- (iii) Now we know that the failure rate is equal to the hazard function and is constant! Therefore the failure rate is the same at 240 h as it is at 360 h and is

$$h(x) = \frac{f(x)}{R(x)} = \frac{1}{85} = \lambda = 0.01176.$$

### The Poisson distribution

You have already met this distribution and used it to model fractional defectives per item. It can also be used to determine failures per hour. In reliability work we use the *Poisson probability density function*, which up till now we have not discussed. This variant of the distribution is given below:

$$P(x, \lambda t) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}, \quad x = 0, 1, 2, 3, \dots, n$$

where  $P(x, \lambda t)$  = the probability of exactly  $x$  occurrences in the interval  $\lambda t$  and  $\lambda$  = the average occurrence rate per unit time.

### Example 6.4.37

A car manufacturer averages five minor defects per two new cars produced. What is the probability of a new car containing: (i) no defects (ii) one defect (iii) two defects, (iv) of more than three defects?

Here all that is required is to recognise the parameters contained in this variant of the Poisson distribution. Where, for *defectives per item*,  $t$  = the number of cars sampled, that is, 2 and so  $\lambda t = (5)(2) = 10$ .

Had we been dealing with failures per unit time, then  $t$  would represent time and  $\lambda$  the failure rate.

- (i) Here  $x = 0$ , so the probability of no defects in a new car is,

$$P(0, 10) = \frac{e^{-10} (10)^0}{0!} = \mathbf{0.0000454}.$$



This is not a very encouraging probability for exactly no defectives!

- (ii) Here  $x = 1$ , so the probability that a new car has exactly one defect is,

$$P(1, 10) = \frac{e^{-10}(10)^1}{1!} = \mathbf{0.0000454}.$$

- (iii) Here  $x = 2$ , so the probability of exactly two defectives is,

$$P(2, 10) = \frac{e^{-10}(10)^2}{2!} = \mathbf{0.00227}.$$

- (iv) The probability of three or more defectives, is given by  $1 - [P(0, 10) + P(1, 10) + P(2, 10)]$  and is,  $1 - [0.0000454 + 0.000454 + 0.00227] = 1 - 0.00277 = \mathbf{0.99723}$ .

Well at least this is some comfort, for the new car owner!

### The Weibull distribution

Finally, no study on reliability engineering would be complete without looking at the Weibull distribution. *The Weibull distribution is a continuous distribution* that has become widely used in the reliability field. In addition to being the most useful density function for reliability calculations. Analysis of the Weibull distribution, provides the information needed for classifying types of failure, fault finding, scheduling of preventive maintenance, and scheduling inspections. Here we look only briefly, at its use for reliability calculations.

The Weibull probability density function is,

$$f(x) = \frac{\beta x^{\beta-1}}{\theta^\beta} \exp\left[-\left(\frac{x}{\theta}\right)^\beta\right], \quad x \geq 0$$

where,  $\beta = \text{the shape parameter}$  and  $\theta = \text{the scale parameter}$ .

Beta ( $\beta$ ) and theta ( $\theta$ ) are continuous in the range from zero to infinity. *Theta is also known as the characteristic life*, because approximately two-third of the population fails by the characteristic life point irrespective of the value of  $\beta$ . By altering the shape parameter,  $\beta$ , the Weibull probability density function takes on a variety of shapes that approximate to *other density functions*. Figure 6.4.17, shows the effect of some important values of the shape parameter.

So for example, when  $\beta = 1$ , the Weibull distribution is identical to the exponential distribution and when  $\beta = 3.6$  the Weibull distribution approximates to the normal distribution.

In a similar manner to the exponential function the Weibull *reliability function* and *hazard function* are given below without proof.

$$\text{Weibull reliability function} \quad R(x) = \exp\left[-\left(\frac{x}{\theta}\right)^\beta\right], \quad x > 0 \quad \text{and}$$

$$\text{Weibull hazard function} \quad h(x) = \frac{\beta x^{(\beta-1)}}{\theta^\beta}, \quad x > 0$$

remembering that here also, the hazard rate is the failure rate.

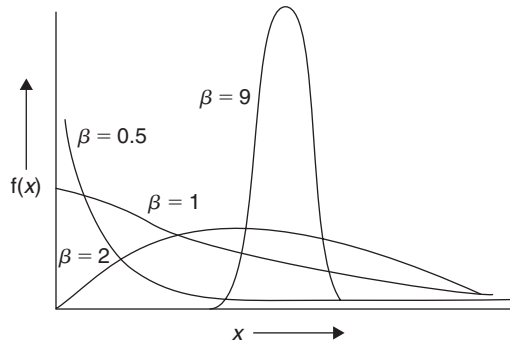


**Figure 6.4.17** The Weibull probability density function and the effect of the shape parameter  $\beta$

The Weibull probability density function and the effect of the shape parameter  $\beta$

### Example 6.4.38

- (i) What is the probability of a food processing system, remaining serviceable for 50 running hours, when it can be modelled using the Weibull distribution, that has parameters  $\beta = 2.0$  and  $\theta = 125$ ?
  - (ii) What is the instantaneous failure rate at 50 h, for the system?
- (i) Here we simply apply the Weibull reliability function,

$$R(50) = \exp \left[ - \left( \frac{50}{125} \right)^2 \right] = \mathbf{0.852}.$$

- (ii) Here all that we need do is recognise that the instantaneous *failure rate* is given by the *hazard function* as:

$$h(50) = \frac{(2.0)(50)^{(2-1)}}{125^{2.0}} = \frac{100}{15625} = \mathbf{0.0064}.$$

This brings our study of statistical applications to an end. There follows a set of questions to test your knowledge.

### Test your knowledge 6.4.4

- (1) A coil of copper wire has a wire diameter of  $0.15 \pm 0.0005$  mm and a total wire length of  $95 \pm 0.03$  m. Given that the specific resistance of copper is  $\rho = 1.7 \times 10^{-5} \Omega \text{mm}$  and  $R = \rho L/A$ . Use the law of error propagation to determine the error range in the total resistance of the coil.

- (2) The relationship between applied temperature and the measured change in volume in an experiment with a piston (assuming constant pressure) is given below:

|                           |       |       |       |       |       |     |     |       |
|---------------------------|-------|-------|-------|-------|-------|-----|-----|-------|
| Temperature (°C)          | 160   | 180   | 200   | 220   | 240   | 260 | 280 | 300   |
| Volume (cm <sup>3</sup> ) | 29.41 | 27.78 | 33.33 | 37.03 | 41.67 | 40  | 50  | 52.63 |

Determine the coefficient of correlation for these values.

- (3) For the data given below:

|   |   |    |    |    |    |    |    |
|---|---|----|----|----|----|----|----|
| X | 2 | 4  | 6  | 8  | 10 | 12 | 14 |
| Y | 5 | 11 | 15 | 19 | 24 | 28 | 33 |

- (i) Determine the equation of the regression line for Y on X correct to three significant figures.
  - (ii) Find the value of Y when X = 11.2.
- (4) Metal rods for concrete reinforcement are delivered in large batches to a supplier. When production is under control, the reinforcing rods have a mean diameter of 2.132 cm, with a standard deviation of 0.001 cm. A random sample of 90 rods from a batch is found to have a mean diameter of 2.1316 cm. Inform management as to whether production is under control or otherwise.
- (5) A random sample of 20 light bulbs was drawn from every 100 light bulbs produced by a production process. The number of defectives per sample in the first 20 samples were: 0,2,0,1,1,2,1,4,2,1,3,1,3,1,1,0,3,1,1,2. Use these samples to calculate the *mean* number of defectives per sample and then, by use of the Poisson distribution, determine the control limit (c) such that the probability of there being c or more defectives per sample is less than 0.1.
- (6) To destroy a military target, at least three missiles must hit the target. The probability that each missile hits the target is 0.9. How many missiles must be launched to ensure with 0.99 reliability, that the target is destroyed?
- (7) Two circuit breakers of the same design, each have a failure rate to open of 0.02. The circuit breakers are placed in the system, so that both must fail to open in order for the circuit breaker system to fail. What is the probability of system failure:
- (i) If the failures are independent.
  - (ii) If the probability of a second failure is 0.1, given that the first failure occurred.
  - (iii) In part (i) what is the probability of one or more circuit breaker failing on demand?
  - (iv) In part (ii) what is the probability of one or more circuit breaker failures on demand?
- (8) Three components A, B, and C each with a reliability of 0.9, 0.75, 0.8, respectively, are placed in series. What is the reliability of the system?

- (9) A flow measurement system consists of a, venturi meter ( $\lambda = 0.85$ ), a pressure transducer and signal conditioning equipment ( $\lambda = 0.9$ ) and a recording instrument ( $\lambda = 0.1$ ). Calculate (using the exponential distribution) the probability of losing the flow measurement after 1 year, for the following situations:
- a single flow measuring system;
  - four identical flow measuring systems in parallel.

*Hint:* Use the overall system failure rate to determine the reliability for both of the above cases.

- (10) (i) A system may be modelled using the Weibull time to failure distribution. What is the probability of the system operating without breakdown for 200 h, given that  $\beta = 1.6$  and  $\theta = 360$ ?
- (ii) What is the instantaneous failure rate at time 200 h?

## 6.5 ADVANCED TOPICS

In this final section we will look briefly at complex numbers, differential equations, LTs, and harmonic analysis using Fourier coefficients. All the appropriate formulae will be found together, after this introduction. These are followed, as in previous sections, with a review of the methods and then the engineering applications essential for the completion of the subject matter.

### Formulae

#### (1) Complex numbers

- $z = x + iy$  where Real  $z = x$  and Imaginary  $z = y$ ,  $i = \sqrt{-1}$ ,  $i^2 = -1$ .
- $\bar{z} = x - iy$  is the *conjugate* of the complex number  $z = x + iy$ .
- $z\bar{z} = x^2 + y^2$ .

(d) Modulus  $|z| = \sqrt{x^2 + y^2}$

(e) Distance between two points  $z_1$  and  $z_2$  is  $|z_1 - z_2| = |z_2 - z_1|$ .

(f) Polar form  $x + iy = r(\cos \theta + i \sin \theta) = r \angle \theta$   $r = |z|$  and the argument  $\theta = \arg z$   $\tan \theta = y/x$

$$z_1 z_2 = r_1 r_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)] = r_1 r_2 \angle \theta_1 + \theta_2$$

$$\frac{z_1}{z_2} = \frac{r_1 [\cos(\theta_1 - \theta_2) + i \sin(\theta_1 - \theta_2)]}{r_2} = \frac{r_1}{r_2} \angle (\theta_1 - \theta_2).$$

#### (g) Exponential form

$$z = r e^{i\theta} = \cos \theta + i \sin \theta$$

$$|e^{i\theta}| = 1$$

$$\text{conjugate of } e^{i\theta} \text{ is } e^{-i\theta}.$$

#### (h) De Moivre's theorem

If  $n$  is any integer, then

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta.$$

#### (i) Hyperbolic form

$$\cosh ix = \frac{1}{2} (e^{ix} + e^{-ix}) = \cos x$$

$$\sinh ix = \frac{1}{2} (e^{ix} - e^{-ix}) = i \sin x.$$

Note:  $i$  notation often used by mathematicians, while  $j$  is often used for applications  $i = j = \sqrt{-1}$ .

(2) *Differential equations*

(a) Linear first order:

$$\frac{dy}{dx} + P(x)y = Q(x)$$

Integrating factor  $e^{\int P(x)dx}$ .

(b) Linear second order with constant coefficients  $a, b, c$ :

$$a \frac{d^2 y}{dx^2} + b \frac{dy}{dx} + cy = f(x).$$

Auxillary equation  $an^2 + bn + c = 0$  roots  $n_1, n_2$ . Complementary function

$$Ae^{n_1 x} + Be^{n_2 x} \quad (\text{real } n_1 \neq n_2)$$

$$(Ax + B)e^{n_1 x} \quad (n_1 = n_2)$$

$$e^{px}(A \cos qx + B \sin qx) \quad (\text{complex } n_1 n_2 = p \pm iq).$$

General solution  $y$  = complementary function plus any particular solution.

$$\text{Harmonic type } \frac{d^2 y}{dx^2} + m^2 y = 0.$$

$$\text{General solution } y = A \cos mx + B \sin mx.$$

(3) *Vectors*

(a) Cartesian co-ordinates:

$$\mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j} + a_z \mathbf{k}$$

$$|\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

where  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  are unit vectors orthogonal (at-right-angles to one another).

(b) Dot product (scalar):

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta = \mathbf{b} \cdot \mathbf{a}$$

$$= a_x b_x + a_y b_y + a_z b_z$$

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b}) + (\mathbf{a} \cdot \mathbf{c}).$$

(c) Cross product (vector):

$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin \theta \mathbf{u} = \mathbf{b} \times \mathbf{a}$$

where  $\mathbf{u}$  = any unit vector.

(4) *Matrices*

(a) Matrix form:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \begin{matrix} m \times n \\ (m \text{ rows} \times n \text{ columns}). \end{matrix}$$

(b) Matrix addition (add each complementary element):

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \end{pmatrix}.$$

*Note:* subtraction carried out in a similar way, except complementary elements are subtracted instead of added.

(c) Matrix multiplication:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}.$$

Multiply each row of first matrix by each column of second matrix.

#### (5) Laplace transforms

The LT of a function of time,  $f(t)$ , is found by multiplying the function by  $e^{-st}$  and integrating the product between the limits of zero and infinity. The result (if it exists) is known as the LT of  $f(t)$ . Thus:

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t) dt.$$

**Table of useful LTs**

| $f(t)$                   | $F(s) = \mathcal{L}\{f(t)\}$                            | <i>Comment</i>        |
|--------------------------|---------------------------------------------------------|-----------------------|
| $a$                      | $\frac{1}{s}$                                           | Step                  |
| $t$                      | $\frac{1}{s^2}$                                         | Ramp                  |
| $e^{at}$                 | $\frac{1}{s - a}$                                       | Exponential growth    |
| $e^{-at}$                | $\frac{1}{s + a}$                                       | Exponential decay     |
| $e^{-at} \sin(\omega t)$ | $\frac{\omega}{(s + a)^2 + \omega^2}$                   | Decaying sine         |
| $e^{-at} \cos(\omega t)$ | $\frac{s + a}{(s + a)^2 + \omega^2}$                    | Decaying cosine       |
| $\sin(\omega t + \phi)$  | $\frac{s \sin \phi + \omega \cos \phi}{s^2 + \omega^2}$ | Sine plus phase angle |
| $\sin(\omega t)$         | $\frac{\omega}{s^2 + \omega^2}$                         | Sine                  |
| $\cos(\omega t)$         | $\frac{s}{s^2 + \omega^2}$                              | Cosine                |

(continued)

Table of useful LTs (*continued*)

| $f(t)$                  | $F(s) = \mathcal{L}\{f(t)\}$                | Comment             |
|-------------------------|---------------------------------------------|---------------------|
| $\sinh \omega t$        | $\frac{\omega}{s^2 - \omega^2}$             | sin h               |
| $\cosh \omega t$        | $\frac{s}{s^2 - \omega^2}$                  | cos h               |
| $e^{at} \sinh \omega t$ | $\frac{\omega}{(s - a)^2 - \omega^2}$       | Shift sin h         |
| $e^{at} \cosh \omega t$ | $\frac{s - a}{(s - a)^2 - \omega^2}$        | Shift cos h         |
| $t \sin \omega t$       | $\frac{2\omega s}{(s^2 + \omega^2)^2}$      | Multiple $t$        |
| $t \cos \omega t$       | $\frac{s^2 - \omega^2}{(s^2 + \omega^2)^2}$ | Multiple $t$        |
| $\frac{d(t)}{dt}$       | $sF(s) - f(0)$                              | First differential  |
| $\frac{d^2 f(t)}{dt^2}$ | $s^2 F(s) - sf(0) - \frac{df(0)}{dt}$       | Second differential |
| $\int f(t) dt$          | $\frac{1}{s} F(s) + \frac{1}{s} F(0)$       | Integral            |

(6) *Fourier series*

(a) Real form:

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{2n\pi x}{T} + \sum_{n=1}^{\infty} b_n \sin \frac{2n\pi x}{T}.$$

where

$$a_n = \frac{2}{T} \int_{-\frac{1}{2}T}^{\frac{1}{2}T} f(x) \cos \frac{2n}{T} x dx, \quad b_n = \frac{2}{T} \int_{-\frac{1}{2}T}^{\frac{1}{2}T} f(x) \sin \frac{2n}{T} x dx.$$

(b) Harmonic analysis – tabular method:

| $\theta$ | $y$              | $2\cos \theta$ | $2y \cos \theta$ | $2s \sin \theta$ | $2y \sin \theta$ | $2\cos 2\theta$ | $2y \cos 2\theta$ | $2\sin 2\theta$ | $2y \sin 2\theta$ | ... |
|----------|------------------|----------------|------------------|------------------|------------------|-----------------|-------------------|-----------------|-------------------|-----|
| 0        |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 30       |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 60       |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 90       |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 120      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 150      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 180      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 210      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 240      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 270      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 300      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 330      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| 360      |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| Total    |                  |                |                  |                  |                  |                 |                   |                 |                   |     |
| Mean     | $\frac{1}{2}a_0$ |                | $a_1$            |                  | $b_1$            |                 | $a_2$             |                 | $b_2$             |     |

$$y = \frac{1}{2} a_0 + a_1 \cos \theta + b_1 \sin \theta + a_2 \cos 2\theta + b_2 \sin 2\theta + \dots$$

(c) Fourier series for special wave forms:

Square wave sine series

$$\frac{4K}{\pi} \left[ \sin \frac{2\pi x}{T} + \frac{1}{3} \sin \frac{6\pi x}{T} + \frac{1}{5} \sin \frac{10\pi x}{T} + \dots \right]$$

Square wave cosine series

$$\frac{4K}{\pi} \left[ \cos \frac{2\pi x}{T} - \frac{1}{3} \cos \frac{6\pi x}{T} + \frac{1}{5} \cos \frac{10\pi x}{T} + \dots \right]$$

Triangular

$$\frac{8K}{\pi^2} \left[ \cos \frac{2\pi x}{T} + \frac{1}{3^2} \cos \frac{6\pi x}{T} + \frac{1}{5^2} \cos \frac{10\pi x}{T} + \dots \right]$$

Sawtooth

$$\frac{2K}{\pi} \left[ \sin \frac{2\pi x}{T} - \frac{1}{2} \sin \frac{4\pi x}{T} + \frac{1}{3} \sin \frac{6\pi x}{T} - \dots \right]$$

In the above  $f(x)$  is periodic with period  $T$ .  $\pm K$  is the amplitude range of wave form.

## Methods

### Complex numbers

The complex number ( $z$ ) consists of a real (Re) part =  $x$  and an imaginary part =  $y$ , the imaginary unit ( $i$  or  $j$ ) multiplies the imaginary part  $y$ . In normal form, complex numbers are written as:

$$z = x + iy \quad \text{or} \quad z = x + jy \quad i = j$$

in all respects,  $j$  often being used for applications.

Example 6.5.1 shows how we apply the formulae to manipulate complex numbers, that is to add, subtract, multiply, and divide them.

#### Example 6.5.1

Add, subtract, multiply, and divide the following complex numbers:

- (a)  $(3 + 2j)$  and  $(4 + 3j)$ ;  
 (b) and in general  $(a + bj)$  and  $(c + dj)$ .

Addition:

$$\begin{aligned} (3 + 2j) + (4 + 3j) &= 3 + 4 + 2j + 3j \\ &= 7 + 5j \end{aligned}$$

$$(a + bj) + (c + dj) = (a + c) + (b + d)j.$$



Subtraction:

$$(3 + 2j) - (4 + 3j) = -1 - j$$

$$(a + bj) - (c + dj) = (a - c) + (b - d)j.$$

Multiplication:

$$\begin{aligned}(3 + 2j) \times (4 + 3j) &= 3(4 + 3j) + 2j(4 + 3j) \\ &= 12 + 9j + 8j + 6j^2.\end{aligned}$$

$$\begin{aligned}\text{Now from definition } j &= \sqrt{-1}, \text{ therefore } j^2 = -1 \text{ so,} \\ &= 12 + 17j + (6)(-1) \\ &= 6 + 17j\end{aligned}$$

$$\begin{aligned}(a + bj) \times (c + dj) &= ac + adj + bcj + bdj^2 \\ &= ac + adj + bcj - bd \text{ (where } j^2 = -1) \\ &= (ac - bd) + (ad + bc)j\end{aligned}$$

so the result of multiplication is still a complex number.

Division:

$$\frac{3 + 2j}{4 + 3j}$$

here we use an algebraic trick to assist us. We multiply top and bottom by the conjugate of the complex number in the denominator.

So here  $z = 4 + 3j$  then  $\bar{z} = 4 - 3j$  and we proceed as follows:

$$\left(\frac{3 + 2j}{4 + 3j}\right) \left(\frac{4 - 3j}{4 - 3j}\right) = \frac{12 - 9j + 8j - 6j^2}{16 + 12j - 12j - 9j^2} = \left(\frac{18 - j}{25}\right)$$

Note denominator becomes real, and in general:

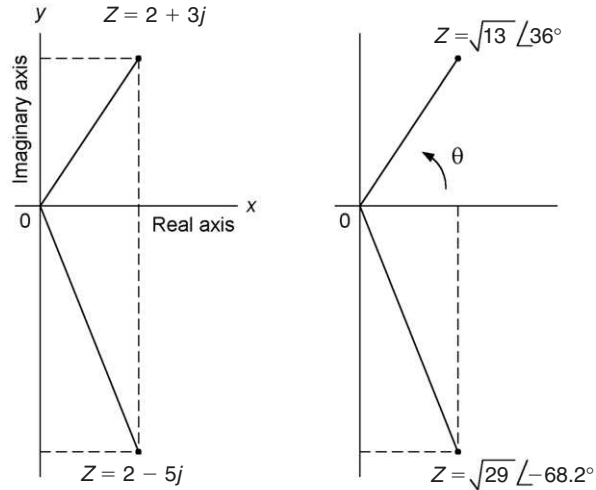
$$\begin{aligned}\frac{a + bj}{c + dj} &= \left(\frac{a + bj}{c + dj}\right) \left(\frac{c - dj}{c - dj}\right) \\ &= \frac{ac - adj + bcj - bdj^2}{c^2 + cdj - cdj - dj^2} \quad \text{and } -bdj^2 = bd, \text{ so} \\ &= \frac{(ac + bd) + (-adj + bcj)}{c^2 + d^2}.\end{aligned}$$

Complex numbers may be transformed from Cartesian to polar form by finding their *modulus and argument*, as the next example shows.

### Example 6.5.2

Express the complex numbers: (i)  $z = 2 + 3j$ ; (ii)  $z = 2 - 5j$  in polar form.

You will remember from your study of co-ordinate systems that polar co-ordinates are represented by an angle  $\theta$  and a magnitude  $r$  (as shown below). Complex numbers may be represented in the same way.



Complex number co-ordinate systems

To express complex numbers in polar form we will first find their modulus and argument. So from formulae 1(d),

if we have for  $z = 2 + 3j$  modulus  $= r = \sqrt{2^2 + 3^2} = \sqrt{13}$   
Argument  $= \theta$  where  $\tan \theta = y/x = 3/2 = 1.5$ ;  $\theta = 56.3^\circ$ . Then

$$z = 2 + 3j = \sqrt{13} (\cos 56.3 + j \sin 56.3)$$

$$z = \sqrt{13} \angle 56.3 \text{ (short hand form).}$$

Similarly for  $z = 2 - 5j$  then modulus  $= |z| = r = \sqrt{2^2 + (-5)^2}$   
so  $\sqrt{29}$  and the argument  $= \theta = -5/2 = -2.5 = -68.2^\circ$ .  
Then

$$z = 2 - 5j = \sqrt{29} (\cos(-68.2) + j \sin(-68.2))$$

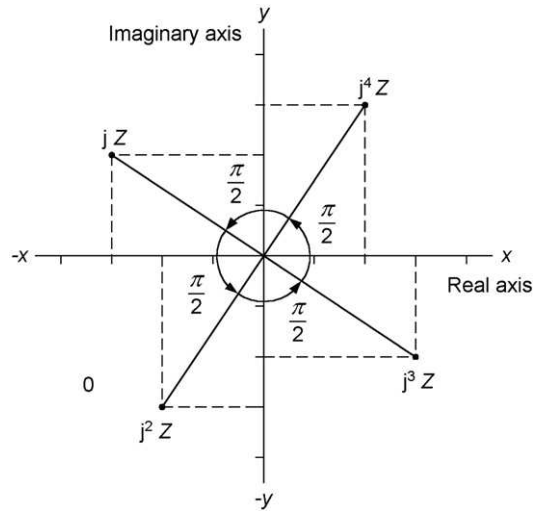
$$z = \sqrt{29} \angle -68.2.$$

The argument of a complex number in polar form, represents the angle  $\theta$  measured anticlockwise from the positive  $x$ -axis. Its tangent is given by  $y/x$ , therefore if we consider the argument  $\theta$  in radians, it can take on an infinite number of values, which are determined up to  $2\pi$  radians.

When we consider complex numbers in Cartesian form, then each time we multiply the complex number by  $(i = j)$ , the complex vector shifts by  $90^\circ$   $\pi/2$  radians. This fact is used when complex vectors represent *phasors* (electrical rotating vectors), successive multiplication by  $j$ , shifts the phase by  $\pi/2$ , as shown in Example 6.5.3. Under these circumstances the imaginary unit  $j$  is known as the  $j$ -operator (see Chapter 3 complex notation).

### Example 6.5.3

Multiply the complex number,  $z = 2 + 3j$  by the  $j$ -operator, three times in succession.

*j-operator rotation*

$$\text{Then } jz = j(2 + 3j) = 2j + 3j^2 = 2j - 3$$

$$j^2z = j(2j - 3) = 2j^2 - 3j = -3j - 2$$

$$j^3z = j(-3j - 2) = -3j^2 - 2j = -2j + 3$$

$$j^4z = j(-2j + 3) = -2j^2 + 3j = 2 + 3j.$$

Note that  $z = j^4z$  we have rotated the vector (phasor) through  $2\pi$  radians, back to its original position, as shown in the figure on the previous page.

We leave this very short introduction to complex numbers by considering the arithmetic operations of multiplication and division of complex numbers in polar form. Addition and subtraction is not considered because we have to convert the complex number from polar to Cartesian form before we can perform these operations!

When we multiply a complex number in polar form; we multiply their moduli and add their arguments. Conversely for division, we divide their moduli and subtract their arguments.

#### Example 6.5.4

For the complex numbers given below find the product of  $z_1$  and  $z_2$  also find  $z_1/z_2$ .

$$z_1 = 3(\cos 120 + j \sin 120)$$

$$z_2 = 4(\cos (-45) + j \sin (-45)).$$

Then  $z_1z_2$  = multiple of their moduli and the addition of their arguments.

$$\text{So } z_1z_2 = (3)(4)[\cos(120 - 45) + j \sin(120 - 45)]$$

$$\mathbf{z_1z_2 = 12(\cos 75 + j \sin 75).}$$

and similarly for division  $z_1/z_2$  = divisor of their moduli and the subtraction of their arguments.

$$\text{So } \frac{z_1}{z_2} = \frac{3}{4}[\cos(120 + 45) + j \sin(120 + 45)].$$

$$\frac{z_1}{z_2} = 0.75(\cos 165 + j \sin 165).$$

For the abbreviated version of complex numbers in polar form, we can multiply and divide in a similar manner. Once again they need to be converted into Cartesian form to be added and subtracted.

So in abbreviated form  $z_1 = 3\angle 120$  and  $z_2 = 4\angle 45$ . Then to multiply  $z_1 z_2$  we multiply their moduli and add their arguments.

$$\text{So } z_1 z_2 = r_1 r_2 \angle \theta_1 + \theta_2 = 12\angle 120 + 45 = 12\angle 165^\circ$$

$$\text{Similarly } \frac{z_1}{z_2} = \frac{r_1}{r_2} \angle \theta_1 - \theta_2 = \frac{3}{4} \angle 120 + 45 = 0.75 \angle 165^\circ$$

as before.

### Differential equations

The theory and solution of differential equations (DEs) is a science in its own right. The methods for their solution include the use of direct integration, separation of variables, integrating factors, D-operators, power series solutions, complex numerical techniques, and many others.

We will only be concerned with DEs, which can be solved using elementary techniques. A DE is simply an equation which contains differential coefficients. The order of a differential equation is given by the order of the highest derivative. For example:

$$\frac{dy}{dx} = 6x + 9 \quad \text{First order}$$

$$\frac{d^2 y}{ds^2} = ms = q \quad \text{Second order}$$

$$\frac{d^3 y}{dx^3} = ky \quad \text{or} \quad \frac{1}{y} \frac{d^3 y}{dx^3} = k \quad \text{Third order.}$$

Differential equations are often found in engineering, and express certain properties or laws relating to rates of change. They are formulated as a result of a situation under investigation or as a result of direct differentiation, their solution requires us to reverse these processes. The method of solution will depend on the number of variables and constants present. Methods involving direct integration and *separation of variables* are given below.

**Example 6.5.5**

Solve the following differential equations:

(a)  $\frac{dy}{dx} = x^3 - 2.$

(b)  $\frac{dy}{dx} = 3y - 6.$

(c)  $\frac{dy}{dx} = \frac{y^2}{x}.$

(a) This equation may be solved by using direct integration.

$$\int \frac{dy}{dx} = \int x^3 - 2 \, dx$$

then  $y = \frac{x^4}{4} - 2x + c.$

(notice the appearance of the constant of integration, we would need to have some information about the function to find a *particular value* for  $c$ . This information is referred to as the *boundary conditions* for the DE, as you will see later).

(b) Again we may integrate directly, after simple rearrangement of the variables.

$$\frac{dy}{dx} = 3y - 6$$

then  $\left(\frac{1}{3y-6}\right) \frac{dy}{dx} = 1$

and  $\int \left(\frac{1}{3y-6}\right) dy = \int 1 \, dx$

$$\frac{1}{3} \ln(3y-6) = x + c \quad (\text{see rules of integration})$$

$$\ln(3y-6) = 3x + 3c$$

$$3y-6 = e^{3x+3c} \quad (\text{antilogging!})$$

$$3y = e^{3x+3c} + 6$$

$$y = \frac{1}{3}e^{3x+3c} + 2.$$

(c)  $\frac{dy}{dx} = \frac{y}{x}$

here we need to separate the variables since there are two present  $x$  and  $y$ . So transposing the equation gives:

$$\frac{dy}{y} = \frac{dx}{x}$$

or  $\frac{1}{y} dy = \frac{1}{x} dx$  and integrating

$$\int \frac{1}{y} dy = \int \frac{1}{x} dx$$

$$\text{so } \frac{-1}{y} = \ln x + c$$

(now since  $c$  is a constant let it equal  $\ln c$ )

$$\text{then } \frac{-1}{y} = \ln x + \ln c$$

$$\frac{-1}{y} = \ln xc \text{ (laws of logs)}$$

$$y = -\frac{1}{\ln xc}.$$

Linear first-order differential equations, which can be put into the form

$$\frac{dy}{dx} + P(x)y = Q(x)$$

where  $P$  and  $Q$  are functions of  $x$ , may be solved using an *integrating factor* (IF).

$$\text{Where } \text{IF} = e^{\int P dx}$$

this assumes that we can evaluate the integral of  $P$ . If we cannot then the method cannot be used.

Now we multiply the whole equation by  $e^{\int P dx}$

$$\text{then } e^{\int P dx} \frac{dy}{dx} + e^{\int P dx} P(x)y = Q(x)e^{\int P dx}.$$

This can be written as:

$$\frac{d}{dx} (e^{\int P dx} y) = Q(x)e^{\int P dx}.$$

The idea is that the IF will produce a 'perfect differential'. This differential (being a product) may be integrated by parts to solve the equation.

The following example illustrates the IF method for solving linear first-order differential equations, if possible.

### Example 6.5.6

Solve the equation

$$x^2 \frac{dy}{dx} + xy = x^3$$

We first place the equation in the correct form which is given by:

$$\frac{dy}{dx} + P(x)y = Qx$$

so for our equation

$$\frac{dy}{dx} + \frac{y}{x} = x \quad (\text{on division by } x^2).$$

We now identify the IF, where we have  $P(x) = 1/x$

$$\text{so IF} = e^{\int \frac{1}{x} dx}.$$

Next we multiply each term of our equation by the IF to give:

$$e^{\int \frac{dx}{x}} \frac{dy}{dx} + e^{\int \frac{dx}{x}} \frac{y}{x} = x e^{\int \frac{dx}{x}}.$$

This may be written as:

$$\frac{d}{dx}(e^{\int P dx} y) = x e^{\int P dx}. \quad (1)$$

Substituting IF into Equation (1) gives:

$$\frac{d}{dx}(xy) = x \cdot x$$

$$\text{or } xy = \int x^2 dx$$

$$\text{and } xy = \frac{x^3}{3} + c. \quad (2)$$

Now if we are given the boundary conditions we are able to find a *particular* solution. Let  $y = 1$  when  $x = 3$ .

Then from (2)

$$3 = \frac{3^3}{3} + c$$

giving

$$c = -6$$

and a particular solution is

$$y = \frac{x^2}{3} - \frac{6}{x} \quad (\text{again from Equation (2)}).$$

A relatively simple method exists for the solution of second-order differential equations with constant coefficients. The method involves finding the roots of the auxiliary equation (the auxiliary equation can be written down directly from the original DE). Then, to find a *general solution*, we find the *complementary function* (CF) and add to it a particular solution (see formula 2(b)).

The whole process is described in the following example.

### Example 6.5.7

Obtain a general solution for the equation

$$\frac{d^2 y}{dx^2} - 5 \frac{dy}{dx} + 6y = 2 \sin 4x \quad (1)$$

Step 1: Write down the auxiliary equation, which has the form  $an^2 + bn + c = 0$ . Where  $n$  corresponds to  $dy/dx$ , then in our case we have  $n^2 - 5n + 6 = 0$ .

Step 2: Find the roots of the auxiliary equation. Factorising gives the roots as  $n = 3$ ,  $n = 2$ .

Step 3: Write down the CF. From formula 2(b) with values of  $n$  being real and not equal we get:

$$y = Ae^{3x} + Be^{2x} \quad (\text{where } n_1 = 3, n_2 = 2).$$

Step 4: Find a particular solution (often called the particular integral, PI).

Here we need to do a little 'guess work'. When  $f(x)$  in equation (1) involves a  $\sin bx$  or a  $\cos bx$ , try the solution

$$y = C \sin 4x + D \sin 4x.$$

Then  $\frac{dy}{dx} = 4C \cos 4x - 4D \sin 4x$

and  $\frac{d^2y}{dx^2} = -16C \sin 4x - 16D \cos 4x$

so substituting these values into Equation (1):

$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = 2 \sin 4x$$

gives

$$[-16C \sin 4x - 16D \cos 4x] - 5[4C \cos 4x - 4D \sin 4x] + \dots 6[C \sin 4x + D \cos 4x] = 2 \sin 4x$$

and on simplifying this expression we get:

$$\begin{aligned} -10C \sin 4x - 10D \cos 4x - 20C \cos 4x + 20D \sin 4x \\ = 2 \sin 4x. \end{aligned}$$

Now equating coefficients! (see partial fractions) we obtain:

$$-10C + 20D = 2 \quad (1)$$

$$-20C - 10D = 0 \quad (2)$$

multiplying Equation (2) by 2 and subtracting from (1):

$$-10C + 20D = 2$$

$$-40C - 20D = 0$$

$$-50C = 2 \quad \text{and} \quad C = -1/25,$$

and substitution into (1)

$$\text{gives} \quad (-10)(-\frac{1}{25}) + 20D = 2$$

$$20D = 2 - 2/5$$

$$\text{and} \quad D = 2/25.$$

So our particular solution is

$$y = -\frac{\sin 4x}{25} + \frac{2 \sin 4x}{25}$$



and the *general solution* to the equation:

$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = 2 \sin 4x$$

$$\text{is } = \mathbf{A}e^{3x} + \mathbf{B}e^{2x} - \frac{\sin 4x}{25} + \frac{2 \sin x}{25}.$$

The engineering application of DEs given in the next section limits their solution to that of direct integration and the determination of the constants of integration for given boundary conditions. Nevertheless the techniques illustrated here will prove useful for your future studies.

### Vectors and matrices

The techniques required for vector addition and the use of the dot and cross-product will be looked at in the next section, when we consider their application to frameworks. We concentrate here on one or two techniques which will enable you to multiply matrices and represent a system of linear equations in matrix form.

The process of matrix multiplication is detailed in formula 4(c). In order to multiply two matrices the number of columns of the first matrix must equal the number of rows of the second matrix. Matrix multiplication requires each matrix to be taken in order. So for two matrices, A and B, the product AB does not equal BA. Matrix multiplication methods are illustrated in Example 6.5.8.

#### Example 6.5.8

Find AB and BA for the following matrices:

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$$

Both are  $2 \times 2$  matrices, that is both have 2 rows and 2 columns, so multiplication may be carried out.

$$\begin{aligned} \text{so } \mathbf{AB} &= \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} (-1)(1) + (0)(3) & (-1)(2) + (0)(0) \\ (2)(1) + (3)(3) & (2)(2) + (3)(0) \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \\ &= \begin{bmatrix} -1 + 0 & -2 + 0 \\ 2 + 9 & 4 + 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix}. \end{aligned}$$

$$\begin{aligned}
\text{and } BA &= \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\
&= \begin{bmatrix} (1)(-1) + (2)(2) & (1)(0) + (2)(3) \\ (3)(-1) + (0)(2) & (3)(0) + (0)(3) \end{bmatrix} \\
&= \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix} \\
&= \begin{bmatrix} (-1 + 4) & (0 + 6) \\ (-3 + 0) & (0 + 0) \end{bmatrix} \\
&= \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}.
\end{aligned}$$

*Note:*  $AB \neq BA$ .

Square matrices with 1s on the lead diagonal are of special interest, such as:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \dots, \text{etc.}$$

A matrix of this form is called an *identity matrix* and is denoted by the symbol  $I$ . Multiplying a matrix  $A$  by the identity matrix  $I$  we get  $IA = A$ . In other words the identity matrix acts like the number 1 in arithmetic. Also, if  $A$  is any square matrix, and if a matrix  $B$  can be found such that  $AB = I$  then  $A$  is said to be *invertible* and  $B$  is called an *inverse* of  $A$ .

### Example 6.5.9

Show that the matrix  $B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$  is an inverse of

$A = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$ . Then we need to show that  $AB = I$

$$\begin{aligned}
AB &= \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} (6 - 5) & (10 - 10) \\ (-3 + 3) & (-5 + 6) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.
\end{aligned}$$

**If  $A$  is invertible as above then the *inverse* of  $A$  is represented as  $A^{-1}$  and so  $AA^{-1} = I$ .**

One more example of matrix multiplication is given in Example 6.5.10, make sure you can obtain the resulting identity matrix.

**Example 6.5.10**

For matrix A and matrix B given below, show that matrix B is the inverse of A.

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 5 \\ -1 & -1 & 2 \end{bmatrix} \quad B = \begin{bmatrix} -3 & 2 & -5 \\ 7 & -4 & 10 \\ 2 & -1 & 3 \end{bmatrix}$$

Then we need to show that  $AB = I$ , where  $B = A^{-1}$ .

$$\begin{aligned} AB &= \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 5 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} -3 & 2 & -5 \\ 7 & -4 & 10 \\ 2 & -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} (-6 + 7) + (4 - 4) + (-10 + 10) \\ (-3 - 7 + 10) + (2 + 4 - 5) + (-5 - 10 + 15) \\ (3 - 7 + 4) + (-2 + 4 - 2) + (5 - 10 + 6) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I. \end{aligned}$$

So  $AB = I$ , or  $AA^{-1} = I$  where  $B = A^{-1}$ .

Systems of linear equations may be put into matrix form and solved using *elementary row operations*. The complete process is illustrated in our final example on matrix methods, given below.

**Example 6.5.11**

Solve the system of linear equations:

$$-x + y + 2z = 2$$

$$3x - y + z = 6$$

$$-x + 3y + 4z = 4.$$

The first step in the process is to write down the *augmented matrix*. All we do is represent the coefficients of the variable and the numbers on the right-hand side of the equality signs, in the form of a matrix array.

$$\text{Then } A = \begin{bmatrix} -1 & 1 & 2 & 2 \\ 3 & -1 & 1 & 6 \\ -1 & 3 & 4 & 4 \end{bmatrix} \text{ note we remove the } - \text{ (minus) sign.}$$

Next we carry out elementary row operations as necessary to produce a matrix of the form

$$ax + by + cz = d \quad (1)$$

$$0 + ey + fz = g \quad (2)$$

$$0 \quad 0 \quad hz = i \quad (3)$$

then we use backward substitution to find a value for  $z$ , from Equation (3), placing value for  $z$  into Equation (2) and solving for  $y$  and so on.

So our first set of row operations need to eliminate  $-1$ , from matrix  $A$  as indicated. The row operations we are allowed to perform, correspond to the operations we can perform on the systems of equations. These are:

- (1) *multiply a row through by a non-zero constant;*
- (2) *interchange two rows;*
- (3) *add a multiple of one row to another.*

Then for our augmented matrix we eliminate  $x$  from row 3, ie.  $-1$ , by adding  $(-1 \times \text{row 1})$  to row 3, to give,

$$\begin{bmatrix} -1 & 1 & 2 & 2 \\ 3 & -1 & 1 & 6 \\ 0 & 2 & 2 & 2 \end{bmatrix}$$

(Notice only the row being operated on changes).

Adding  $(3 \times \text{row 1})$  to row 2 gives:

$$\begin{bmatrix} -1 & 1 & 2 & 2 \\ 0 & 2 & 7 & 12 \\ 0 & 2 & 2 & 2 \end{bmatrix}$$

Adding  $(-1 \times \text{row 2})$  to row 3 gives:

$$\begin{bmatrix} -1 & 1 & 2 & 2 \\ 0 & 2 & 7 & 12 \\ 0 & 0 & -5 & -10 \end{bmatrix}$$

Now our system of equations looks like:

$$-x + y + 2z = 2 \quad (1)$$

$$2y + 7z = 12 \quad (2)$$

$$-5z = -10 \quad (3)$$

and using *back substitution* we find that from Equation (3)  $z = 2$  and from Equation (2)  $y = -1$  and finally from Equation (1)  $x = 1$ .

### Laplace transforms

Differential equations are difficult to formulate and to solve as well as being difficult to manipulate. To help us overcome these problems we can revert to the use of *operators*. So, for example, if we find multiplication of numbers difficult we use the logarithm operator to convert multiplication into the simpler process of addition. We can adopt this idea by using the *s-operator*, or *LT*. This enables us to take our differential equation or other complex function which is in the time domain, convert it into the  $s$  domain (a complex domain where  $s = j\omega$ ) manipulate the simpler result, then convert it back to the time domain using Inverse LTs to obtain our desired solution.

The definition of the LT is given at the beginning of the table in the formula sheet, it is included again here with some further comments.

The LT of a function  $f(t)$  is given by:

$$\mathcal{L}[f(t)] = \int_0^{\infty} e^{-st} f(t) dt = F(s)$$

whenever the integral exists.

Note that the LT is denoted by the capital letter corresponding to the original function and is a function of  $s$  rather than  $t$ .

We will assume that any required transforms do actually exist. We will also assume that we are considering  $t \geq 0$  only.

The above definition may be used to evaluate a LT of a function. For example to find the LT of  $f(t) = t$ , then:

$$\mathcal{L}[t] = \lim_{k \rightarrow \infty} \int_0^k e^{-st} t dt.$$

This integral is a product, so integrating by parts we get:

$$\begin{aligned} &= \left[ -\frac{1}{s^2} e^{-st} (st + 1) \right]_0^k \\ &= -\frac{1}{s^2} e^{-sk} (sk + 1) + \frac{1}{s^2} \end{aligned}$$

and as  $k \rightarrow \infty$  the first term  $\rightarrow 0$  (true if  $s > 0$ ). Then:

$$\mathcal{L}[t] = \frac{1}{s^2}.$$

As you can see, obtaining LTs from first principles can be a tedious and often difficult process, fortunately there exist tables of standard LT to which we may refer. In practice we nearly always look up LTs in tables and, you should always try this first. The following rules may then help you to evaluate LTs of more complicated functions using the results for elementary functions, which may be read directly from the tables.

$$(1) \quad \mathcal{L}[f(t) + g(t)] = \mathcal{L}[f(t)] + \mathcal{L}[g(t)]$$

$$(2) \quad \mathcal{L}[af(t)] = a\mathcal{L}[f(t)] \quad a = \text{constant}.$$

The above rule indicates that  $\mathcal{L}$  is a *linear operator* (the  $s$ -operator).

$$(3) \quad \text{If} \quad \mathcal{L}[f(t)] = F(s)$$

$$\text{then} \quad \mathcal{L}[f(t)] = \frac{1}{a} F\left(\frac{s}{a}\right) \quad \text{the scale rule.}$$

$$(4) \quad \text{If} \quad \mathcal{L}[f(t)] = F(s)$$

$$\text{then} \quad \mathcal{L}[e^{at}f(t)] = F(s - a).$$

This rule is the *shift rule*, so called because the transform function  $F(s)$  is 'shifted' a distance ' $a$ ' along the  $s$ -axis by the presence of  $e^{at}$ .

$$(5) \quad \text{If } a > 0 \quad \mathcal{L}[f(t)] = 0 \quad t < 0 \quad \text{and} \quad g(t) = f(t - a) \\ \text{then} \quad \mathcal{L}[g(t)] = e^{-as} \mathcal{L}[f(t)].$$

This rule is equivalent to:

$$\mathcal{L}[u(t - a)f(t - a)] = e^{-as} \mathcal{L}[f(t)]$$

where  $u(t)$  is the limit step function

$$u(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0. \end{cases}$$

$$(6) \mathcal{L}[t^n] = F\left(\frac{n!}{s^{n+1}}\right) \text{ for integer } n > 0.$$

The above rule enables us to apply LTs to powers.

The following example shows the first stage of the LT process, finding the LT for a given function using tables.

### Example 6.5.12

Find the LT of the following functions:

- (a)  $\mathcal{L}[t^3 - 8t^2 + 1]$ .
- (b)  $\mathcal{L}[e^{2t} \sin 3t]$ .
- (c)  $\mathcal{L}[3et^2 + 2e^{-t} - 4 \cos t]$
- (d)  $\mathcal{L}[e^{-3t} \sin 2t]$ .
- (e)  $\mathcal{L}[t \cos 3t]$ .

$$(a) \mathcal{L}[t^3 - 8t^2 + 1] = \mathcal{L}(t^3) - 8\mathcal{L}(t^2) + \mathcal{L}(1)$$

$$\text{Tables and rule } 6 = \frac{6}{s^4} - \frac{16}{s^3} + \frac{1}{s}.$$

$$(b) \mathcal{L}[e^{2t} \sin 3t] = \frac{3}{(s-2)^2 + 9} \text{ (shift rule 4).}$$

$$(c) \mathcal{L}[3t^2 + 2e^{-t} - 4 \cos t] = 3\mathcal{L}(t^2) + 2\mathcal{L}(e^{-t}) - 4 (\cos t)$$

$$\text{Tables and rule } 6 = \frac{6}{s^3} + \frac{2}{s+1} - \frac{4s}{s^2+1}.$$

$$(d) \mathcal{L}[e^{-3t} \sin 2t] = \frac{2}{(s+3)^2 + 4} = \frac{2}{s^2 + 6s + 13}$$

decaying sine:

Note that you are given the shift rule in the tables for this function. Otherwise it requires you to apply it to  $\mathcal{L}[\sin 2t] = 2/(s^2 + 4)$ , then by shift rule 4 with  $a = -3$ , the above result follows.

$$(e) \mathcal{L}[t \cos 3t] = \frac{s^2 - 9}{(s^2 + 9)^2} \text{ for multiple } t.$$

Note the meaning of the multiple in the LT. The LT for  $\cos 3t$  is,

$$\mathcal{L}[\cos 3t] = \frac{s}{s^2 + 9}.$$

on multiplication by 't' when transformed we get:

$$\begin{aligned}\mathcal{L}[t \cos 3t] &= -\frac{d}{ds} \left( \frac{s^2}{s^2 + 9} \right) \quad \text{differential of transform so} \\ &= -\frac{9 - s^2}{(s^2 + 9)^2} = \frac{s^2 - 9}{(s^2 + 9)^2}\end{aligned}$$

that is the result from the tables.

We now turn our attention to the techniques required for finding inverse LTs. This requires us to use our rules in reverse and to apply the odd algebraic trick to get the function into the form we require. The important *reverse rules* for finding the inverse LTs are given below.

Rules for finding the inverse LT:

- (1)  $\mathcal{L}^{-1}[F(s) + G(s)] = \mathcal{L}^{-1}[F(s)] + \mathcal{L}^{-1}[G(s)];$
- (2)  $\mathcal{L}^{-1}[aF(s)] = a\mathcal{L}^{-1}[F(s)];$
- (3)  $\mathcal{L}^{-1}[F(s - a)] = e^{at}\mathcal{L}^{-1}[F(s)];$
- (4)  $\mathcal{L}^{-1}[e^{-as}F(s)] = u(t - a)\mathcal{L}^{-1}[F(s)].$

$u(t)$  is the step function.

Note the symmetry between the LTs and their inverses.

The following example shows how we put these rules into practice.

### Example 6.5.13

Find the inverse LTs for the following functions:

(a)  $\mathcal{L}^{-1}\left(\frac{4}{s^2 + 22} + \frac{12}{s - 5}\right);$

(b)  $\mathcal{L}^{-1}\left(\frac{2}{s^2 + 2s + 5}\right);$

(c)  $\mathcal{L}^{-1}\left(\frac{s + 1}{s(s^2 + 1)}\right);$

(a)  $\mathcal{L}^{-1}\left(\frac{4}{s^2 + 22} + \frac{12}{s - 5}\right).$

The only difficulty here is being able to put the constants into a form which enables us to read the inverse directly from the tables. This means that the constant in the numerator needs to be the square root of the constant in the denominator. Then:

$$\begin{aligned}\mathcal{L}^{-1}\frac{4}{s^2 + 22} + \mathcal{L}^{-1}\frac{12}{s - 5} &= 4\mathcal{L}^{-1}\frac{1}{s^2 + 22} + 12\mathcal{L}^{-1}\frac{1}{s - 5} \\ &= 4\mathcal{L}^{-1}\frac{1}{s^2 + 22} + 12e^{5t} \\ &\quad \text{(direct from tables).}\end{aligned}$$

The trick for the left-hand inverse is to multiply top and bottom by the square root of 22, that is  $\sqrt{22}/\sqrt{22}$ .

So we get

$$4\mathcal{L}^{-1}\frac{\sqrt{22}}{(s^2 + 22)\sqrt{22}}$$

which gives

$$\frac{4}{\sqrt{22}}\mathcal{L}^{-1}\frac{\sqrt{22}}{s^2 + 22}$$

this is the form required from the tables

$$\frac{4}{\sqrt{22}}\mathcal{L}^{-1}\frac{\sqrt{22}}{s^2 + 22} = \left(\frac{4}{\sqrt{22}}\right)\sin\sqrt{22}t + 12e^{5t}$$

$$(b) \quad \mathcal{L}^{-1}\left(\frac{2}{s^2 + 2s + 5}\right).$$

We need this inverse LT in shift form so the denominator looks like  $(s + a)^2 + \omega^2$ . We achieve this by 'completing the square', the quadratic expression in the denominator.

*To complete the square we use the following rule:*

$$s^2 + as + b = \left(s + \frac{a}{2}\right)^2 + b - \frac{a^2}{4}.$$

Then using this rule we get:

$$s^2 + 2s + 5 = (s + 1)^2 + 5 - \frac{4}{4}.$$

$$\text{So } \mathcal{L}^{-1}\left(\frac{2}{s^2 + 2s + 5}\right) = \left(\frac{2}{(s+1)^2 + 4}\right) = e^{-t} \sin 2t$$

direct from tables.

$$(c) \quad \text{Find } \mathcal{L}^{-1}\left(\frac{s+1}{s(s^2+1)}\right).$$

This requires us to use partial fractions (PFs) which you met when studying the algebra in this chapter.

$$\text{so } \mathcal{L}^{-1}\left(\frac{s+1}{s(s^2+1)}\right) = \mathcal{L}^{-1}\left[\frac{A}{s} + \frac{Bs+c}{s^2+1}\right].$$

Then the PFs are given from:

$$s + 1 = A(s^2 + 1) + (Bs + c)s$$

and equating coefficients:

$$s^0 \text{ gives } 1 = A$$

$$s^1 \text{ gives } 1 = C$$

$$s^2 \text{ gives } 0 = A + B$$



from which  $A = 1$ ;  $B = -1$ ;  $C = 1$ .

$$\begin{aligned}\mathcal{L}^{-1}\left(\frac{s+1}{s(s^2+1)}\right) &= \mathcal{L}^{-1}\left[\frac{1}{s} + \left(\frac{-s+1}{s^2+1}\right)\right] \\ &= \mathcal{L}^{-1}\left[\frac{1}{s} - \frac{s}{s^2+1} + \frac{1}{s^2+1}\right] \\ &= \mathbf{1 - \cos t + \sin t.}\end{aligned}$$

Our final technique involving the use of LTs is concerned with their use in the solution of differential equations (DEs). In our LT table (formula (5)) we have transforms for the first and second derivatives. The extension to these formulae is given below, where we are able to find the LT for  $n$ th order differential equations.

The LT for  $n$ th derivative ordinary differential equations, may be found by using:

$$\begin{aligned}\mathcal{L}[f^{(n)}(t)] &= s^n \mathcal{L}[f(t)] - s^{n-1}f(0) - s^{n-2}f'(0) \dots \\ &\quad - s^{n-3}f''(0) \dots - sf_{(0)}^{(n-2)} \dots - f_{(0)}^{(n-1)}.\end{aligned}$$

where  $f^{(n)}(t) = \frac{d^n f}{dt^n}$  is the  $n$ th derivative.

In the case of ordinary differential equation problems where the function ( $f$ ) and its derivatives are specified at time  $t = 0$  (as in the above rule), they are referred to as *initial value problems*. As opposed to problems where ( $f$ ) and its derivatives are specified at other values of ( $t$ ) which are called *boundary value problems*.

The procedure for applying LTs to ordinary differential equations is as follows:

- (1) Initial value problem given in the form of an ordinary differential equation (ODE).
- (2) The LT is applied to the ODE to produce an algebraic equation.
- (3) The algebraic equation is solved.
- (4) The inverse LT is applied, giving the solution of the initial value problem.

### Example 6.5.14

Solve  $y'' - 4y = t$  given  $y(0) = 1$ ,  $y'(0) = -2$ .

(1) We apply the LT to the ODE using the rules:

$$\begin{aligned}\mathcal{L}[y''] - 4\mathcal{L}[y] &= \mathcal{L}[t] \\ \mathcal{L}[y''] &= s^2 \mathcal{L}[y] - sy(0) - y'(0) \\ &\quad \text{(applying differential rule)} \\ &= s^2 \mathcal{L}[y] - s + 2.\end{aligned}$$

Since the ODE is second order (note the use of the initial value conditions),

and the  $\mathcal{L}[t] = \frac{1}{s^2}$ .

So the transformed equation is:

$$s^2 \mathcal{L}[y] - s + 2 - 4 \mathcal{L}[y] = \frac{1}{s^2}.$$

(2) Rearranging and manipulating the algebraic equation gives:

$$(s^2 - 4) \mathcal{L}[y] = \frac{1}{s^2} + s - 2.$$

$$\text{So } \mathcal{L}[y] = \frac{1}{s^2(s^2 - 4)} + \frac{s - 2}{s^2 - 4}.$$

(3) Now applying inverse LT:

$$y = \mathcal{L}^{-1} \left[ \frac{1}{s^2(s^2 - 4)} + \frac{s - 2}{s^2 - 4} \right].$$

Again we require the use of PFs for the term:

$$\frac{1}{s^2(s^2 - 4)} = \frac{A + Bs}{s^2} + \frac{s + D}{s^2 - 4}.$$

where  $A = -1/4$ ,  $D = 1/4$ ,  $B = C = 0$ , (make sure you are able to find these coefficients).

$$\text{Then } y = \mathcal{L}^{-1} \left[ -\frac{1}{4s^2} + \frac{1}{4(s^2 - 4)} + \frac{1}{s + 2} \right]$$

$$\text{since } \frac{s - 2}{s^2 - 4} = \frac{(s - 2)}{(s - 2)(s + 2)}$$

$$y = -\frac{1}{4} \mathcal{L}^{-1} \frac{1}{s^2} + \frac{1}{4} \mathcal{L}^{-1} \frac{1}{s^2 - 4} + \mathcal{L}^{-1} \frac{1}{s + 2}.$$

One last piece of algebraic manipulation on the middle term gives:

$$y = -\frac{1}{4} \mathcal{L}^{-1} \frac{1}{s^2} + \frac{1}{8} \mathcal{L}^{-1} \frac{2}{s^2 - 4} + \mathcal{L}^{-1} \frac{1}{s + 2}$$

$$\text{and } y = -\frac{1}{4}t + \frac{1}{8} \sinh 2t + e^{-2t} \quad (\text{from tables}).$$

Perhaps you can now see why manipulating algebra is so fundamental to all of your analytical study!

This concludes our study on advanced methods. The methods used for studying wave forms using harmonic analysis is treated briefly in the next section, as an application.

### Engineering applications

We begin our study of advanced applications, with a brief look at the use of harmonic analysis for determining the nature of wave forms, which is

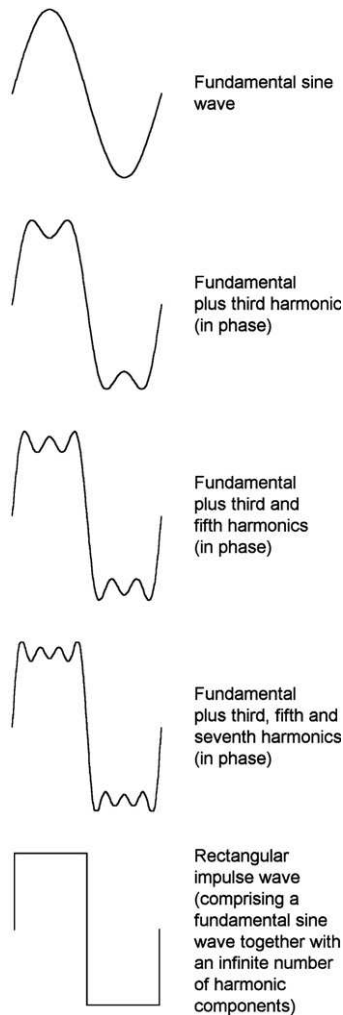


**Figure 6.5.1** Harmonic synthesis of rectangular impulse wave

a skill needed by those wishing to become Electrical Engineers. We then return to look at the application of complex numbers, differential equations, vectors, and finally LTs.

### Harmonic analysis

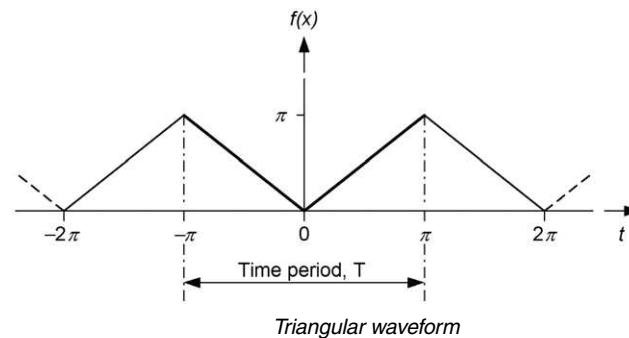
In Chapter 4.3 Fourier analysis is used in the study of complex waves. All wave forms whether continuous (like sinusoidal waves) or discontinuous (sawtooth waves), can be expressed in terms of a *convergent series*. You will remember from your study of series, that convergence guarantees the series is *bounded*.

The technique used for harmonic analysis, allows us to produce a better and better approximation to the shape of any wave form by successive addition of a series of sine waves, until the parent wave is replicated to the desired degree of accuracy. The more terms of the series we use, then the greater the number of waves, which may be added to produce the parent wave shape. This process is illustrated for the synthesis of a rectangular impulse wave in Figure 6.5.1.

For an illustration of the application of Fourier and harmonic analysis to a typical engineering application, we consider a triangular wave form.

#### Example 6.5.15

Find the Fourier series for the triangular wave form shown below.



From the formula given in 6(a), we know that the Fourier series for our triangular wave may be found using:

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{2\pi nt}{T} + \sum_{n=1}^{\infty} b_n \sin \frac{2\pi nt}{T}. \quad (1)$$

Now for our problem the time period  $T = 2\pi$ , so the above equation becomes:

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nt + \sum_{n=1}^{\infty} b_n \sin nt.$$

Note also that  $\omega = 2\pi/T$  which may be used in the series if desired.

We now need to find the Fourier coefficients  $a_0$ ,  $a_n$  and  $b_n$  for the series. If we were given tabular data (or more

numerical information about the triangular wave) we could use the tabular methods given in Example 4.3.3.

Even so, knowing whether the function is EVEN or ODD enables us to *halve* our work, since the integrals involving sine or cosine will equal zero. This can be determined by inspection of the wave form under consideration. If the integral of the function between the limits  $-\pi$  to  $\pi = 0$  (in other words the areas under the graph are both positive and negative) the function is ODD. The formal definition for ODD functions is that  $f(x) = -f(-x)$ , under these circumstances all the coefficients  $a_n$  vanish since  $f(x) \cos nx$  is ODD.

If between the limits  $-\pi$  to  $\pi$  the function has areas that produce a value, the function is EVEN. Or formally, if  $f(x) = f(-x)$  giving symmetry about the  $y$ -ordinate the function is EVEN. For similar reasons as before, the coefficients  $b_n$  vanish.

The result of all this is that, when the function is EVEN we need only consider the Fourier coefficients  $a_0, a_n$ . If the function is ODD then we need only consider the Fourier coefficient  $b_n$ .

I hope you can see that the *triangular function* given above is EVEN.

So coefficients  $b_n = 0$  ( $n = 1, 2, \dots$ ).

$$\text{Coefficients } a_n = \frac{2}{T} \int_{-\frac{1}{2}T}^{\frac{1}{2}T} f(x) \cos \frac{2n\pi x}{T} dx \quad (\text{form } 6(a))$$

and since  $T = 2\pi$  then:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt$$

$$\text{or } a_n = \frac{2}{\pi} \int_0^{\pi} f(t) \cos nt \, dt \quad (\text{notice the limits of integration!}).$$

Now  $f(t)$  for the interval  $-\pi$  to  $\pi$  gives:

$$f(t) = \begin{cases} -t & \text{if } -\pi \leq t \leq 0 \\ t & \text{if } 0 \leq t \leq \pi. \end{cases}$$

This is the *definition for the triangle function*.

$$\text{So, } a_n = \frac{2}{\pi} \int_0^{\pi} t \cos nt \, dt \quad (\text{since } f(t) = t \text{ from } 0 \text{ to } \pi).$$

Now to find  $a_n$  we will need to use integration by parts! For this example we will show the working in full.

$$\text{Then using } \int uv' \, dx = uv - \int u'v \, dx$$

$$\text{where } u = t, u' = 1$$

$$v' = \cos nt \quad \text{so } v = \frac{1}{n} \sin t$$

then

$$\begin{aligned}
 \frac{2}{\pi} \int_0^{\pi} t \cos nt \, dt &= \frac{t}{n} \sin t - \int \frac{1}{n} \sin nt \, dt \\
 &= \frac{t}{n} \sin t - \frac{1}{n^2} \cos nt \\
 &= \left[ \frac{t}{n} \sin t + \frac{1}{n^2} \cos nt \right]_0^{\pi} \\
 &= \left[ \frac{\pi}{n} \sin \pi + \frac{1}{n^2} \cos n\pi \right] - \left[ 0 + \frac{1}{n^2} \cos n0 \right] \\
 &= \left[ 0 + \frac{1}{n^2} \cos n\pi \right] - \left[ 0 + \frac{1}{n^2} \right] \\
 &= \frac{1}{n^2} \cos n\pi - \frac{1}{n^2}
 \end{aligned}$$

$$\begin{aligned}
 a_n &= \frac{2}{\pi} \int_0^{\pi} t \cos nt \, dt \\
 &= \frac{2}{\pi n^2} [\cos n\pi - 1].
 \end{aligned}$$

Now if integer  $n$  is EVEN  $\cos n\pi = +1$  hence  $a_n = 0$ . If integer  $n$  is ODD  $\cos n\pi = -1$  hence  $a_n = -4/(\pi n^2)$ :

We still need to find

$$a_0 = \frac{2}{\pi} \int_0^{\pi} t \cos 0 \, dt = \frac{2}{\pi} \int_0^{\pi} t \, dt$$

$a_0 = \pi$  (solving integral between limits).

So finally collecting the coefficients and putting them into Equation (1) yields:

$$f(t) = \frac{1}{2}(\pi) + \sum_{n=1}^{\infty} -\frac{4}{\pi n^2} \cos nt \quad (\text{for ODD 'n'})$$

and expanding

$$f(t) = \frac{\pi}{2} - \frac{4}{\pi} \left[ \frac{\cos t}{1^2} + \frac{\cos 3t}{3^2} + \frac{\cos 5t}{5^2} + \dots \right].$$

Note that the Fourier series expansion does not include EVEN values of ' $n$ ' within the bracket, remembering that when  $n$  is EVEN  $a_n = 0$ .

If you have managed to plough your way through the last example, you will realise that to find the coefficients using integration, requires care, skill, and extreme patience! This is why, whenever possible you should resort to the tabular method given in Example 4.3.3. Even the tabular calculations associated with this method can be eliminated with the aid of specially designed computer software (see Appendix).

### Complex numbers

You may have already applied complex numbers in your study of AC. We will complement this work by looking at one or two examples, similar to those given in Chapter 3.3.

**Example 6.5.16**

Find the resultant impedance in a circuit where the impedances are connected in parallel and are defined as follows:

$Z_1$ : Comprises a  $20\Omega$  resistance connected in series with an inductive reactance (at  $+90^\circ$ ) of  $8\Omega$ .

$Z_2$ : Comprises a  $15\Omega$  resistance connected in series with a capacitive reactance (at  $-90^\circ$ ) of  $10\Omega$ .

Complex numbers are required to find an expression for the resultant (total) impedance  $Z_T$  of the circuit. Before this can be done we need to write down  $Z_1$  and  $Z_2$  remembering that the  $j$ -operator shifts the complex variable by  $90^\circ$ .

$$\text{Then } Z_1 = 20 + j8$$

$$Z_2 = 15 - j10.$$

Now impedances (from Chapter 3) in parallel are found using:

$$\frac{1}{Z_T} = \frac{1}{Z_1} + \frac{1}{Z_2} + \dots$$

so we add  $\frac{1}{Z_1} + \frac{1}{Z_2}$  using fractions!

$$\frac{1}{Z_T} = \frac{1}{Z_1} + \frac{1}{Z_2}$$

$$\frac{1}{Z_T} = \frac{Z_2 + Z_1}{Z_1 Z_2} \quad (\text{Now we can turn upside down})$$

$$Z_T = \frac{Z_1 Z_2}{Z_2 + Z_1}.$$

$$\text{Now } Z_T = \frac{(20 + j8)(15 - j10)}{(20 + j8) + (15 - j10)}$$

Applying complex multiplication and addition gives:

$$Z_T = \frac{300 + j120 - j200 - j^2 80}{(35 - j2)}$$

$$Z_T = \frac{380 - 80j}{35 - 2j} \quad (\text{now multiplying top and bottom by the conjugate of denominator})$$

$$\begin{aligned} \text{then } Z_T &= \frac{(380 - 80j)(35 + 2j)}{(35 - 2j)(35 + 2j)} \\ &= \frac{13\,300 + 760j - 2800j - 160j^2}{1225 + 70j - 70j - 4j^2} \\ &= \frac{13\,460 - 2040j}{1229} \end{aligned}$$

$$\text{so } Z_T = 10.95 - 1.66j.$$

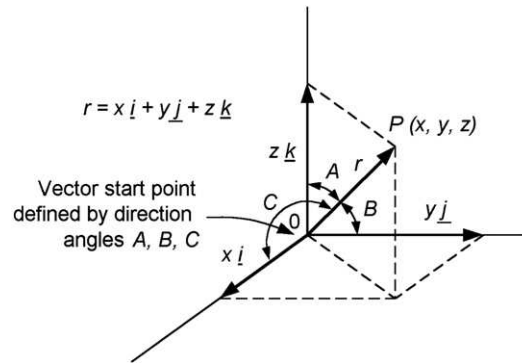


**Figure 6.5.2** Position vector system

### Example 6.5.17

An impedance  $Z = (200 + j100)\Omega$  is connected to a 50 V supply. Find an expression for the current in complex form given that  $I = V/Z$  and  $I$ ,  $V$ , and  $Z$  are complex variables.

Then  $I = V/Z$  simply requires complex division

$$\begin{aligned} I &= \frac{50}{200 + j100} \\ &= \frac{50(200 - j100)}{(200 + j100)(200 - j100)} \\ &= \frac{10000 - j5000}{40000 - j^2 10000} \end{aligned}$$

$$\begin{aligned} I &= \frac{10000 - j5000}{50000} \\ &= 0.2 - j0.1 \end{aligned}$$

so  $I = (0.2 - j0.1) \text{ A}$ .

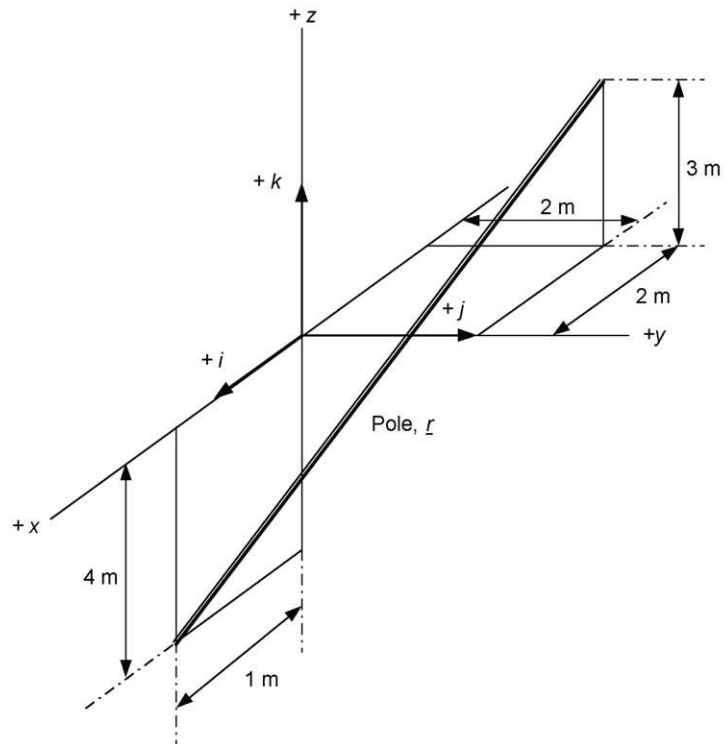
### Vectors and matrices

You should be familiar with the concept of a vector and vector addition of forces. Look at the formulae on vectors at the beginning of this section. The Cartesian co-ordinate system in three dimensions is represented by the position vectors  $i$ ,  $j$ , and  $k$  which act at right angles to each other, they are *orthogonal*.

A *position vector* system is directly comparable with the Cartesian vector system and uses the same direction vectors  $i$ ,  $j$ , and  $k$ . The position vector  $r$  is defined as a fixed vector which locates a point in space relative to another point. The position vector can be expressed in Cartesian form as  $r = xi + yj + zk$  (Figure 6.5.2). Let us consider an example, which will illustrate the power of the position vector system.

### Example 6.5.18

Determine the magnitude and direction of the pole  $r$ , shown in Figure (Example 6.5.18).



The vector representing the pole can be found by determining the co-ordinates of its start and finish points A and B.

A(1, 0, -4); B(-2, 2, 3).

Now  $\mathbf{r} = \mathbf{r}_B - \mathbf{r}_A$  (in other words position at finish minus position at start, defines size and position of  $\mathbf{r}$ ).

$$\text{so } \mathbf{r} = (-2 - 1)\mathbf{i} + (2 - 0)\mathbf{j} + (3 - (-4))\mathbf{k}$$

$$\mathbf{r} = -3\mathbf{i} + 2\mathbf{j} + 7\mathbf{k}.$$

Then the magnitude of  $|\mathbf{r}| = \sqrt{(-3)^2 + (2)^2 + (7)^2} = 7.9 \text{ m}$ .  
The unit vector  $\mathbf{u}$  is defined as  $\mathbf{u} = \mathbf{r}/r$  that is the unit vector in the direction of  $\mathbf{r}$  is formed by dividing the individual components of  $\mathbf{r}$  by the size of  $|\mathbf{r}|$ .

This is identical to dividing a vector by its scalar multiplier to obtain the unit vector. Then the unit vector:

$$\mathbf{u} = \frac{\mathbf{r}}{r} = \frac{-3}{7.9}\mathbf{i} + \frac{2}{7.9}\mathbf{j} + \frac{7}{7.9}\mathbf{k}.$$

Now the components of the unit vector give the cosines of the co-ordinate direction angles, of the strut from its start point. So these angles position the strut in its fixed direction as it leaves A for B.

$$\text{Then } \cos \theta_i = -0.3797 \quad \text{so } \theta_i = 112.3^\circ$$

$$\cos \theta_j = 0.2538 \quad \theta_j = 75.3^\circ$$

$$\cos \theta_k = 0.866 \quad \theta_k = 27.6^\circ$$

So we have determined both the size of the strut and its direction in space.



The important points to note in this example are:

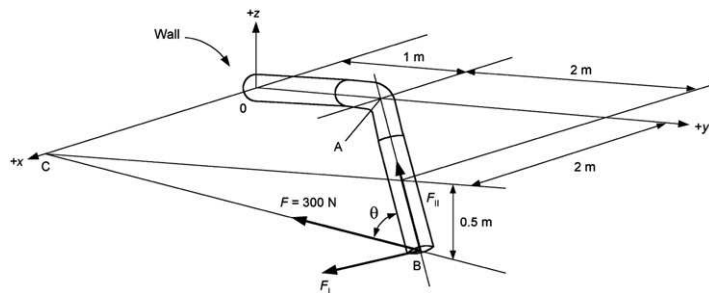
- (i) the technique for finding the vector representing a *line in space*, considering its start and finish points that is  $\mathbf{r} = \mathbf{r}_B - \mathbf{r}_A$ ;
- (ii) finding the *direction cosines* using the unit vector  $\mathbf{u}$  (this always works);
- (iii) noting that the magnitude of the vector  $|\mathbf{r}|$  was found in the normal way (formula 3(a)).

If you find difficulty in following this application of vectors you will need to ensure that you review the theory associated with vector geometry.

The next example shows the use of the vector dot product. The dot product has two very useful applications. It enables us to find the angle formed between two vectors or intersecting lines and, when used with the unit vector it enables us to find the component of a vector parallel to a line. These applications are illustrated for the force system shown in the next example.

### Example 6.5.19

An overflow pipe is subject to a force  $F = 300 \text{ N}$  at its end B. Determine the angle  $\theta$  between the force and the pipe section AB and, find the magnitudes of the components of the force  $F$ , which act parallel and perpendicular to B. Figure example 6.5.19 shows the situation.



To find the angle  $\theta$  we use the dot product as given in the formula sheet.

$$A \cdot B = |A||B| \cos \theta$$

$$\text{or } \cos \theta = \frac{A \cdot B}{|A||B|}.$$

To locate the point of contact of force vector  $\mathbf{F}$  we first find the position vector along BA and BC. Thus by inspection

$$\mathbf{r}_{BA} = (-2\mathbf{i} - 2\mathbf{j} + 0.5\mathbf{k})\text{m}.$$

Note that to get from B to the point A, we travel back 2 m along the  $x$ -axis, 2 m back along the  $y$ -axis, and 0.5 m up the  $z$ -axis.

Similarly:  $\tau_{BC} = (0\mathbf{i} - 3\mathbf{j} + 0.5\mathbf{k})$ . That is from B to C we travel 3 m along the  $y$ -axis and up 0.5 m along the  $z$ -axis, C and B are the same distance along the  $x$ -axis.

We are now in a position to find  $\cos \theta$  using the dot product.

$$\begin{aligned} A \cdot B &= [(-2)(0) + (-2)(-3) + (0.5)(0.5)] \\ &= (0 + 6 + 0.25) \end{aligned}$$

$$\text{and } |r_{BA}| = \sqrt{(-2)^2 + (-2)^2 + (0.5)^2} = \sqrt{8.25}$$

$$|r_{BC}| = \sqrt{(0)^2 + (-3)^2 + (0.5)^2} = \sqrt{9.25}$$

$$\begin{aligned} \text{then } \cos \theta &= \frac{A \cdot B}{|r_{BA}| |r_{BC}|} = \frac{6.25}{\sqrt{8.25} \sqrt{9.25}} = \frac{6.25}{8.736} \\ &= 0.7155 \end{aligned}$$

$$\text{giving } \theta = 44.3^\circ.$$

Now we have  $\theta$ , then force parallel to pipe section  $F_{\parallel}$  is given by  $F_{\parallel} = F \cos \theta = (300)(0.7155) = \mathbf{214.7 \text{ N}}$  and  $F_{\perp} = F \sin \theta = (300)(0.6984) = \mathbf{209.5 \text{ N}}$ .

We now look at an application for elementary row operations, for the solution of linear equations. This method is sometimes referred to as *Gaussian elimination*.

### Example 6.5.20

The equations shown below were obtained by modelling an electric circuit having two e.m.f.s of 24 V and 12 V, using Kirchoff's laws. Using Gaussian elimination find the values of the currents  $I_1$ ,  $I_2$ ,  $I_3$ .

$$12I_1 + 4(I_1 - I_2) = 24$$

$$4(I_2 - I_1) + 3(I_2 - I_3) = 0$$

$$3I_3 + 3(I_3 - I_2) = 12.$$

We first simplify the equations to give:

$$16I_1 - 4I_2 = 24$$

$$-4I_1 + 7I_2 - 3I_3 = 0$$

$$-3I_2 + 6I_3 = 12.$$

Now write down the augmented matrix

$$\begin{array}{cccc|l} 16 & -4 & 0 & 24 & \text{(row 1)} \\ -4 & 7 & -3 & 0 & \text{(row 2)} \\ 0 & -3 & 6 & 12 & \text{(row 3)} \end{array}$$

Now using elementary row operations we get the matrix into *row-echelon* form, that is zeros to the left and below the lead diagonal. We need to eliminate the  $-4$  term from row 2. This can be achieved by adding a quarter of row 1 to row 2.

So adding  $(\frac{1}{4} \times \text{row 1})$  to row 2 gives:

$$\begin{array}{cccc} 16 & -4 & 0 & 24 \end{array} \quad (1)$$

$$\begin{array}{cccc} 0 & 6 & -3 & 6 \end{array} \quad (2)$$

$$\begin{array}{cccc} 0 & -3 & 6 & 12 \end{array} \quad (3)$$

Adding  $(\frac{1}{2} \times \text{row 2})$  to row 3 gives:

$$\begin{array}{cccc} 16 & -4 & 0 & 24 \end{array} \quad (1)$$

$$\begin{array}{cccc} 0 & 6 & -3 & 6 \end{array} \quad (2)$$

$$\begin{array}{cccc} 0 & 0 & 4.5 & 15 \end{array} \quad (3)$$

We are now in row-echelon form, so replacing matrix with original equations gives:

$$16I_1 - 4I_2 = 24$$

$$6I_2 - 3I_3 = 6$$

$$4.5I_3 = 15.$$

Now using backward substitution we get:

$$I_3 = \frac{15}{4.5} = 3\frac{1}{3},$$

$$6I_2 - (3)(3\frac{1}{3}) = 6 \quad \text{so } I_2 = 2\frac{1}{6}.$$

$$\text{and finally } 16I_1 - (4)(2\frac{2}{3}) = 24 \quad \text{so } I_1 = 2\frac{1}{6}.$$

Then required current values are:

$$I_1 = 2\frac{1}{6} \text{ A}, \quad I_2 = 2\frac{2}{3} \text{ A}, \quad I_3 = 3\frac{1}{3} \text{ A}.$$

### Differential equations

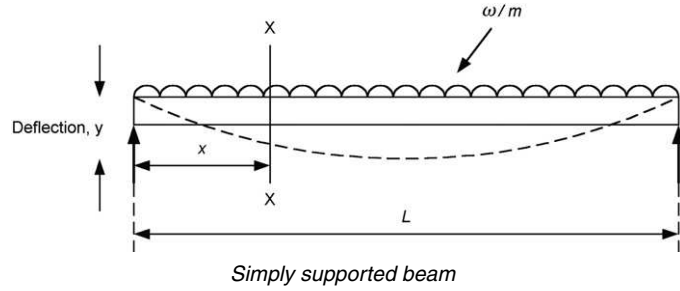
We know from our previous work that a first-order differential equation will have one constant of integration, a second order will have two, and so on. In order to solve these equations completely we need to be given boundary conditions. In Chapter 5, we considered the use of the Macaulay method for the bending of beams, this required the formulation of differential equations, for different situations. To solve these equations the boundary conditions were considered, as the next example shows.

#### Example 6.5.21

The figure shows a simply supported beam subject to a uniformly distributed load, causing a deflection  $y$ .

If the differential equation modelling the situation is given by:

$$EI \frac{d^2y}{dx^2} - m = 0$$



where the bending moment  $m$  at section  $x-x$  is

$$\frac{\omega Lx}{2} - \frac{\omega x^2}{2}.$$

The product  $EI$  may be treated as a constant. Solve the DE given the boundary conditions:

$$x = 0, y = 0$$

$$x = L, y = 0.$$

Then we have:

$$EI \frac{d^2 y}{dx^2} = \frac{\omega Lx}{2} - \frac{\omega x^2}{2} \left( m = \frac{\omega Lx}{2} - \frac{\omega x^2}{2} \right).$$

This is a second-order differential equation which can be solved by direct integration.

So integrating both sides once gives:

$$EI \frac{dy}{dx} = \frac{\omega Lx^2}{4} - \frac{\omega x^3}{6} + A \quad (A \text{ is constant of integration}).$$

and integrating again gives:

$$EI_y = \frac{\omega Lx^3}{12} - \frac{\omega x^4}{24} + Ax + B. \quad (1)$$

Now using the boundary conditions to find constant  $A$  and  $B$ . Then when  $x = 0, y = 0$  and substituting these values into (1) gives  $B = 0$ . Also when  $x = L, y = 0$  and from Equation (1) again we get:

$$0 = \frac{\omega L^4}{12} - \frac{\omega L^4}{24} + AL$$

$$-\frac{\omega L^4}{24} = AL$$

$$\text{and } A = -\frac{\omega L^3}{24}.$$

$$\text{Therefore } EI_y = \frac{\omega Lx^3}{12} - \frac{\omega x^4}{24} - \frac{\omega xL^3}{24}$$

$$\text{hence } y = \frac{1}{EI} \left[ \frac{\omega Lx^3}{12} - \frac{\omega x^4}{24} - \frac{\omega xL^3}{24} \right].$$

The next example involves a differential equation of first order and relies on knowledge of Newton's second law of motion for its solution.

### Example 6.5.22

From Newton's second law  $F = ma$  it can be shown that for a body falling freely in a vacuum then:

$$m \frac{dv}{dt} = mg$$

where  $m$  = mass of body,  
 $g$  = acceleration due to gravity ( $9.81 \text{ m/s}^2$ ),  
 and  $v$  = velocity of body ( $\text{m/s}$ ).

Given that  $v = 0$ , when  $t = 0$ , solve the differential equation and find the velocity of the body, when  $t = 3.5 \text{ s}$ .

The differential equation  $m \frac{dv}{dt} = mg$  is easily solved by separation of variables.

Rearranging gives  $dv = g dt$

$$\begin{aligned} \text{so } \int 1 dv &= g \int dt \\ &= v = gt + A. \end{aligned} \quad (1)$$

Now using boundary condition  $v = 0$  when  $t = 0$  in Equation (1) gives  $A = 0$ .

so  $v = gt$  and at time  $3.5 \text{ s}$ .

$$v = (9.81)(3.5) = \mathbf{34.3 \text{ m/s}}.$$

### Laplace transforms

LTs play a particularly important role in the modelling of engineering control systems. A control system transfer function  $G$  gives the ratio of the output signal to the input signal for the system. Since the modelling of engineering systems often involves the formulation of DEs, then it becomes necessary to manipulate these equations in order to modify system behaviour. This difficult process is best achieved by using LTs and carrying out the desired manipulation in the  $s$ -plane.

The analysis of AC electrical circuits involves the representation of system variables such as capacitive reactance, as complex functions. Again, the DEs formed using these variables are best manipulated in the  $s$ -plane, after the Laplace operator has been applied.

The final examples in this section, illustrate the above uses.

### Example 6.5.23

The closed loop transfer function for a control system is given below. Determine the time domain response to a unit-step input.

$$G = \frac{2}{s(s - 5)}.$$

Note first that  $G$  is already in the  $s$ -domain, that is, the LT has been performed on  $G$ .

To convert to the time domain we are required to find the  $\mathcal{L}^{-1}$ .

Before we do this we are told that the system is subject to a unit step response. From our LT table a unit step response in the ' $s$ ' domain is given as  $1/s$ . Therefore, the system transfer function  $G$ , is now:

$$F(G) = \frac{1}{s} \left( \frac{2}{s(s-5)} \right) = \frac{2}{s^2(s-5)}$$

therefore, we require

$$\mathcal{L}^{-1} \frac{2}{s^2(s-5)}$$

this cannot be found directly from the tables, so we need to simplify  $G$ , by finding the PFs.

$$\text{Then } \frac{2}{s^2(s-5)} = \frac{A}{s^2} + \frac{B}{s} + \frac{C}{s-5}.$$

$$\text{So } 2 = As - 5A + Bs^2 - 5Bs + Cs^2.$$

If we equate coefficients then:

$$s^2: 0 = B + C$$

$$s^1: 0 = A - 5B$$

$$s^0: 2 = -5A.$$

$$\text{So } A = -\frac{2}{5}, B = -\frac{2}{25}, C = \frac{2}{25}.$$

$$\text{Then we require } \mathcal{L}^{-1} \left[ \frac{2}{25(s-5)} - \frac{2}{25s} - \frac{2}{5s^2} \right]$$

$$\text{or } \frac{2}{25} \mathcal{L}^{-1} \frac{1}{s-5} - \frac{2}{25} \mathcal{L}^{-1} \frac{1}{s} - \frac{2}{5} \mathcal{L}^{-1} \frac{1}{s^2}.$$

These may now be read directly from the tables to give our function  $f(G)$  in the time domain as:

$$f(G) = \frac{2}{25} e^{5t} = \frac{2}{25} - \frac{2t}{5}.$$

### Example 6.5.24

An electric circuit is modelled by the DE:

$$V = Ri(t) + L \frac{di(t)}{dt} + \frac{1}{C} \int i(t) dt.$$

Solve the equation *in terms of the current ( $i$ )* using LTs, given that  $i = 0$  at  $t = 0$ , and  $v = 12$ ,  $L = 2$ ,  $R = 8$  and  $C = 0.025$ .

Applying the LT gives:

$$\mathcal{L}[V] = R\mathcal{L}[i(t)] + L\mathcal{L}\left[\frac{di(t)}{dt}\right] + \mathcal{L}\left[\frac{1}{C}\int i(t)dt\right].$$

Using our table then,

$$\frac{V}{s} = Rl(s) + sL[l(s) - i(0)] + \frac{1}{sC}[ls + l(0)]$$

and applying initial conditions  $i = 0$ , at  $t = 0$  we get  $i(0) = 0$  and  $l(0) = 0$ .

$$\text{Therefore } \frac{V}{s} = Rl(s) + sL[l(s) - 0] + \frac{1}{sC}[ls + l(0)]$$

$$\frac{V}{s} = Rl(s) + sLl(s) + \frac{1}{sC}ls$$

$$\frac{V}{s} = l(s)\left(R + sL + \frac{1}{sC}\right).$$

So in terms of  $l(s)$ :

$$l(s) = \frac{V}{s\left(R + sL + \frac{1}{sC}\right)} = \frac{V}{s^2L + sR + \frac{1}{C}}$$

and on substitution of given values:

$$l(s) = \frac{12}{2s^2 + 8s + 40} = \frac{6}{s^2 + 4s + 20}.$$

Now we need to find the *inverse LT*

$$\mathcal{L}^{-1}\frac{6}{s^2 + 5s + 20}$$

We cannot obtain this inverse directly from the tables but, using the method of completing the square gives:

$$\mathcal{L}^{-1}\frac{6}{\left(s + \frac{5}{2}\right)^2 + 20 - \frac{16}{4}} \quad \text{or}$$

$$\mathcal{L}^{-1}\frac{6}{(s + 2.5)^2 + 16} \quad \text{and on multiplication by } \frac{\sqrt{16}}{\sqrt{16}}$$

$$6\mathcal{L}^{-1}\frac{\sqrt{16}}{[(s + 2.5)^2 + 16]\sqrt{16}} \quad \text{or}$$

$$\frac{6}{\sqrt{16}}\mathcal{L}^{-1}\frac{\sqrt{16}}{(s + 2.5)^2 + 16} \quad \text{which is required form}$$

for direct use of tables where  $F(s) = \frac{\omega}{(s + a)^2 + \omega^2}$

$$\text{then } I(s) = \left( \frac{6}{\sqrt{16}} \right) \frac{\sqrt{16}}{(s + 2.5)^2 + 16}$$

$$\text{so } i(t) = \frac{6}{\sqrt{16}} e^{-2.5t} \sin \sqrt{16} t.$$

Compare with Example 4.4.3!

We end this section and our study of analytical methods with a number of problems of varying difficulty.

### Questions 6.5.1

- (1) Perform the required calculation and express the answer in the form  $a + ib$  for:

(a)  $(3 - 2i) - i(4 + 5i)$ ; (b)  $(7 - 3i)(5i + 3)$ ; (c)  $i^5$ ;

(d)  $\frac{1 + 2i}{3 - 4i}$ ; (e)  $\frac{(4 - i)(1 - 3i)}{-1 + 2i}$ .

- (2) Represent the following complex numbers in polar form:

(a)  $6 - 6i$ ; (b)  $3 + 4i$ ; (c)  $(4 + 5i)^2$ ; (d)  $\frac{1}{(1 - 2i)^2}$ .

- (3) Express the following complex numbers in Cartesian form:

(a)  $\sqrt{30} \angle 60^\circ$ ; (b)  $\sqrt{13} \angle \frac{\pi}{4}$ ; (c)  $2 \cos \frac{\pi}{6} + j 2 \sin \frac{\pi}{6}$ .

- (4)  $Z_1 = 20 + 10j$ ;  $Z_2 = 15 - 25j$ ;  $Z_3 = 30 + 5j$ .

Find: (a)  $|Z_1||Z_2|$ ; (b)  $Z_1 Z_2 Z_3$ ; (c)  $\frac{Z_1 Z_2}{Z_3}$ ; (d)  $\frac{Z_1 Z_2}{Z_1 Z_3}$ .

- (5) An impedance  $Z$  of  $(200 + j80)\Omega$  is connected to a 100 V rectified current (R.C.) supply. Given that  $I = \frac{V}{Z}$ , determine the current flowing and express your answer in complex form.

- (6) Solve the following differential equations:

(a)  $\frac{dy}{dx} = \frac{2x}{2x + y}$ ;

(b)  $x \frac{dy}{dx} = -y + 2x$  given  $y = 0$  at  $x = 2$ .

- (7) Obtain the general solution of the differential equation:

$$\frac{d^2 y}{dt^2} = -\omega^2 \cos \omega t.$$

- (8) Obtain the particular integral for the equation:

$$\frac{d^2 y}{dx^2} + 4 \frac{dy}{dx} + 2y = 2x + 3.$$

- (9) Solve  $y'' + 4y = 0$ , given that  $y(0) = 0$  and  $y'(0) = 1$ .



- (10) The bending moment for a beam is related by the equation:

$$M = EI \frac{d^2 y}{dx^2} \quad \text{where } M = \frac{WL}{4} - \frac{Wx}{2}.$$

Find an equation for  $y$  given that when  $x = 0$ ,  $y' = 0$  and when  $x = \frac{6}{2}$ ,  $y = 0$ .

- (11) The second order differential equation:

$$\frac{d^2 q}{dt^2} + \frac{6dq}{dt} + 40q = 115$$

models a CRL series circuit with an e.m.f of 115 V. Find an equation for  $q$  (the charge) in terms of  $t$ .

- (12) Solve the system of linear equations:

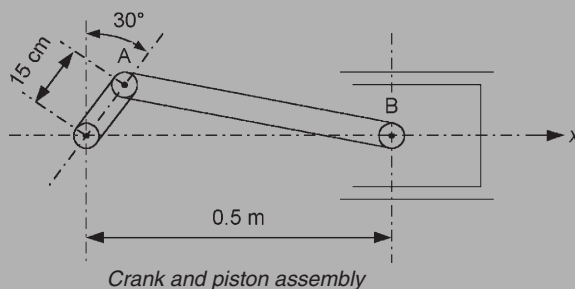
$$3x - 7y + 4z = 10$$

$$x + y + 2z = 8$$

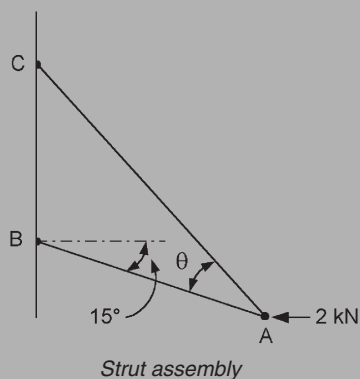
$$-x - 2y + 3z = 1$$

using Gaussian elimination.

- (13) Determine the length of the crankshaft AB by first formulating a Cartesian position vector from A to B and then determining its magnitude.



- (14) Determine the angle  $\theta$  (acute) between the two struts so that the horizontal force (2 kN) has a component of 2.3 kN directed from A towards C.



- (15) Using an analytical method show that the Fourier series for a fully rectified sine wave is given by:

$$\frac{4x}{\pi} \left[ \frac{1}{2} - \frac{1}{1.3} \cos \frac{4\pi x}{T} - \frac{1}{3.5} \cos \frac{8\pi x}{T} - \frac{1}{5.7} \cos \frac{12\pi x}{T} - \dots \right]$$

for limits  $-\frac{1}{2}T \leq x \leq \frac{1}{2}T$ , amplitude  $k$ .

- (16) Use harmonic analysis, with tabulated values at 30° steps to find the Fourier series up to the fourth harmonic for the wave form given by the following data.

---

|          |    |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $\theta$ | 0  | 30  | 60  | 90  | 120 | 150 | 180 | 210 | 240 | 270 | 300 | 330 | 360 |
| $y$      | 10 | 9.5 | 8.5 | 7.5 | 7.0 | 6.5 | 6.0 | 5.0 | 4.5 | 3.0 | 1.5 | 1.0 | 0   |

---

- (17) Find the inverse LTs for the following system transfer functions:

(a)  $G(s) = \frac{100}{s^2 + 5s + 200}$

(b)  $G(s) = \frac{200}{s(s+5)(s-4)}$ .

- (18) A control system can be described by the second-order differential equation:

$$\frac{d^2\theta_0}{dt^2} + 2\xi\omega \frac{d\theta_0}{dt} + \omega^2\theta_0 = k\omega^2\theta_i.$$

Solve the equation for  $\frac{\theta_0}{\theta_i}$  using LTs, given that  $\xi = 0.7$

(damping ratio), the angular velocity ( $\omega$ ) applied to the system is 15 rad/s and  $k = 4$ .

$$\frac{\theta_0}{\theta_i} = \text{the gain of system,}$$

where  $\theta_0$  = magnitude of output  
 $\theta_i$  = magnitude of input.

This Page Intentionally Left Blank

# Appendix

## Mapping of contents to units and reference material

### 1. Business management techniques

| <i>Outcome</i> | <i>Topic</i>                    | <i>Pages</i> | <i>Reference material</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|----------------|---------------------------------|--------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.1            | Managing work activities        | 1–22         | Cather, Morris and Wilkinson, <i>Business Skills for Engineers and Technologists</i> (Butterworth-Heinemann, 2001) – Chapter 1 and 2.                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 1.2            | Costing systems and techniques  | 22–36        | <p>Henderson, Illige, and McHardy, <i>Management for Engineers</i> (Butterworth-Heinemann, 1994) – Chapter 8.</p> <p>Curtis, T., <i>Business and Marketing for Engineers and Scientists</i> (McGraw-Hill, 1994) – Chapter 9.</p> <p>Microsoft Excel spreadsheet models for various costing exercises can be downloaded from: <a href="http://www.key2study.com">www.key2study.com</a></p>                                                                                                                                                                              |
| 1.3            | Financial planning and control  | 36–48        | <p>Henderson, Illige, and McHardy, <i>Management for Engineers</i> (Butterworth-Heinemann, 1994) – Chapters 6, 7, and 9.</p> <p>Curtis, T., <i>Business and Marketing for Engineers and Scientists</i> (McGraw-Hill, 1994) – Chapter 9.</p> <p>Lock, D., <i>Handbook of Engineering Management</i> (Newnes, 1992) – Chapters 19 and 20 provide information on costing and budgeting processes.</p> <p>Microsoft Excel spreadsheet models for various budgetary planning exercises can be downloaded from: <a href="http://www.key2study.com">www.key2study.com</a></p> |
| 1.4            | Project planning and scheduling | 48–58        | <p>Curtis, T., <i>Business and Marketing for Engineers and Scientists</i> (McGraw-Hill, 1994) – Chapter 9.</p> <p>Microsoft Excel spreadsheet models for various budgetary planning exercises can be downloaded from: <a href="http://www.key2study.com">www.key2study.com</a></p> <p>A Visio stencil containing symbols for constructing network diagrams can be downloaded from: <a href="http://www.key2study.com">www.key2study.com</a></p>                                                                                                                        |

## 2. Engineering design

| <i>Outcome</i> | <i>Topic</i>                               | <i>Pages</i> | <i>Reference material</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|----------------|--------------------------------------------|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2.1            | The design specification                   | 60–70        | <p>Pugh, <i>Total Design</i> (Addison Wesley, 1991) – Chapter 3.</p> <p>BS8888:2002, <i>Technical Product Documentation (TPD)</i>. Specification for defining, specifying and graphically representing products.</p> <p>Pahl and Beitz, <i>Engineering Design</i> (Springer, 1995) – Chapters 3–5 provide information on general design process and formulating specification.</p> <p>BS 7373-1:2001, <i>Guide to the Preparation of Specifications</i> (British Standards Institute).</p> <p>BS7373-2:2001, <i>Guide to identifying criteria for a product</i> (British Standards).</p>                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 2.2            | The design report                          | 70–82        | <p>Pahl and Beitz, <i>Engineering Design</i> (Springer, 1996) – Chapters 4 and 5.</p> <p>BS 4811, <i>The Presentation of Research and Development Reports</i> (British Standards Institute).</p> <p>Rudd, D., <i>Report Writing, a Guide to Organisation and Style</i>, at: <a href="http://www.bolton.ac.uk/learning/pubs/csu/reportw.pdf">http://www.bolton.ac.uk/learning/pubs/csu/reportw.pdf</a></p> <p>Norman, <i>Advanced Design and Technology</i> (Addison Wesley Longman, 2nd ed., 1995) – Chapter 2.</p> <p>Lock, D., <i>Handbook of Engineering Management</i> (Newnes, 1992) – Chapters 19 and 20 provide information on costing and budgeting processes.</p>                                                                                                                                                                                                                                                                                                                                |
| 2.3            | Computer technology and the design process | 83–101       | <p>Giesecke, F.E., et al., <i>Modern Graphics Communication</i> (Prentice Hall, 2003) and Giesecke, F.E., et al., <i>Engineering Graphics</i> (Prentice Hall, 2003).</p> <p>Lock, D., <i>Handbook of Engineering Management</i> (Newnes, 1992) – Chapter 31 provides information on applying computers to engineering.</p> <p>Grimston, C., <i>PC Maintenance</i> (Arnold, 1996) – Chapters 1–3 provide a useful introduction to computer hardware and microprocessor architecture.</p> <p>Moeler, Phillips, and Davis, <i>Project Management with CAM, PERT and Precedence Diagramming</i> – Chapter 2 provides an introduction to networks.</p> <p>Access to CAD and analytical packages for engineering, such as Pro-Engineer AutoCAD, CadKey, MathCAD, and CODAS.</p> <p>Software packages for the computer analysis of structural systems, such as Analysis by Ing. Frank Crylaerts, can be downloaded from: <a href="http://www.club.innet.be/~year1335">http://www.club.innet.be/~year1335</a></p> |

The student's version of QuickField, a software package for field analysis, can be downloaded from: [http:// www.tor.ru/quickfield](http://www.tor.ru/quickfield) and also from Simtelnet at: <http://www.simtel.net/pub/simtelnet/Win95/engin>

A Visio stencil containing symbols for constructing network diagrams can be downloaded from: [www.key2study.com](http://www.key2study.com)

### 3. Engineering science

| <i>Outcome</i> | <i>Topic</i>                           | <i>Pages</i> | <i>Reference material</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|----------------|----------------------------------------|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3.1            | Static engineering systems             | 102–139      | <p>Hearn, E.J., <i>Mechanics of Materials – Volume 1</i> (Butterworth-Heinemann, 3rd ed., 1997) – Chapters 3 and 4 provide comprehensive coverage of shear force, bending moments, and bending theory; Chapter 8 provides coverage of torsion.</p> <p>Meriam, J.L., <i>Engineering Mechanics: Statics</i> VI, L.G. Kraige. John Wiley (WIE, 2001).</p> <p>W. Morgan et al., <i>Structural Mechanics</i> (Longman, 1996). A thoroughly comprehensive coverage of bending of beams and columns, including beam selection and design.</p>                                            |
| 3.2            | Dynamic engineering systems            | 140–174      | <p>Bacon and Stephens, <i>Mechanical Technology</i> (Butterworth-Heinemann, 3rd ed., 1998) – Chapters 1–5, 10, and 18.</p> <p>Bolton, W., <i>Higher Engineering Science</i> (Elsevier, 2004). This book covers both static and dynamic systems.</p> <p>Hibbeler, R., <i>Engineering Mechanics: Dynamics Pack</i> (Prentice Hall, 2004).</p>                                                                                                                                                                                                                                       |
| 3.3            | DC and AC theory                       | 175–249      | <p>Bird, J., <i>Higher Electrical Technology</i> (Newnes, 2nd ed., 1996) – Chapters 1–4.</p> <p>Hughes, <i>Electrical Technology</i> (Addison Wesley Longman, 7th ed., 1995) – Chapters 10–14.</p> <p>Silvester, P., <i>Electric Circuits</i> (Macmillan, 1993) – Programmes 4 and 5.</p> <p>Bolton, W., <i>Engineering Science</i> (Newnes, 3rd ed., 1998) – Chapter 16 provides a useful introduction to AC circuits; Chapter 18 develops this further.</p> <p>Tooley, M., <i>Electronic Circuits, Fundamentals and Applications</i>, 2nd edn (Newnes, 2002). Chapters 1–4.</p> |
| 3.4            | Information and energy control systems | 249–284      | <p>Hughes, <i>Electrical Technology</i> (Addison Wesley Longman, 7th ed., 1995) – Chapters 17–19 and 29.</p> <p>Mazda, F., <i>Power Electronics Handbook</i> (Newnes, 3rd ed., 1997) – this book provides numerous examples of electrical power and energy controllers. It also contains a useful introduction to a variety of semiconductor power control devices.</p>                                                                                                                                                                                                           |

Tooley, M., *PC-based Instrumentation and Control* (Newnes, 2nd ed., 1995) – this book provides numerous examples of interfacing sensor, transducers, and other control devices to personal computers.

Bolton, W., *Engineering Science* (Newnes, 3rd ed., 1998) – Chapter 1 provides a general introduction to engineering systems.

## 4. Electrical and electronic principles

| <i>Outcome</i> | <i>Topic</i>                           | <i>Pages</i> | <i>Reference material</i>                                                                                                                                                                                                                                                                                                                                                                                      |
|----------------|----------------------------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4.1            | Circuit theory                         | 285–331      | <p>Bird, J., <i>Higher Electrical Technology</i> (Newnes, 2nd ed., 1996) – Chapter 8.</p> <p>Bird, J., <i>Electrical Circuit Theory and Technology</i> (Newnes, 1997) – Chapters 11, 13, and 27–32.</p> <p>Hughes, <i>Electrical Technology</i> (Addison Wesley Longman, 7th ed., 1995) – Chapters 4, 13, and 33.</p> <p>Silvester, P., <i>Electric Circuits</i> (Macmillan, 1993) – Programmes 1–3 and 5.</p> |
| 4.2            | Networks                               | 331–342      | <p>Bird, J., <i>Higher Electrical Technology</i> (Newnes, 2nd ed., 1996) – Chapters 7, 9, and 13.</p> <p>Bird, J., <i>Electrical Circuit Theory and Technology</i> (Newnes, 1997) – Chapters 27 and 38.</p> <p>Hughes, <i>Electrical Technology</i> (Addison Wesley Longman, 7th ed., 1995) – Chapters 4, 15, and 30.</p>                                                                                      |
| 4.3            | Complex waves                          | 342–354      | <p>Bird, J., <i>Higher Electrical Technology</i> (Newnes, 2nd ed., 1996) – Chapter 10.</p> <p>Bird, J., <i>Electrical Circuit Theory and Technology</i> (Newnes, 1997) – Chapters 23, 33, and 34.</p> <p>A Microsoft Excel spreadsheet model for Fourier analysis can be downloaded from: <a href="http://www.key2study.com">www.key2study.com</a></p>                                                         |
| 4.4            | Transients in $R$ – $L$ – $C$ circuits | 355–362      | <p>Silvester, P., <i>Electric Circuits</i> (Macmillan, 1993) – Programme 7.</p> <p>Bird, J., <i>Electrical Circuit Theory and Technology</i> (Newnes, 1997) – Chapter 42.</p>                                                                                                                                                                                                                                  |

## 5. Mechanical principles

| <i>Outcome</i> | <i>Topic</i>            | <i>Pages</i> | <i>Reference material</i>                                                                                                     |
|----------------|-------------------------|--------------|-------------------------------------------------------------------------------------------------------------------------------|
| 5.1            | Complex loading systems | 363–373      | <p>Bolton, W., <i>Mechanical Science</i> (Blackwell, 1995) – Chapter 8 provides information on complex stress and strain.</p> |

|     |                              |         |                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-----|------------------------------|---------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     |                              |         | Hearn, E.J., <i>Mechanics of Materials – Volume 1</i> (Butterworth-Heinemann, 3rd ed., 1997) – Chapters 2 and 4 cover all the requirements of this outcome.                                                                                                                                                                                                                                                                         |
|     |                              |         | Gere and Timoshenko, <i>Mechanics of Materials</i> , 5th edn (Brooks Cole, 2000). Covers all aspects of complex loading systems.                                                                                                                                                                                                                                                                                                    |
| 5.2 | Loaded beams and cylinders   | 373–385 | Hearn, E.J., <i>Mechanics of Materials – Volume 1</i> (Butterworth-Heinemann, 3rd ed., 1997) – Chapters 9 and 10.<br><br>Benham and Crewford, <i>Mechanics of Engineering Materials</i> (Addison Wesley Longman, 2nd ed., 1996) – Chapters 2 and 15.                                                                                                                                                                                |
| 5.3 | Pressure vessels             | 385–397 | BS PD 5500:2003, <i>Specification for Unrefined Fusion Welded Pressure Vessels</i> (Access through Institute Library Service).<br><br>Hearn, E.J., <i>Mechanics of Materials</i> , 3rd edn (1997 Butterworth-Heinemann). Good coverage of thin and thick walled cylinders.<br><br>Benham, P.P., et al., <i>Mechanics of Engineering Materials</i> (Prentice Hall, 1996). More advanced coverage of thin and thick walled cylinders. |
| 5.4 | Dynamics of rotating systems | 397–430 | Waldron, K. and Kinzel, G., <i>Kinematics, Dynamics and Design of Machinery</i> (John Wiley, 2003). This book uses MATLAB to solve realistic engineering problems. It is, however, expensive.<br><br>Wilson, C. and Sadler, J., <i>Kinematics and Dynamics of Machinery</i> (Prentice Hall, 2003). This book covers all aspects of dynamic engineering systems, if perhaps at a slightly advanced level.                            |

## 6. Analytical methods

| <i>Outcome</i> | <i>Topic</i> | <i>Pages</i> | <i>Reference material</i>                                                                                                                                                                                                                                                                                     |
|----------------|--------------|--------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 6.1            | Algebra      | 432–459      | Tooley, M. and Dingle, L., <i>BTEC National Engineering</i> (Newnes, 2002). Unit 4 provides an introduction for those with limited knowledge.<br><br>Bolton, W., <i>Essential Mathematics for Engineering</i> (Newnes, 1997) – Chapter 2 provides a useful introduction to polynomials and partial fractions. |
| 6.2            | Trigonometry | 459–477      | Tooley, M. and Dingle, L., <i>BTEC National Engineering</i> (Newnes, 2002). Unit 4 provides an introduction for those with limited knowledge.<br><br>Bird, J., <i>Higher Engineering Mathematics</i> 4th edn (Elsevier 2004). Covers the trigonometric functions.                                             |



## 630 Appendix

- |     |                            |         |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-----|----------------------------|---------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 6.3 | Calculus                   | 477–515 | <p>Bird, J., <i>Higher Engineering Mathematics</i> 4th edn (Elsevier, 2004). Covers all aspects of the calculus.</p> <p>Bolton, W., <i>Essential Mathematics for Engineering</i> (Newnes, 1997) – Chapters 22–28.</p> <p>Stroud, K.A., <i>Engineering Mathematics</i>, 5th edn (Palgrave Macmillan, 2001). Covers all aspects of calculus.</p> <p>Yates, J.C., <i>National Engineering Mathematics – Volume 3</i> (Macmillan, 1996) – Chapters 4–10.</p> |
| 6.4 | Statistics and probability | 515–585 | <p>Bird, J., <i>Higher Engineering Mathematics</i> (Newnes, 3rd ed., 1998) – Chapters 33–36 deal with binomial, Poisson and normal distributions, linear correlation, and linear regression.</p> <p>Bolton, W., <i>Essential Mathematics for Engineering</i> (Newnes, 1997) – Chapters 42–45 cover probability and statistics.</p> <p>Dodson, B. and Nolan, D., <i>Reliability Engineering Handbook</i> C Marcel Dekker, 1999).</p>                      |
| 6.5 | Advanced topics            | 585–623 | <p>Attenborough, M., <i>Mathematics for Electrical Engineering and Computing</i> (Elsevier, 2003). Good coverage of Laplace and Fourier Series.</p> <p>Stroud, K.A. and Booth, D.J., <i>Advanced Engineering Mathematics</i> (Palgrave Macmillan, 2003). Covers all aspects of advanced topics.</p>                                                                                                                                                      |

# Abbreviations

2D, two-dimensional  
3D, three-dimensional  
4PCO, four-pole changeover  
  
AC, alternating current  
ACWM, anticlockwise moments  
ADC, analogue-to-digital converter  
AF, audio frequency  
ALU, arithmetic and logic unit  
AM, amplitude modulated  
ASCII, American standard code for information interchange  
  
BCD, binary coded decimal  
BJT, bipolar junction transistor  
BM, bending moment  
BS, British Standards  
  
CAD, computer-aided design  
CAM, computer-aided manufacture  
CdS, calcium sulphide  
CIMS, computer-integrated manufacturing system  
CNC, computer numerical controlled  
COSHH, control of substance hazardous to health  
CPM, critical path method  
CPU, central processing unit  
C-R, capacitor and resistor  
CRT, cathode ray tube  
CWM, clockwise moment  
  
DAC, digital-to-analogue converter  
DC, direct current  
DE, differential equation  
DFM, digital frequency meter  
DPCO, double-pole changeover  
DPM, digital panel meter  
DVST, direct view storage tube  
  
EMF, electromotive force  
ESC, control character

FC, fixed cost  
FEA, finite element analysis  
FEM, finite element model  
FET, field effect transistor  
FM, frequency modulated  
FMS, flexible manufacturing system  
FO, fixed overheads  
  
HSE, health and safety executive  
  
I/O, input/output  
  
JIT, just-in-time  
  
KE, kinetic energy  
  
LCD, liquid crystal display  
L-C-R, inductor, capacitor and resistor  
LDR, light dependent resistor  
LED, light emitting diode  
LHS, left-hand side  
L-R, inductor and resistor  
LSB, least-significant bit  
LT, Laplace transform  
LVDT, linear variable differential transformer  
  
MOS, metal oxide semiconductor  
MSB, most-significant bit  
MTBF, mean time between failure  
  
ODE, ordinary differential equation  
  
PC, personal computer  
PCB, printed circuit board  
PDS, product design specification  
PE, potential energy  
PERT, programme evaluation and review technique  
PF, partial fraction  
PLA, programmed logic array  
PLC, programmable logic controllers  
PLD, programmed logic device  
PROM, programmable read-only memory  
PTFE, polytetrafluoroethylene  
PWM, pulse width modulated  
  
QC, quality control  
  
R&D, research and development  
RAM, random access memory  
RF, radio frequency  
RHS, right-hand side  
RMS, root mean square  
ROCE, return on capital employment  
ROM, read-only memory  
RPM, revolutions per minute  
  
SF, shear force  
SHM, simple harmonic motion  
S/N, signal-to-noise

SNR, signal-to-noise ratio  
SPCO, single-pole changeover  
SV, stationary value  
  
TC, total cost  
TP, turning point  
TQ, total quality  
TQM, total quality management  
  
UDL, uniformly distributed load  
  
VA, volt-amperes  
VC, variable cost  
VDU, visual display unit  
  
WD, work done

This Page Intentionally Left Blank

# Answers

## Chapter 1

### Question 1.2.1

(page 26)

Hardware, £2.85; Semiconductors, £3.33;  
Passive components, £3.51; Miscellaneous, £2.41;  
Total cost, £11.83 (overall reduction in cost = 16 p).

### Question 1.2.2

(page 30)

Manual data switch, £37; automatic data switch, £85.

### Questions 1.3.1

(page 41)

12 month figures are as follows:

- |                                                                                                               |                                                                                                               |
|---------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|
| (a) Battery charger, £66,000<br>Trolley jack, £466,500<br>Warning triangle, £88,200<br><b>Total, £620,700</b> | (b) Battery charger, £69,000<br>Trolley jack, £464,250<br>Warning triangle, £82,000<br><b>Total, £615,250</b> |
| (c) Battery charger, £40,000<br>Trolley jack, £390,000<br>Warning triangle, £73,000<br><b>Total, £503,000</b> | (d) Battery charger, £58,500<br>Trolley jack, £431,750<br>Warning triangle, £79,625<br><b>Total, £569,875</b> |

### Questions 1.3.2

(page 43)

Totals are as follows:

- |                                                                                                                   |                                                                                                                    |
|-------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| (a) Battery charger, £22,456.5<br>Trolley jack, £161,540<br>Warning triangle, £32,227.5<br><b>Total, £216,224</b> | (b) Battery charger, £19,906.5<br>Trolley jack, £152,150<br>Warning triangle, £32,227.5<br><b>Total, £204,284</b>  |
| (c) Battery charger, £21,337.5<br>Trolley jack, £152,150<br>Warning triangle, £25,770<br><b>Total, £199,257.1</b> | (d) Battery charger, £18,787.13<br>Trolley jack, £161,540<br>Warning triangle, £25,770<br><b>Total, £206,097.1</b> |

**Questions 1.3.3**

(page 44)

Total labour costs are as follows:

- |                              |                                |
|------------------------------|--------------------------------|
| (a) Battery charger, £15,675 | (b) Battery charger, £13,537.5 |
| Trolley jack, £113,281.25    | Trolley jack, £113,281.25      |
| Warning triangle, £11,531.25 | Warning triangle, £10,890.63   |
| (c) Battery charger, £15,675 | (d) Battery charger, £13,537.5 |
| Trolley jack, £122,343.75    | Trolley jack, £122,343.75      |
| Warning triangle, £10,890.63 | Warning triangle, £11,531.25.  |

**Questions 1.3.4**

(page 46)

Total overheads are as follows:

- |                |                |                |                  |
|----------------|----------------|----------------|------------------|
| (a) £98,094.38 | (b) £98,894.38 | (c) £97,623.81 | (d) £101,823.81. |
|----------------|----------------|----------------|------------------|

**Question 1.3.5**

(page 46)

Additional semi-variable cost; cost of plant and equipment maintenance.

Additional fixed cost; energy cost for factory and office lighting and heating.

**Questions 1.4.1**

(page 52)

1. Critical path; B1, B2, C, D, F.
2. Critical path; C, D, H, J,  $t_e = 15$  days.

(Please see the Tutor Resource Pack for a model answer with critical path diagram.)

**Questions 1.4.2–1.4.4**

(pages 55–58)

Please see the Tutor Resource Pack for model answers.

**Chapter 3****Problems 3.1.1**

(page 112)

1.  $R_B = 51.25 \text{ kN}$ ,  $R_A = 18.75 \text{ kN}$
2.  $\text{BM} = -300 \text{ N}$
3.  $R_A = 25.333 \text{ kN}$ ,  $R_B = 26.666 \text{ kN}$ ,  $\text{BM} = 56 \text{ kNm}$
5.  $R_A = 120 \text{ kN}$ ,  $R_B = 100 \text{ kN}$ , From  $R_A$ ; SF = 120, 95, -25, -80 (kN),  
BM at left-hand support = 268.75 kNm,  
BM at right-hand support = -20 kNm.

**Test Your Knowledge**

(page 117)

- |                             |                                            |
|-----------------------------|--------------------------------------------|
| 1. (i) 162 kNm.             | 2. (iii) BM = 450 kNm.                     |
| (ii) 1350 cm <sup>3</sup> . | (iv) $\sigma_{\max} = 128.6 \text{ mPa}$ . |

### Problems 3.1.2

(pages 132–133)

1. (a) 73.71 mm from base  
(b)  $19.48 \times 10^6 \text{ mm}^4$
2. (a) 3 kN at each support  
(b)  $11.5 \times 10^6 \text{ mm}^4$   
(c) 12.0 kN m  
(d)  $62.5 \text{ N/mm}^2$
3. (a)  $2.011 \times 10^6 \text{ mm}^4$   
(b) 4.022 kN m  
(c) 2681 N
4. 12.4 kN
5.  $13.02 \text{ MN/m}^2$
6. 42.384 kN m
7. (a) 116.6 mm  
(b) 24
8. (a) 36.17 mm  
(b) 58.1  
(c) 2.73 MN

### Problems 3.1.3

(page 140)

1. 6.6267 kN m  
2.248 MW  
 $\tau_{\max} = 49.73 \text{ MN/m}^2$ ,  $\theta = 0.9^\circ$   
 $\tau_{\max} = 51.49 \text{ MN/m}^2$ ,  $\tau_{\min} = 41.19 \text{ MN/m}^2$ ,  $\theta = 1.57^\circ$

### Problems 3.2.1

(pages 152–153)

1.  $\omega = 50 \text{ rad/s}$ ,  $v = 13.5 \text{ m/s}$
2. (a)  $\alpha = 0.33 \text{ rad/s}$   
(b) 1800 rad or 286.48 revs
3. (a) 1.57 rad/s  
(b) 204.4 Nm
4. 53.3 rad/s
5.  $35.07 \text{ kg m}^2$
6. retardation =  $2.2 \text{ rad/s}^2$ , braking torque = 2.475 Nm

### Problems 3.2.2

(pages 164–165)

1. 400 W
2. 0.36 J
3. (i) 67.5 kJ  
(ii) 3 m/s
4. 123.46 kJ
5. 1.004 MJ
6. (a) 12.9 kph  
(b) 4070 J
7. (a)  $13.09 \text{ rad/s}^2$   
(b) 0.53 Nm

### Problems 3.2.3

(page 174)

1. Time period = 0.513 s, velocity = 0.92 m/s.
3. Velocity = 0.105 m/s.
4. Velocity = 1.57 m/s, acceleration =  $9.87 \text{ m/s}^2$ .
5. Velocity of piston = 2.45 m/s, acceleration of piston =  $53.5 \text{ m/s}^2$ .



### Problems 3.3.1

(page 184)

1. 4 A
2. 89.1 V
3. 1.25 A, 7.5 V
4.  $10.5 \Omega$
5. 5.5 W
6. 18.144 kJ

### Problems 3.3.2

(page 195)

1.  $0.25 \mu\text{F}$
2. 100 V
3. 956.2 pF
4. (a)  $2.5 \mu\text{F}$  (b)  $4 \mu\text{F}$  (c)  $5 \mu\text{F}$
5. 1.97 V

### Problems 3.3.3

(page 200)

1. 553 turns
2. 88.5 mH
3. 1.15 A
4.  $-120 \text{ V}$

### Problems 3.3.4

(page 211)

1. 27.5 V
2. 550 turns
3. 1.023 A
4. 5.5%
5. 96.1%

### Problems 3.3.5

(page 221)

1. (a) 250 mA; (b) 176.75 mA; (c) 100 Hz; (d) 10 ms; (e) 237 V.
2.  $v = 160 \sin(377t)$ , 144.8 V.
3.  $i = 4.713 \sin(2513t)$ , 3.33 A.
4. (a)  $585.43 \Omega$ ; (b)  $11.709 \Omega$ .
5. (a)  $7.536 \Omega$ ; (b)  $1,507.2 \Omega$ .
6.  $233.24 \Omega$ , 0.857 A.
7.  $85.32 \Omega$ , 1.29 A.
8.  $V_C = 21.44 \text{ V}$ ,  $V_R = 10.8 \text{ V}$ .
9. 159 kHz.
10.  $13.63 \mu\text{H}$ .

### Problems 3.3.6

(page 234)

1.  $V = 200 + j240 \text{ V}$ .
2.  $I = 2 + j5 \text{ A}$ .

3.  $V = 0.4 = j0.75 \text{ V}$ .
4.  $I = 0.04 = j0.0126 \text{ A}$ .
5.  $V_L = 40 \text{ V}$ ,  $V_C = 40 \text{ V}$ ,  $V_R = 0.3 \text{ V}$ .
6.  $f_0 = 47.7 \text{ kHz}$ ,  $Z_d = 1,000 \Omega$ .
7.  $C_{\max} = 1.25 \text{ nF}$ ,  $C_{\min} = 312.5 \text{ pF}$ .
8.  $f_{\max} = 711 \text{ kHz}$ ,  $Q_{\max} = 89.35$ .  
 $f_{\min} = 224 \text{ kHz}$ ,  $Q_{\min} = 28.15$ .

### Problems 3.3.7

(page 242)

1. Reactive component = 1.75 A, active component = 1.785 A.
2. Power factor = 0.6, apparent power = 1.148 kW.
3. True power = 662.9 W, reactive power = 499.5 W.
4. Power factor = 0.193, capacitance = 24.4  $\mu\text{F}$ .
5. Power factor = 0.178, current = 0.459 A.
6. Power factor = 0.606, phase angle = 52.7°.
7. Power factor = 0.598, current = 6.313 A.
8. Capacitance = 116.7  $\mu\text{F}$ .

### Question 3.3.1

(page 221)

- (a and b) Please see Tutor Resource Pack for model answers.  
 (c)  $v = 15.14 \text{ V}$ ,  $i = 0.22 \text{ A}$ .

### Question 3.3.2

(page 222)

- (a)  $V_L = 1.036 \text{ V}$ ; (b)  $V_C = 3.388 \text{ V}$ ; (c)  $V_R = 1.65 \text{ V}$ ;  
 (d)  $Z = 57.46 \Omega$ ; (e)  $V = 2.873 \text{ V}$ ; (f)  $\phi = -55^\circ$ .

## Chapter 4

### Questions 4.1.1

(page 300)

1.  $0.4 \Omega$ , 20 V.
2.  $-18 \text{ V}$  in series with  $3 \Omega$ .
3. 3.692 A in parallel with  $4.333 \Omega$ .
4. 0.98 A (Thévenin equivalent 2.5 V in series with  $1.55 \Omega$ ).
5. 4.62 V (Norton equivalent 2.7 A in parallel with  $10.91 \Omega$ ).
6. 8.076 V.
7.  $R_A = 206.7 \Omega$ ,  $R_B = 310 \Omega$ ,  $R_C = 124 \Omega$ .
8.  $R_1 = 40 \Omega$ ,  $R_2 = 10 \Omega$ ,  $R_3 = 8 \Omega$ .
9.  $R_1 = 6.67 \Omega$ ,  $R_2 = 3.33 \Omega$ ,  $R_3 = 4.44 \Omega$ .
10.  $R_A = 37 \Omega$ ,  $R_B = 29.6 \Omega$ ,  $R_C = 24.67 \Omega$ .

### Questions 4.1.2

(page 306)

1. (a)  $33.54 \angle 63.44^\circ \Omega$ ; (b)  $12 \angle 90^\circ \Omega$ ; (c)  $0.5 \angle -53.13^\circ \Omega$ ;  
 (d)  $1000 \angle -90^\circ \Omega$ .
2. (a)  $12.5 + j21.65 \Omega$ ; (b)  $88.39 - j88.39 \Omega$ ;  
 (c)  $56.72 + j647.52 \Omega$ .

3.  $0.618 - j0.379 \text{ A}$  or  $0.725 \angle -31.52^\circ \text{ A}$ .
4. (a)  $1.118 \angle 26.57^\circ \text{ A}$ ; (b)  $0.988 \angle -18.44^\circ \text{ A}$ .  
(c)  $0.771 \angle 40.6^\circ \text{ A}$ .
5.  $Z_1 = Z_2 = -j120$ ,  $Z_3 = j160$ .
6.  $Z_A = Z_B = Z_C = 30 + j30$ .
7.  $1.212 + j0.308 \text{ A}$ .

### Questions 4.1.3

(page 311)

1. Turns ratio, 22.4:1; Primary voltage = 141.1 V.
2.  $8 - j2 \Omega$ .

### Questions 4.1.4

(page 316)

1.  $I_1 = 0.118 - j0.344 \text{ A}$ ,  $I_2 = -0.123 - j0.615 \text{ A}$ ,  
 $I_3 = 0.241 + j0.271 \text{ A}$ .
2.  $V_A = 5.05 - j3.21 \text{ V}$ ,  $V_B = 0.08 + j0.72 \text{ V}$ .

### Questions 4.1.5

(page 325)

1.  $18 \mu \text{Wb}$ .
2.  $-40 \text{ V}$ .
3. 0.133.
4. 0.26.
5. 0.2 A.
6.  $1.146 \text{ mWb}$ , 20% increase.
7. (a) 55 V; (b) 5 A; (c) 275 W.

### Questions 4.1.6

(page 330)

1. (a) 237 kHz; (b) 22.3; (c) 10.62 kHz.
2.  $I = 0.05 - j1.825 \text{ mA}$ ,  $V_L = 2578 + j71 \text{ mV}$ ,  
 $V_C = -580 - j16 \text{ mV}$ ,  $V_R = 1.5 - j54.75 \text{ mV}$ .
3. (a) 12 mA; (b) 40.5 kHz.
4. 22.75.
5. (a) 425.6 kHz, 140.3, 3.03 kHz; (b) 45.5 kHz.

### Questions 4.2.1

(page 341)

1. (a)  $32.25 \Omega$ ; (b)  $53.3 \Omega$ ; (c)  $35.8 + j27.9 \Omega$ .
2.  $R_1 = 40.9 \Omega$ ,  $R_2 = 10.1 \Omega$ .
3.  $R_1 = 1.273 \text{ k}\Omega$ ,  $R_2 = 945.8 \Omega$ .
4.  $\gamma = 0.672 + j0.087$ .
5. 2.08 Nepers, 18.06 dB.

### Questions 4.3.1

(page 347)

$$(b) y = 1 - \frac{2}{\pi} \sin \omega t - \frac{1}{\pi} \sin 2\omega t - \frac{2}{3\pi} \sin 3\omega t - \dots$$

$$(d) \ y = \frac{2}{\pi} - \frac{4}{3\pi} \cos 2\omega t - \frac{4}{15\pi} \cos 4\omega t - \dots + \sin \omega t$$

$$(e) \ y = \frac{20}{\pi} \cos \omega t - \frac{20}{3\pi} \cos 3\omega t + \frac{20}{5\pi} \cos 5\omega t - \dots$$

### Questions 4.3.2

(page 350)

Fourier constants are as follows:  $A_0 = 4.55$ ,  $A_1 = -4.79$ ,  
 $A_2 = -1.85$ ,  $B_1 = 4.79$ ,  $B_2 = 1.79$ , etc.

### Questions 4.3.3

(page 354)

- (a) 100 W, 16 W, 4 W; (b)  $i = 2 \sin \omega t + 0.8 \sin(2\omega t - \pi/2) + 0.4 \sin(4\omega t - \pi/2)$ ; (c)  $V_{\text{rms}} = 77.46 \text{ V}$ ,  $I_{\text{rms}} = 1.55 \text{ A}$ ; (d) 120 W.
- (a) 21.87 V; (b) 0.715 A; (c) 15.28 W; (d) 15.63 W; (e) 0.977.
- (a)  $i = 100 \sin(314t - 0.64) + 0.51 \sin(942t - 1.15) - 0.26 \sin(1570t - 1.31) \text{ A}$ .  
 (b) 118.74 V; (c) 1.47 A;  
 (d) 87.93 W; (e) 174.5 W; (f) 0.504.

### Questions 4.4.1

(page 362)

- $I(s) = \frac{0.354s + 70.7}{s^2 + 10^4}$ .
- $i = e^{-40t}$ , 0.607 mA.
- $i = 0.6(1 - e^{-5t})$
- $i = e^{-2t} \sin(14t)$ .

## Chapter 5

### Questions 5.1.1

(page 365)

- Change in length = 0.05 mm (extension),  
 Change in dimension on 60 mm surface = 0.0036 mm (compression),  
 Change in dimension on 25 mm surface = 0.0015 mm (compression).
- Change in length = 0.0714 mm.
- $\sigma_1 = 35.32 \text{ MPa}$ ;  $\sigma_2 = 17.96 \text{ MPa}$ .

### Questions 5.1.2

(page 373)

- $\sigma_1 = 95.3 \text{ MPa}$ ,  $\sigma_2 = 72.69 \text{ MPa}$ .
- $\nu = 0.212$ ,  $K = 46.3 \text{ GPa}$ .
- $\nu = 0.303$ ,  $E = 57.87 \text{ GPa}$ ,  $G = 22.2 \text{ GPa}$ ,  $K = 48.96 \text{ GPa}$ .

### Questions 5.2.1

(page 384)

- Change in diameter = 17.7  $\mu\text{m}$ ; change in length = 16.67  $\mu\text{m}$ .
- Maximum safe gas pressure = 1375 kPa.

3. (i)  $P = 427.9 \text{ kPa}$ .  
(ii)  $\sigma_h = 34.23 \text{ MPa}$ ,  $\sigma_r = 17.12 \text{ MPa}$  (compression).
4.  $\sigma_h = 28.4 \text{ MPa}$ ,  $\sigma_r = 15 \text{ MPa}$  (compression).
5. (i) At inner surface:  $\sigma_h = 38 \text{ MPa}$ ,  $\sigma_r = 70 \text{ MPa}$  (compression).  
At outer surface:  $\sigma_h = 8 \text{ MPa}$ ,  $\sigma_r = 40 \text{ MPa}$  (compression).  
(ii)  $\sigma_L = 16 \text{ MPa}$ .

### Questions 5.3.1

(page 396)

### Questions 5.4.1

(page 429)

1. (i)  $\theta_1 = 168.52^\circ$ ,  $\theta_2 = 191.48^\circ$ ; (ii)  $2.95 \text{ kW}$ .
2. (i)  $T_1 = 962 \text{ N}$ ,  $T_2 = 143 \text{ N}$ ; (ii)  $6.1 \text{ MPa}$ .
3.  $112 \text{ Nm}$ .
4. (i)  $t_A = 5$ ; (ii)  $1466 \text{ W}$ ; (iii)  $1173 \text{ W}$  (iv)  $1454 \text{ N}$ .
5.  $T_1 = 1769 \text{ N}$ ;  $T_2 = 977 \text{ N}$  and  $L = 4.48 \text{ m}$ .

## Chapter 6

### Questions 6.1.1

(page 457)

1.  $A = 1.63 \times 10^{-3}$
2.  $4.49$  (see worked solution).
3.  $P_2 = 374.28$ .
4.  $P_{\text{noise}} = 0.02$  (see worked solution).
5.  $r = 0.247$ .
6.  $x = 2$ ,  $y = 4$ .
7. Law of graph is  $H = 0.05 V^3$
8. (i)  $\frac{1}{(x+2)(x-3)} = \frac{-1}{5(x+2)} + \frac{1}{5(x-3)}$ ;  
(ii)  $\frac{x^2}{(x+1)(x^2+2x+6)} = \frac{-1}{5(x+1)} + \frac{x+6}{5(x^2+2x+6)}$ ;  
(iii)  $\frac{x^2}{(x+1)(x+2)} = \frac{1}{9(x-1)} + \frac{8}{9(x+2)} - \frac{4}{3(x+2)}$ .
9.  $A = 2$ ,  $b = 2.772$  (to 4 s.f.).
10. We know that the general condition for convergence is that in the limit as  $n \rightarrow \infty$  the modulus of:  
$$\frac{U_n + 1}{U_n} < 1$$
 where  $U_n$  is the  $n$ th term.

Then by expanding each series and applying the above criteria, we are able to establish convergence or other wise.

Thus both of these series are *convergent*.

11. Coefficients are 1, 6, 15, 20, 15, 6, 1 (see worked solution).
12.  $Q$  increases by 6% (see worked solution).
13. Using the exponential series for  $e^x$  we get:

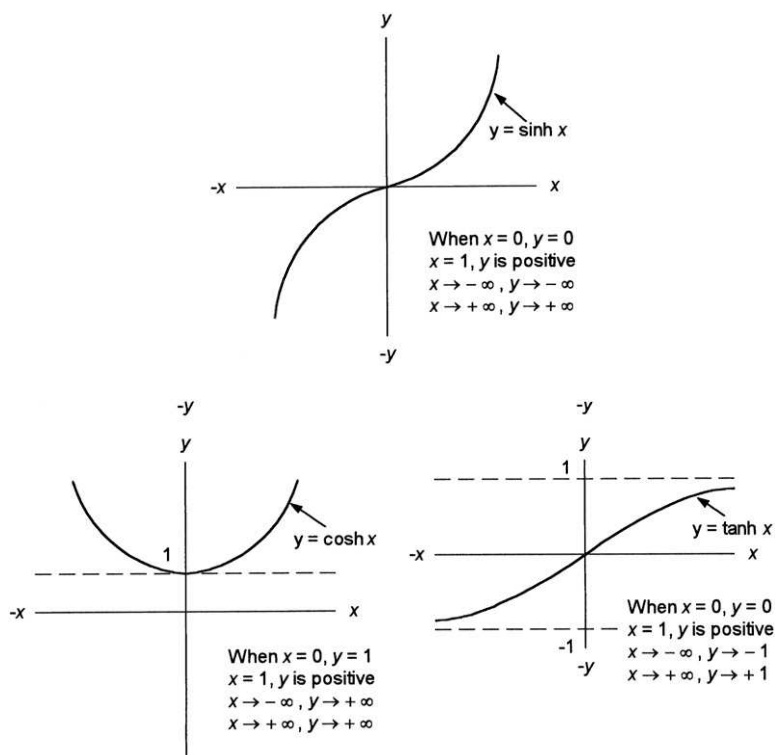
$$x^2 e^{-x} = x^2 - x^3 + \frac{x^4}{2!} - \frac{x^5}{3!} + \frac{x^6}{4!} - \dots$$

$$\text{also } -e^{-x} = -1 + x - \frac{x^2}{2!} + \frac{x^3}{3!} - \frac{x^4}{4!} \dots$$

$$\text{then } x^2 e^{-x} - e^{-x} = -1 + x + \frac{x^2}{2} - \frac{5x^3}{6} + \frac{11x^4}{24} - \dots$$

are the first five terms of required series.

14.



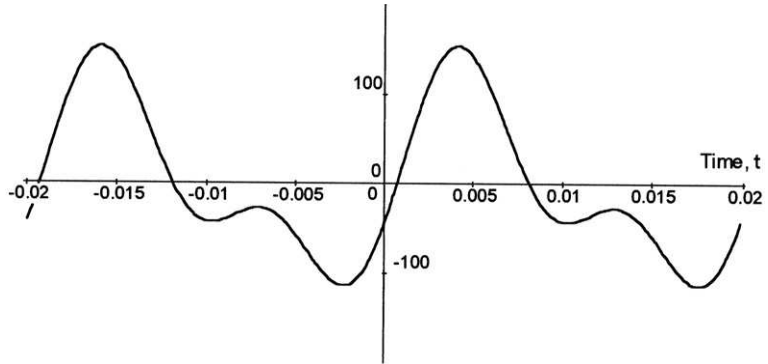
15. (i)  $\sinh x = 1.5$  then  $x = 1.195$  (correct to 4 s.f.)  
 (ii)  $\cosh x = 1.875$  then  $x = 1.242$  (correct to 4 s.f.);  
 (iii)  $\sinh^{-1} 1.375 = x$  then  $x = 1.123$  (correct to 4 s.f.);  
 (iv)  $\tanh x = 0.32$  then  $x = 0.3316$  (correct to 4 d.p.).

## Questions 6.2.1

(page 476)

- (i)  $s = 9.43, A = 84.8$ ; (ii)  $s = 56.5, A = 679$ .
- (i)  $5, 306.9^\circ$ ; (ii)  $1.5, 56.3^\circ$ .
- (i)  $x = 2.828, y = 2, 828$ ; (ii)  $x = \sqrt{3}, y = -1$ .
- (i) Amplitude = 6, phase angle =  $30^\circ$  lagging;  
 (ii) Amplitude = 4, phase angle =  $60^\circ$  lagging.
- (i) Amplitude = 3 A  
 Frequency = 100 Hz  
 Periodic time =  $1/f = 1/100$  s  
 Phase angle =  $45^\circ$  lagging  
 Time to first maximum  $t = 3/800$  s.  
 (ii) Amplitude = 0.7 V  
 Frequency = 200 Hz  
 Period time =  $1/200$  s.  
 Phase angle =  $60^\circ$  leading  
 Time to first maximum =  $1/2400$  s or 0.416 ms.

8. (i) Amplitude = 40 amps  
 Periodic time =  $1/50$  s.  
 Frequency = 50 Hz  
 Phase angle =  $18.3^\circ$  lagging.  
 (ii)  $i = -12.55$  A.  
 (iii)  $t = 1.13$  ms.
- 9.



Peak – peak voltage  $\approx 260$  V.

12.  $0 < \mu < 0.67$ .  
 13. 443.4 kN m.  
 15. (i)  $\sin 6\theta$ ; (ii)  $\cos 7t$ .

### Questions 6.3.1

(page 512)

1. (a)  $6(3x + 1)$ ; (b)  $-\frac{15}{2}(2 - 5x)^{1/2}$ ; (c)  $-8x(4x^2 + 3)^{-2}$ ;  
 (d)  $8 \sin 4x \cos 4x$ ; (e)  $5 \sin(2 - 5x)$ ;  
 (f)  $\frac{1}{x}$ ; (g)  $\frac{2}{3}(1 - 2t)^{-4/3}$ .
2. (a)  $\sin x + x \cos x$ ; (b)  $e^x \sec^2 x + e^x \tan x$ ; (c)  $1 + \ln x$ ;  
 (d)  $\cos^2 t - \sin^2 t$ ; (e)  $12te^{3t} + (t^2 - 1)18e^{3t}$ ;  
 (f)  $1 - 3s + \log_e s - 6s \log_e s$ .
3. (a)  $\frac{1}{(1 - x)^2}$ ;  
 (b)  $\frac{3e^x \sin 2x - 6e^x \cos 2x}{9 \sin^2 x}$ ;  
 (c)  $\frac{4x(x^2 - 2)(3x + 4) - 6(x^2 - 2)^2}{(3x + 4)^3}$ ;  
 (d)  $\frac{-e^{2t}(2 \sin 2t + e^{2t} \cos 2t)}{e^{4t}}$ ;  
 (e)  $-\operatorname{cosec}^2 \theta$ .
4. -12.836.  
 5. -16.49.  
 6. 5.88.  
 7. -0.3254.  
 8. -54.

9. 12.96.
10. 90.16.
11. (a) 36 m/s; (b) 30 m/s<sup>2</sup>; (c)  $t = 1$  and 2 s.
12. (a) 27 rad/s; (b) 72 rad/s<sup>2</sup>; (c) at time  $t = 0$  and  $t = 1$  s.
13. Minimum at  $y = 1$ , maximum at  $y = 33$ .
14.  $x = 12.5992$  m.
15. Kinetic energy = 62.5 kJ.
16. (a)  $-\frac{4}{3}e^{-3x} - \frac{2}{3}\sin 3x + \frac{5}{2}\cos 2x + c$ ;  
 (b)  $-xe^{-x} - e^{-x} + c$ ;  
 (c)  $\frac{1}{17}\ln(5x + 1) - \frac{1}{17}\ln(-2x + 3) + c$ ;  
 (d)  $\frac{1}{2}[(\ln b)^2 - (\ln a)^2]$ ;  
 (e)  $-\ln(\cos x) + c$ ;  
 (f)  $\arcsin x + c$ .
17. (a) 6.298; (b)  $\frac{2}{3}$ ; (c) 147.07; (d) 1.15.
18. 0.423 to 3 d.p.
19.  $T = 2.399$  Nm.
20. Mean value = 1.074.
21.  $\bar{x} = 3.2$ ,  $\bar{y} = 16.38$ .
22.  $40\pi$  cm<sup>4</sup>.
23.  $y = \frac{M}{2EI}x^2 + Ax + B$ .
24. A is given, B = 0.
25. 4 Henry.

## Questions 6.5.1

(page 621)

1. (a)  $8 - 6i$ ; (b)  $36 + 26i$ ; (c)  $i$ ; (d)  $-\frac{1}{5} + \frac{2}{5}i$ ; (e)  $-\frac{27}{5} + \frac{11}{5}i$ .
2. (a)  $\sqrt{72} \angle -45^\circ$ ; (b)  $\sqrt{25} \angle 36.9^\circ$ ; (c)  $\sqrt{1681} \angle -12.7^\circ$ ;  
 (d)  $\sqrt{\frac{1}{25}} \angle -36.9^\circ$ ;
3. (a)  $2.74 + 4.74i$ ; (b)  $\frac{\sqrt{13}}{\sqrt{2}} + \frac{\sqrt{13}}{\sqrt{2}}i$ ; (c)  $\sqrt{3} + i$ .
4. (a) 651.92 (b)  $21.250 - 7250j$ ; (c)  $\frac{710}{37} - \frac{550}{37}i$ ;  
 (d)  $\frac{87}{185} - \frac{181}{185}i$ .
5.  $I = 0.431 - 0.172j$  amps.
6. (a)  $(2x - y)^2 = Ax^3$ ; (b)  $xy = x^2 + c$  and  $y = x - \frac{4}{x}$ .
7.  $y = \cos \omega t + ct + D$ .
8.  $y = x - 0.5$ .
9.  $y = -\frac{1}{4}e^{-4x}$  with respect to  $x$ .
10.  $y = \frac{I}{EI} \left[ \frac{WLx^2}{8} - \frac{Wx^3}{12} + \frac{27W}{12} - \frac{9WL}{8} \right]$ .



11.  $q = Ae^{4t} + Be^{-10t} + c.$
12.  $x = 3, y = 1, z = 2.$
13. 44.4 cm.
14. –
15. –
16. See model answer.
17. (a)  $G(t) = \frac{100}{193.75} e^{-2.5t} \sin \sqrt{193.75}t;$   
(b)  $G(t) = -10 + 4.44e^{-5t} + 5.55e^{4t}.$
18.  $G(t) = \frac{900}{\sqrt{4.5}} e^{-10.5t} \sin \sqrt{4.5}t.$

## Answers to TYK

### TYK 6.4.1

- (1) Business and administration = 29.23%, Humanities and social science = 42.28%, Physical and life sciences = 15.74%, Technology = 12.75%.
- (2)

|     |    |    |    |    |    |    |    |    |    |    |    |
|-----|----|----|----|----|----|----|----|----|----|----|----|
| $x$ | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 |
| $f$ | 1  | 2  | 4  | 5  | 7  | 5  | 4  | 7  | 2  | 2  | 1  |

- (3) Percentage height of column relates to average for class interval.

|                |      |       |    |       |       |      |
|----------------|------|-------|----|-------|-------|------|
| Class interval | 62   | 67    | 72 | 77    | 82    | 87   |
| Percentage     | 6.67 | 18.33 | 30 | 26.67 | 11.67 | 6.66 |

### TYK 6.4.2

- (1)  $\bar{x} = 127$
- (2) mean = 20, median = 8.5, mode = 9
- (3)  $x = 38.6$  cm, mean deviation = 1.44 cm
- (4)  $x = 169.075$  mm, mean deviation = 0.152 mm
- (5) median = 8.5,  $\sigma = 34.73$
- (6)  $\bar{x} = 3.42$ ,  $\sigma = 0.116$
- (7) (i)  $\bar{x} = 3.2$ ,  $\sigma = 0.19$ ; (ii)  $3.2 \pm 0.063$ .

### TYK 6.4.3

- (1)  $s = 0.56 \text{ m} + 0.76$
- (2)  $I = -0.48 + 2.27 V$
- (3) (i)  $Y = 0.256X + 2.226$ ; (ii) 0.079 low; (iii)  $Y = 0.619 X^2 + 3.22$ ; (iv) 0.53 more appropriate.

### TYK 6.4.4

- (1)  $91.4 \pm 0.61 \Omega$
- (2) 0.96
- (3) (i)  $Y = 1.12 + 2.27 X$ ; (ii) 26.54
- (4)  $t = 3.79$  out of control

- (5)  $\bar{x} = 1.5, c = 4.0$
- (6) Five missiles
- (7) (i) 0.0004; (ii) 0.1; (iii) 0.0396; (iv) 0.38
- (8) 0.54
- (9) (i) 0.842; (ii) 0.504
- (10) (i) 0.677; (ii) 0.0025.

This Page Intentionally Left Blank

# Index

- 2D drafting 83
- 3D drafting 83
- 3D wire-frame modelling 90
- AC coupled amplifier 258
- AC electrical principles 175
- AC generator 202
- AC in a capacitor 215
- AC in a resistor 214
- AC in an inductor 215
- ADC 254
- ALU 85
- AM signal 255
- ASCII code 251, 252
- Absorption costing 28
- Acceleration 140
- Accounting 4
- Accumulator 85
- Active component of current 238
- Active sensor 274
- Activity-based costing 29
- Aesthetics 68
- Allocation of work 4
- Alternating current 211
- Alternating voltage 211
- Amplifier 258, 259
- Amplifier bias 261
- Amplifier gain 260
- Amplitude 213
- Analogue signal 255
- Analogue-to-digital converter 254
- Angle of twist 135
- Angular equations of motion 146, 147
- Angular frequency 462
- Angular momentum 161
- Angular motion 146, 149
- Angular position sensor 276
- Angular velocity sensor 276
- Anti-derivative 477
- Anti-differentiation 492
- Apparent power 237, 238, 240
- Apportionment 28
- Arcing 281
- Area of revolution 502
- Argand diagram 223, 224
- Arithmetic logic unit 85
- Arithmetic mean 524
- Armature 205
- Assembly line production 24
- Astable multivibrator 270, 271
- Astable oscillator 268, 272, 273
- Attenuation 341
- Attenuation coefficient 340
- Attenuator 335
- Audio frequency amplifier 259
- Audio signals 249
- Automated drafting 83
- Automatic energy control system 277
- Average cost 34
- Average value 213
- Axial strain 363
- BS 60
- BS4811 81
- BS5750 13, 69
- BS5950 125, 129, 130
- BS7373 60
- Balance sheet 47
- Balanced network 331, 332
- Balancing 416
- Bandwidth 231, 232, 263, 264
- Bar chart 516
- Bayes theorem 541
- Beam 374, 378
- Beam selection 124
- Bel 260
- Belt drive 398, 401
- Belt tension 399
- Bending moment 105, 106, 108, 109
- Bias 261
- Bias point 262
- Bill of materials 23
- Bimodal data 527
- Binary 252
- Binomial distribution 544
- Binomial theorem 448, 451
- Bistable multivibrator 270
- Boundary value problem 606
- Bounded series 608
- Brainstorming 72
- Break-even analysis 31
- Break-even chart 31
- British Standards 60
- Budget centres 40
- Budgetary control 22, 37
- Budgetary planning 39
- Bulk modulus 367
- Buoyancy 509
- Burn-in 580
- Bus controller 85
- Business functions 30
- Business plan 38
- Buyer 20
- C-R circuits 191
- C-R ladder network oscillator 269
- CAD 83, 84, 88, 93
- CAD workstation 84
- CAD/CAM system 93, 94
- CAM 93
- CIMS 93
- CNC 83
- COSHH 68, 69
- CPM 49
- CPU 85
- CRT 86, 253
- Cache 86
- Calculus 477
- Capacitance 185
- Capacitive reactance 217
- Capacitor 187
- Capacitors in series and parallel 188
- Capacity planning 11
- Cartesian co-ordinates 460, 461
- Cash-flow budget 46
- Cashier 5
- Cathode ray tube 86
- Central processing unit 85
- Central tendency 522
- Centroid 118, 502, 503
- Characteristic impedance 331, 332
- Charge 175, 185
- Charts 516
- Chronological bar chart 518
- Class A 261
- Class AB 261
- Class B 262
- Class C 262
- Class midpoint 525

## 650 Index

- Class of operation 261
- Class width 521
- Class-C frequency multiplier 248
- Classification scheme 73
- Clock circuit 85
- Closed-loop gain 264
- Coefficient of coupling 317, 319
- Columns 113
- Combination 448, 542
- Combined errors 563
- Complementary function 596
- Complex admittance 226
- Complex impedance 223, 301
- Complex loading systems 363
- Complex notation 220, 222
- Complex numbers 585, 589, 610
- Complex waveform 211, 212, 243, 342
- Compound gear train 411
- Compound wound motor 205
- Computer aided design 83
- Computer graphics 88
- Computer hardware 84
- Computer modelling 90
- Computer-aided analysis 93
- Computer-aided planning 95
- Computer-integrated manufacturing system 93
- Concentrically loaded columns 127
- Conceptual design 71
- Conductance 226
- Cone clutch 406
- Confidence limit 569
- Conformance quality 12
- Conglomerate structure 17
- Conservation of energy 153, 158
- Constant current source 291
- Constant of integration 490
- Constant voltage source 287
- Consumer markets 6
- Contact angle 401
- Contact bounce 281
- Continuous distribution 547
- Continuous production 24
- Contract costing 24
- Contribution analysis 36
- Control 3
- Control character 251
- Controlled device 277
- Controller 275
- Controlling element 275
- Convergent series 608
- Copper loss 209
- Coprocessor 86
- Copyright Act 69
- Core losses 324
- Correlation 553
- Cost accounting 21
- Cost centre 40
- Cost control 21
- Cost object 29
- Costing sheets 79
- Costing systems 22
- Costing techniques 27
- Coulomb 175
- Coupled circuit 319
- Coupled magnetic circuit 317
- Coupled system 421
- Coupling 317, 319
- Crash time 52
- Credit controller 5
- Critical path 49, 50, 51
- Critical path method 49
- Cumulative frequency distribution 521
- Current 175, 176
- Current divider 181
- Current gain 259, 260
- Current law 177
- Current source 285
- Custom built 24
- Customer requirements 60
- Cut-off point 261, 262
- DAC 253
- DC coupled amplifier 258
- DC electrical principles 175
- DC generator 204
- DC motor 204
- DVST 86, 87
- Damping 172, 173, 174, 330
- Data 254
- De-layering 14
- Decibel 260, 341
- Decimal 251
- Deflection 374
- Deflection curve 105
- Deformation in bending 113
- Degree of correlation 553
- Degree of damping 173
- Delta-network 294
- Demodulation 249
- Density function 547, 548
- Derivative 486
- Design 30, 59
- Design parameters 60
- Design problem 71
- Design process 70, 81, 100
- Design quality 12
- Design report 59, 70, 81
- Design requirements 66
- Design solution 72, 74
- Design specification 59, 60, 71
- Development process 10
- Diac 283
- Dielectric 186
- Dielectric constant 185
- Differential calculus 477, 478
- Differential equation 586, 593, 616
- Differentiation rules 478
- Digital processing 253
- Digital signal 255
- Digital-to-analogue converter 253
- Digitiser 84
- Direct current 175
- Direct labour budget 43
- Direct view storage tube 86
- Direction 3
- Disc clutch 403
- Discounted cashflow 21
- Dispersion 522
- Distribution 6, 8
- Divisional structure 17
- Dot notation 321
- Double entry accounts 4, 5
- Dummy activity 51, 95
- Dynamic microphone 275
- Dynamic systems 140
- EMF 175
- EMF equation for a transformer 321
- Earliest expected time 53
- Eddy current 209, 324
- Effective length 129
- Effective noise voltage 256
- Effective value 213
- Efficiency 415
- Electrical power 183
- Electrical principles 175
- Electromechanical relay 281
- Electromotive force 175
- Electronic digitiser 84
- Encastre beam 105
- Energy 183
- Energy control system 275, 277, 278
- Energy storage 187, 198
- Energy transfer 153
- Engineering services 31
- Engineers' theory of bending 113
- Engineers' theory of torsion 134, 137
- Environment 66
- Epicyclic gear train 413
- Equations of motion 142
- Equilibrium 141
- Equivalent circuit of a transformer 323
- Ergonomics 66
- Error 558, 563
- Error analysis 560
- Error in sample mean 534
- Error signal 277
- Evaluating design solutions 74
- Evaluation matrix 75
- Expected time 53
- Exponential distribution 580
- Exponential equations 434

- Exponential functions 432
- Exponential series 454
- FEA 91, 93
- FEM 93
- FM signal 255
- FMS 93
- Facilities management 11
- Factor 432
- Factory schedules 19
- Failure rate 580
- Federal structure 16
- Feedback 264
- Field 205
- Finance 4
- Finance director 4
- Financial accountant 4
- Financial control 36, 37
- Financial planning 36, 37
- Finite element 93
- Finite element analysis 91, 93
- Finite element model 93
- First order differential equation 593
- First-order system 355
- Fixed cost 32, 44
- Flat structure 14
- Flexible manufacturing system 93
- Flow sensor 276
- Flux 195
- Flux density 196
- Flux waveform 248
- Flywheel 419
- Force 141
- Forced response 355
- Forcing frequency 174
- Forcing function 355
- Forecasting 2
- Formulae 432
- Fourier analysis 342
- Fourier coefficient 343, 347, 349
- Fourier series 343
- Frequency 213, 462
- Frequency distribution 519
- Frequency multiplier 248
- Frequency response 263
- Friction clutch 403, 404
- Full-load 208
- Full-wave control 283
- Function of a function rule 481
- Fundamental 243
- Fundamental component 242, 243
- Gain 259, 260
- Gantt chart 55, 95
- Gaussian elimination 615
- Gaussian law 565
- Gear train 409, 410
- General solution 596
- Generalised Gaussian law 565
- Generator principles 201
- Geometric modelling 83
- Goals 1
- Grouped data 525
- HASAWA 69
- Half-power frequency 231
- Half-wave rectifier 247
- Hard copy 19
- Harmonic 243, 245, 246, 247
- Harmonic analysis 608
- Harmonic component 243, 264, 347
- Harmonic motion 165
- Hedging of risk 4
- Hexadecimal 251
- Histogram 521, 527
- Hollow shaft 137
- Hoop stress 386
- Human factors 66
- Hysteresis loss 209, 324
- ISO9000 13, 69
- Ideal transformer 323
- Imaginary axis 224
- Imaginary part 222
- Impact 160
- Impedance 217, 218, 220
- Impedance triangle 217, 218, 219
- Impulse noise 255
- Indices 432
- Indirect cost 28
- Inductance 197
- Inductive reactance 216
- Inductor 195
- Industrial markets 8
- Inertia 141
- Inflection 488
- Information 18
- Information engineering 249
- Information system 249, 253, 254
- Information theory 250
- Initial value problem 606
- Input device 84, 253
- Input resistance 262, 263
- Input setting device 275
- Input transducer 275
- Instantaneous power 235
- Integral calculus 477, 490
- Integration by parts 491, 497
- Integration by substitution 491, 496
- Interaction 18
- Intermittent flow 11
- Internal resistance 285, 288, 292
- Intranet 19
- Inventory control 21
- Inventory levels 21
- Investment appraisal 20, 21
- Iron loss 209
- j-operator 222, 591
- JIT 21
- Job costing 22
- Joystick 84
- Just-in-time 21
- KE 157
- Keyboard 84
- Kinetic energy 157, 158
- Kirchhoff's current law 177
- Kirchhoff's voltage law 177
- L-R circuits 200
- LCD screen 254
- LDR 276
- LVDT 276
- Labour budget 43
- Ladder network 269
- Lagging 241
- Lap angle 401
- Laplace transform 356, 587, 601, 618
- Large-signal amplifier 259
- Latched state 282
- Lateral strain 363
- Latest allowable time 53
- Layout design 71
- Leadership 3
- Leading 241
- Leakage flux 323
- Least radius of gyration 128
- Legal implications 68
- Lenz's Law 199
- Light dependent resistor 276
- Light level sensor 276
- Light pen 84
- Limit 449
- Line flow 11
- Line of best fit 553
- Linear equations of motion 142, 147
- Linear factors 443
- Linear momentum 160
- Linear position sensor 276
- Linear variable differential transformer 276
- Link arm 413
- Liquid level sensor 276
- Loaded beam 373
- Loading 330
- Logarithmic functions 432
- Logarithms 432, 433
- Longitudinal stress 386, 394
- Loop gain 265
- Loudspeaker 275
- Low-noise amplifier 259
- Lower threshold voltage 274
- MOSFET 284
- MTBF 13

## 652 Index

- Macaulay's method 379, 382
- Magnetic field strength 195, 196
- Magnetic flux density 195
- Magnetising current 324
- Mail order 7
- Maintenance 67
- Make or buy 21
- Management accountant 5
- Management accounting 21
- Manufacture 67
- Manufacturing 31
- Manufacturing overhead budget 44
- Marginal cost 34
- Marginal costing 28
- Marketing 5
- Mass moment of inertia 120
- Master budget 47
- Matching 309, 310
- Materials budget 41
- Matrices 586, 598, 612
- Matrix structure 16, 17
- Maxima and minima 478
- Maximum point 488
- Maximum power transfer 210, 307
- Mean 523
- Mean deviation 527, 528
- Mean time between failure 13
- Mean value 502, 508, 523
- Measures and evaluation 31
- Median 526
- Menu tablet 84
- Mesh current analysis 312
- Microprocessor 85
- Mid-band 262
- Minimum point 488
- Modal class 527
- Modal value 527
- Mode 523, 526
- Modelling 83, 88, 90, 91
- Modulation 249
- Moment of inertia 149, 150, 502, 507
- Moments 103
- Momentum 141
- Monitor 86
- Monostable multivibrator 270
- Most likely time 53
- Motion 142, 146
- Motion of a simple pendulum 171
- Motion of a spring 169
- Mouse 84
- Multivibrator 268, 270
- Mutual inductance 199, 317
- NC part programming 94
- Natural frequency 173, 174
- Negative feedback 264, 265
- Neper 341
- Network 294, 300, 331
- Network diagram 50
- Network model 331
- Network theorem 304
- Network transformation 295
- Newton's laws of motion 141
- Node voltage analysis 314
- Noise 255, 256, 257, 258
- Non-linear laws 435
- Normal distribution 548
- Norton's theorem 292
- Numerical integration 491, 499
- Objectives 2
- Ogive 521
- Ohm's Law 176
- One-shot circuit 270
- Open-circuit voltage 288
- Optimistic time 53
- Optimum design solution 71
- Organisation chart 15
- Organisational structures 14
- Oscillating mechanical systems 165
- Oscillator 267, 268
- Oscillator types 268
- Oscillatory motion 173
- Output device 87, 253
- Output resistance 262, 263
- Output transducer 275
- Over-spend 3
- Overhead 44
- Overheads 28, 29
- PD6112 64
- PDS 64
- PE 154, 155
- PERT 49, 50, 53
- PF 443
- PLA 253
- PLC 275
- PWM signal 255
- Pads 309
- Parallel axes theorem 123
- Parallel capacitors 188
- Parallel circuit calculations 178
- Parallel resistors 180
- Parallel resonance 229
- Parent population 532
- Partial fractions 442, 443, 495
- Parts costing 24
- Passive sensor 274
- Peak value 213
- Peak-peak value 213
- Pendulum 171
- Per unit cost 24
- Per unit regulation 208
- Performance 66
- Periodic functions 462
- Periodic time 213
- Permeability 196
- Permeability of free space 197
- Permittivity 185
- Permittivity of free space 186
- Permutation 448, 542
- Pessimistic time 53
- Phase angle 217, 218, 219
- Phase change coefficient 340
- Phase shift 264
- Phasor 591
- Phasor diagram 214, 215
- Pi-network 294, 334
- Pi-network attenuator 338
- Pie chart 519
- Pillars 127
- Planet carrier 413
- Planet gear 413
- Planetary gear train 413
- Planning 1
- Plans 2, 3
- Plc 16
- Plotter 87
- Point load 104
- Point of inflection 488
- Poisson distribution 546, 566, 574, 581
- Poisson's ratio 363, 364, 366, 369
- Polar co-ordinates 460, 461
- Polar second moment of area 506
- Position vector 612
- Positive feedback 264, 267
- Posts 127
- Potential divider 181
- Potential energy 154, 155, 158
- Potentiometer 275
- Power 163, 183
- Power control 279
- Power controller 279
- Power factor 237, 238, 239, 352
- Power factor correction 240
- Power gain 259, 260
- Power in AC circuits 235
- Power in a pure reactance 236
- Power switching device 280
- Power transmitted by a shaft 138
- Power transmitted by a torque 163
- Precedence planning 95
- Pressure sensor 276
- Pressure vessel 385, 389, 392
- Price structure 7, 8
- Primary 206
- Prime function 492
- Principle of superposition 379
- Printable character 251
- Probability 515, 536, 538
- Probability distribution 536, 542
- Process costing 26
- Processing device 253
- Product design specification 60
- Product development 8

- Product rule 484
- Production process 11
- Profit 32
- Profit and loss statement 47
- Profit centre 40
- Profitability 36
- Program 275
- Programme evaluation and review technique 49
- Programme timing 8
- Programmed logic array 253
- Project cost 52
- Project costs 52
- Project manager 18
- Project network 95
- Project planning 48
- Project scheduling 48
- Project-based production 11
- Propagation coefficient 339
- Proportionate bar chart 517
- Proximity sensor 276
- Public limited company 16
- Pulse wave 212, 347
- Purchasing 13
- Pure tone 243
  
- Q-factor 231, 232, 327, 329, 330
- QC 69
- Quadratic factors 446
- Quality control 12, 69, 566, 569
- Quality factor 231
- Quality systems 13
- Quotient rule 484
  
- RMS value 213, 350, 502, 508
- Radian measure 459, 461
- Radio frequency amplifier 259
- Radius of gyration 127, 150
- Random error 560
- Random noise 255
- Random sample 532
- Raster refresh 87
- Rate of change 486
- Reactance 215, 216
- Reactive component of current 238
- Real axis 224
- Real part 222
- Reduction of non-linear laws 435
- Refresh system 86
- Register 85
- Regression 553
- Relative permeability 197, 198
- Relative permittivity 186
- Relay 281
- Reliability 13, 576
- Reliability theory 559
- Resistance 175, 176
- Resistors in parallel 180
- Resistors in series 180
  
- Resonance 173
- Resonant circuit 227, 232, 325
- Resonant frequency 227, 228
- Resource scheduling 95
- Responsibility accounting 29
- Retail outlets 6
- Right first time 69
- Rotating system 397
- Rotational KE 157
- Rotational kinetic energy 161
- Rules of differentiation 478
  
- s-operator 601
- SV 488
- Safety 69
- Sales 6
- Sales forecast 19
- Sample error 534
- Sampling 558, 566
- Schedules 19
- Script file 89
- Second harmonic 243
- Second moment of area 118, 120, 502, 505
- Second moment of mass 502
- Second order differential equation 593
- Second-order system 355
- Secondary 206
- Self-inductance 199
- Selling price 32
- Sensor 253, 274, 276, 277
- Sequence of numbers 449
- Series 448, 454
- Series AC circuit 225
- Series resistors 180
- Series capacitors 188
- Series circuit calculations 178
- Series coupling 320
- Series resonance 227
- Series wound motor 205
- Shear force 105, 106, 108, 109
- Shear modulus 369
- Shear strain 133
- Shear stress 105, 133, 136, 394
- Shunt wound motor 205
- Sigma notation 448, 449
- Signal 249, 254, 255
- Signal conditioning 253
- Signal-to-noise ratio 256, 257
- Silicon controlled rectifier 282
- Simple harmonic motion 165
- Simply supported beam 102
- Sine wave 212
- Single use plans 2
- Sinusoidal oscillator 268
- Sinusoidal waveform 242
- Slack time 51, 53
- Slenderness ratio 127
  
- Slope 374
- Slope and deflection 374, 378
- Small-signal amplifier 259
- Solid modelling 91
- Specification 60
- Speed 140
- Spring 169
- Spur gear 409
- Square wave 212, 245
- Standard derivatives 478, 479
- Standard deviation 523, 529, 530
- Standard integral 490, 492
- Standard supply multiple 24
- Standard time 43
- Standing plans 2
- Star-network 294
- Static systems 102
- Stationary value 488
- Statistical probability 536
- Statistical process control 21
- Statistics 515
- Step-down transformer 208
- Step-up transformer 208
- Stepper motor 275
- Stock control 20
- Storage device 88
- Strain energy 156
- Strain sensor 276
- Strategic plans 3
- Stress due to bending 113
- Struts 113
- Sun gear 413
- Superposition 460
- Superposition theorem 293
- Supervision of work 4
- Supply 13
- Surface modelling 90
- Susceptance 226
- Systematic error 560
- Systematic search method 73
  
- T-network 294, 333
- T-network attenuator 336
- TP 487, 488
- TPV 206
- TQ process 69
- TQM 69
- Tablet 84
- Tactical plans 3
- Tall structure 15
- Targets 18
- Teething troubles 580
- Temperature sensor 277
- Theorem of Pappus 502
- Thermocouple 275
- Thevenin's theorem 288
- Third order differential equation 593
- Three-dimensional loading 366
- Threshold voltage 273, 274



## 654 Index

- Thyristor 282  
Time constant 191  
Time planning 95  
Torque 134, 149  
Torque in gear trains 414  
Torsion 133  
Total Quality Management 69  
Total cost 32  
Total flux 195  
Transducer 253, 274  
Transformation 295, 297  
Transformation equations 146  
Transformer 206, 321, 323  
Transformer applications 208  
Transformer efficiency 208, 209  
Transformer matching 310  
Transformer principles 205  
Transformer regulation 208  
Transient 355  
Transistor 284  
Translation KE 157  
Transportation 67  
Trapezoidal rule 500  
Triac 283  
Triangle formulae 460  
Triangle wave 212  
Triangles 463  
Triangular wave 245  
Trigonometric identities 469  
Trigonometric ratios 460  
Trigonometry 459  
True cost 28  
True power 237, 238, 240  
Turning point 487  
Turns ratio 206  
Turns-per-volt 206  
Twisting 133  
Two-dimensional loading 364  
Two-dimensional stress system 364  
UDL 104, 106, 109, 382  
Unbalanced network 331  
Uniformly accelerated motion 140  
Uniformly distributed load 104  
Unimodal data 527  
Universal resonance curve 326, 327  
Upper threshold voltage 273  
Useful life 580  
V-belt 398, 400  
VA rating of a transformer 208  
VDU 84, 86, 251  
Value-added 21  
Variable cost 32, 44  
Variance 523  
Vector 598  
Vector refresh 87  
Vectors 586, 612  
Velocity 140, 143  
Vibration sensor 277  
Video signals 249  
Voltage 175, 176, 185  
Voltage gain 259, 260  
Voltage law 177  
Voltage regulation 208  
Voltage source 285  
Volume of revolution 502  
Volumetric strain 367  
Waveform 211, 212, 242  
Wear-out 580  
Weibull distribution 582  
Weight sensor 277  
What-if analysis 31  
Wholesale outlets 7  
Wideband amplifier 259  
Wien bridge network 269  
Wien bridge oscillator 269  
Winding resistance 323  
Wire-frame modelling 90  
Work 183  
Work done 153, 154  
Work-in-progress 20  
Work-oriented culture 4  
Workforce management 12  
Workstation 84