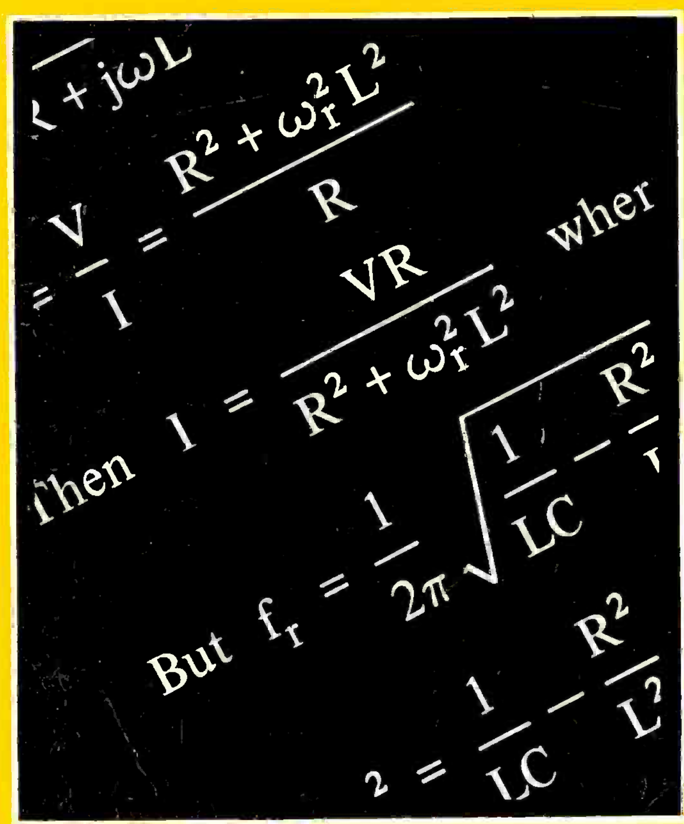
Further Practical Electronics Calculations and Formulae

F.A. WILSON



FURTHER PRACTICAL ELECTRONICS CALCULATIONS AND FORMULAE

OTHER BOOKS OF INTEREST

160	Coil Design and Construction Manual
228	Essential Theory for the Electronics Hobbyist
BP53	Practical Electronics Calculations and Formulae
BP62	The Simple Electronic Circuit and Components (Elements of Electronics – Book 1)
BP63	Alternating Current Theory (Elements of Electronics – Book 2)
BP64	Semiconductor Technology (Elements of Electronics – Book 3)
BP77	Microprocessing Systems and Circuits (Elements of Electronics – Book 4)
BP89	Communication (Elements of Electronics – Book 5)
BP111	Audio (Elements of Electronics – Book 6)

FURTHER PRACTICAL ELECTRONICS CALCULATIONS AND FORMULAE

by
F. A. WILSON
CGIA, C.Eng, FIEE, FIERE, FBIM

BERNARD BABANI (publishing) LTD
THE GRAMPIANS
SHEPHERDS BUSH ROAD
LONDON W6 7NF
ENGLAND

PLEASE NOTE

First Published – December 1986

British Library Cataloguing in Publication Data
Wilson, F.A.
Further practical electronics calculations and formulae (BP144)
1. Electronics – Mathematics
I. Title
621.381'01'51 TK7835

ISBN 0 85934 119 4

Printed and Bound in Great Britain by Cox & Wyman Ltd, Reading

PREFACE

Throughout the ages men and women of science have managed to describe so much through mathematical formulae but for us the inheritors, faced also with a technology which is growing apace, is the thorny problem of how to keep in touch with so much. We cannot remember these formulae one and all, that is too much to ask so we must fall back on such a book as this which has formulae galore at the turn of a page. But there is even more, most of the formulae are developed from first principles with the whys and wherefores clarified and because it is not until we put figures into formulae that the true meaning emerges, many examples are also included.

This is the companion book to BP53 under the same title, carrying on where the first leaves off. Together the two books constitute a complete *ABC of Electronics and Communication*, a work not only for reference but also of learning. Each of the two books stands on its own however, the first embracing components, elementary circuit analysis, networks and measurements, this one continuing in those branches of electronics where formulae abound.

There are some who actually enjoy plodding through formulae to unveil their secrets. But there are many more who because of their calling or because they are students, have this thrust upon them. Irrespective of the category into which each of us falls (and it could be both), this book is written to provide help. Electronics does not stay still, it is continually producing new ideas, systems and other wonders. To keep up we must retain an intimate knowledge of the basics for once these are at our fingertips the whole world of electronics is within our grasp.

The book is therefore of interest to all with an affection for electronics. For the *student*, working through examples with the author brings a greater intimacy with the subject. For the *practising engineer* there is always the need to recall formulae and how they arise. Even the *expert* cannot be indifferent to what goes on in other spheres of activity.

The mathematics employed are deliberately kept at low key and should not frighten most readers. Seldom do we go beyond the inevitable trigonometry, moreover for the many of us with short memories all the algebraic and trigonometrical formulae used reappear in an Appendix. The early chapters are perhaps more theoretical than practical but they are included because we may at times need to refresh our memories on the underlying fundamental principles. Introductions to the techniques of Statistics and Reliability are also here for without these no electronics engineer is truly complete. For easy reference all key formulae developed or quoted in each chapter are repeated in a table at the end.

A note about the tools for the job is in order. The aim is to study or remind ourselves of electronics formulae, not to practise arithmetic so we get by with any device we can, logarithm tables, slide-rule, ordinary home calculator, computer or best of all a scientific calculator. None of these is essential however because all examples are worked out.

A bracketed, raised reference [A3] refers to Appendix 3 (A3.1) and equations with numbers commencing A2 are to be found in Appendix 2.

F. A. Wilson

CONTENTS

Page

Page

Page

Page

Page

1. ELECTRICITY

Indubitably the whole of this book is based on the escapades of that tireless little worker, the electron. Tireless because it never ever stops moving and it is entirely through its motion that it achieves so much. It is fitting therefore that for a better appreciation of the electron and its habits, the book should open with an examination of things pertaining to motion itself. The laws of motion were developed long before the electron was discovered and these same simple laws explain the action in everything that moves. Without motion the electrons and even we ourselves, achieve nothing.

1.1 DYNAMICS

Electric current flow is due to particles in motion under the influence of forces. *Dynamics* is the subject title embracing both motion and force and the appropriate laws follow.

1.1.1 Newton's Laws

In the late 1600's Newton[(A3)] published his "Laws of Motion and Gravitation". The *First Law* states that matter will continue in its state of rest or of uniform motion in a straight line unless acted upon by some external force. So arises the definition of *force* as that which changes the state of rest or motion of a body. If speed changes then acceleration is involved and from Newton's *Second Law* which states that the rate of change of momentum (mass x velocity) is proportional to the applied force:

$$F = m.a \text{ newtons (N)} \tag{1.1}$$

where m is the mass of the body in kilogrammes (kg) and a is the acceleration in metres per second per second (m/s^2), i.e. a *newton* is defined as that force which, applied to a mass of 1 kg, impresses upon it an acceleration of 1 m/s^2.

Gravity is the force which any object exerts on any other. It is directly proportional to the masses of the objects and inversely proportional to the square of the distance between them, expressed through Newton's *Law of Gravitation* as:

$$F = \frac{Gm_1 m_2}{d^2} \text{ newtons} \qquad (1.2)$$

where m_1 and m_2 (kg) are the two masses and d the distance between their centres in metres.

F is the mutual force of attraction between the masses and G is the *gravitational constant* for all masses. It has a value which changes minutely as further experimentation produces more precise answers but a round figure of 6.67×10^{-11} $N.m^2/kg^2$ is sufficiently accurate for most calculations. It is the gravitational pull of the Earth because of its relatively enormous size which is noticed in everyday life rather than the minute attraction between objects on the Earth.

Measurements have shown that gravity due to the Earth causes a body to fall freely with an acceleration of 9.81 m/s^2 and from $F = m.a$, for a mass of 1 kg, $F = 9.81$ N so it can be said that the weight of a mass of 1 kg is equal to a force of 9.81 newtons (hence roughly, a 100 g or 3½ oz weight exerts a force of one newton).

EXAMPLE 1.1:
Using Newton's Law of Gravitation, what is the mass ("weight") of the Earth?

Let m_1 be a mass of 1 kg, m_2 the mass of the Earth
" r be the radius of the Earth (approx. 6370 km)
" F be the force exerted by the Earth on a mass of 1 kg (9.81 N)

Then, from Equation (1.2):

$$F = \frac{Gm_1 m_2}{r^2}$$

i.e.

$$m_2 = \frac{Fr^2}{Gm_1} = \frac{9.81 \times (6370 \times 10^3)^2}{6.67 \times 10^{-11} \times 1}$$

i.e.

$$m_2 = 5.968 \times 10^{24} \text{ kg } (\approx 6 \times 10^{21} \text{ tonnes}).$$

Note that from Einstein's work there is a modification to the law, but so small that for everyday calculations it can be ignored.

1.2 ENERGY

Energy is an intangible something capable of generating action and it is present in all electronic activity. It is difficult to define precisely but usually is said to be the "capacity for doing work" and accordingly is measured in the same units.

1.2.1 Work

When a force moves a body, work is done and the measure is the product of the force (F) and the distance (s) its point of application moves. The unit is the work done by a force of one newton acting through a distance of one metre, giving the metre-newton or *joule* (J).

$$\text{Work done } (W) = F \times s. \qquad (1.3)$$

EXAMPLE 1.2:
How much work is done when a load of 10 kg is lifted 2 metres?

The load is raised against the Earth's gravity which exerts a pull downwards of 9.81 newtons per kilogramme. The total force F is therefore $10 \times 9.81 = 98.1$ N. Hence from Equation (1.3):

$$W = 98.1 \times 2 = 196.2 \text{ joules}$$

1.2.2 Conservation of Energy

According to the energy conservation principle, the total amount of energy in the universe is constant. A given quantity of energy in one form can only be changed into an equivalent quantity in one or more other forms.

1.2.3 Mechanical Energy

Two forms exist which are interchangeable. They are the energy of motion, known as *kinetic* energy (from Greek, to move), the other is not due to motion but is capable of creating it and is called *potential* energy (from Latin, to be able).

1.2.3.1 The Equations of Motion

Consider a moving body having: an initial velocity u (m/s), a final velocity v (m/s) and an acceleration a (m/s^2). If moving with uniform acceleration for a time t secs, then the final velocity is equal to the initial velocity plus the increase in velocity ($a \times t$), i.e.:

$$v = u + at \tag{1.4}$$

and the mean velocity $= \dfrac{u + v}{2}$ (1.5)

Let the distance covered be s, then

$$s = \left(\frac{u + v}{2}\right) \times t \tag{1.6}$$

$$\therefore \quad s = \left[\frac{u + (u + at)}{2}\right] t$$

$$\therefore \quad s = ut + \tfrac{1}{2}at^2 \tag{1.7}$$

Squaring Equation (1.4):

$$v^2 = (u + at)^2 = u^2 + 2uat + a^2t^2 \text{ [from Eqn.(A2.4)]}$$

$$= u^2 + 2a(ut + \tfrac{1}{2}at^2)$$

$$\therefore \quad v^2 = u^2 + 2as \tag{1.8}$$

1.2.3.2 Kinetic Energy

If a body of mass m is acted upon by a force F, then from Section 1.1.1 [Eqn.(1.1)] :

$$F = m.a$$

Let the body start from rest ($u = 0$) and reach a velocity v over a distance s. Multiplying the above by s gives:

$$F.s = m.a.s$$

From Equation (1.8):

$$\text{since } u = 0,\ v^2 = 2as \quad \text{i.e. } as = \frac{v^2}{2}$$

$$\therefore \quad F.s = \frac{mv^2}{2}$$

Fs is the work done and is a measure of the *kinetic energy* (E_k) of the body, that is,

$$E_k = \tfrac{1}{2}m.v^2 \text{ joules} \tag{1.9}$$

(m in kg, v in m/s). Note especially that the k.e. is proportional to the velocity *squared.*

EXAMPLE 1.3:

A body of mass 20 kg is lifted to a height of 10 m and allowed to fall. What is its k.e. immediately before hitting the ground?

$$m = 20 \text{ kg}; \quad s = 10 \text{ m}; \quad a = 9.81 \text{ m/s}^2$$

Let final velocity (before hitting ground) = v, initial velocity $u = 0$. From Equation (1.8):

$$v^2 = u^2 + 2as = 2as = 2 \times 9.81 \times 10$$

and from Equation (1.9):

$$\text{final k.e.} = \tfrac{1}{2}m.v^2 = \tfrac{1}{2} \times 20 \times 20 \times 9.81$$

$$= 1962 \text{ joules}$$

When the body in the above example comes to rest on the ground its k.e. must be zero because it no longer has motion. The 1962 joules of energy are dissipated in the work done in deforming the ground and the body itself.

1.2.3.3 Potential Energy

When a body can do work, not because of its motion but by virtue of its position, for example a body raised to a height above ground, it is said to possess *potential energy* (E_p). Simply put, it has the potential to fall to a lower position and thereby do work, the same amount as was necessary to raise it in the first place. A wound or compressed spring is another example.

EXAMPLE 1.4:
A clock is driven by a weight of 10 kg. The total travel of the weight is 5 m. What is the p.e. of the weight when the clock is fully wound?

Force due to gravity acting on 1 kg = 9.81 N

∴ total force on clock weight = 98.1 N

$$\therefore E_p = \text{work possible} = F \times s = 98.1 \times 5 = 490.5 \text{ J}$$

1.2.3.4 Conversion between k.e. *and* p.e.

When a body possesses energy, whether kinetic or potential, it has the capability of changing its energy over from one type to the other and assuming no losses at the change, the total energy (k.e. + p.e.) is constant.

EXAMPLE 1.5:
In Example 1.3 what is the p.e. of the body (i) before it is allowed to fall, (ii) at 4 m above ground?

(i) Force due to gravity on 20 kg $= 9.81 \times 20 = 196.2$ N

$\therefore$ Work possible (p.e.) $= 196.2 \times 10 = 1962$ J

(ii) At 4 m above ground, from Equation (1.9):

$$v^2 = u^2 + 2as = 2 \times 9.81 \times 6 = 117.72$$

and from Equation (1.9):

$$\text{k.e.} = \tfrac{1}{2}mv^2 = \tfrac{1}{2} \times 20 \times 117.72 = 1177.2 \text{ J}$$

Because at 4 m above ground 1177.2 J of the original p.e. of 1962 J have been changed to k.e., the remaining p.e. $=$ $1962 - 1177.2 = 784.8$ J. Check:–

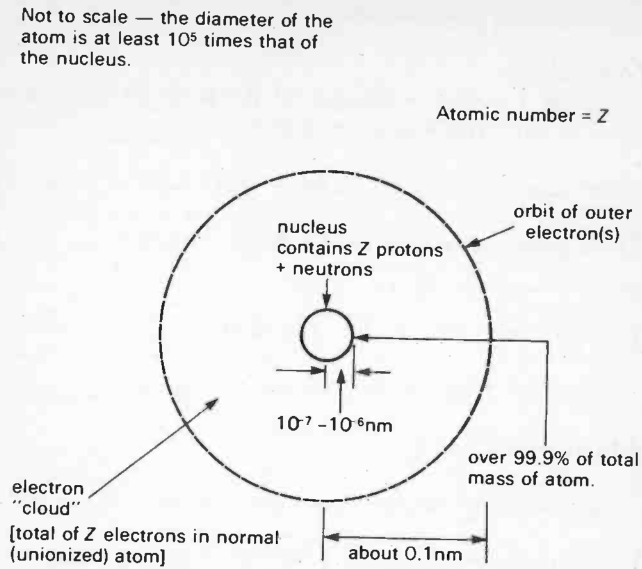
Force due to gravity on body $= 196.2$ N

$\therefore$ remaining work possible $= 196.2 \times 4 = 784.8$ J

As the body falls therefore, its store of p.e. is being continually converted into k.e. such that (k.e. + p.e.) is constant.

1.3 ATOMIC STRUCTURE

An atom is the smallest particle of an element which can take part in a chemical reaction. Within it are elementary particles carrying *charges* which are of major interest in electronics. An atom consists of a central dense *nucleus* around which its electrons travel in their individual orbits,



electron charge = -1.6×10^{-19}C, mass = 9.1×10^{-31}kg.

proton charge = $+1.6 \times 10^{-19}$C, mass = 1.67×10^{-27}kg.

Fig. 1.1 Some general characteristics of the atom

some general details are shown in Figure 1.1. Any substance consisting entirely of atoms of the same kind is known as an *element.* Different atoms combining form *molecules* which are the basis of *compounds.* As an example, water is a compound in which atoms of the elements hydrogen and oxygen combine to form the water molecule (H_2O). There are just over 100 elements each differing in the number of protons in the nucleus. This is known as the atomic number (Z) and Table 1.1 gives the values for elements associated with electronics. We label the proton charge positive and that of the electron negative and being of exactly the same magnitude

Table 1.1: ATOMIC NUMBERS AND SHELLS

Element	Symbol	Atomic Number, Z	Shell Occupancy (maxima shown in brackets)						
			K-1 (2)	L-2 (8)	M-3 (18)	N-4 (32)	O-5 (50)	P-6 (72)	Q-7 (98)
Aluminium	Al	13	2	8	3				
Antimony	Sb	51	2	8	18	18	5		
Arsenic	As	33	2	8	18	5			
Boron	B	5	2	3					
Carbon	C	6	2	4					
Copper	Cu	29	2	8	18	1			
Gallium	Ga	31	2	8	18	3			
Germanium	Ge	32	2	8	18	4			
Gold	Au	79	2	8	18	32	18	1	
Hydrogen	H	1	1						
Indium	In	49	2	8	18	18	3		
Iron	Fe	26	2	8	14	2			
Manganese	Mn	25	2	8	13	2			
Mercury	Hg	80	2	8	18	32	18	2	
Neon	Ne	10	2	8					
Nickel	Ni	28	2	8	16	2			
Oxygen	O	8	2	6					
Palladium	Pd	46	2	8	18	18			
Phosphorus	P	15	2	8	5				
Platinum	Pt	78	2	8	18	32	17	1	
Selenium	Se	34	2	8	18	6			
Silicon	Si	14	2	8	4				
Silver	Ag	47	2	8	18	18	1		
Tantalum	Ta	73	2	8	18	32	11	2	
Thorium	Th	90	2	8	18	32	18	10	2
Tungsten	W	74	2	8	18	32	12	2	
Zinc	Zn	30	2	8	18	2			

(denoted by *e*), one proton and one electron together are electrically neutral. An atom with its normal complement of electrons contains a nucleus carrying a +ve charge equal to *Z.e* surrounded by *Z* electrons each carrying a charge $-e$. Losing one or more electrons makes an atom a positive *ion* (from Greek, wanderer), equally gaining an electron creates a negative ion.

1.3.1 The Electron

Although usually pictured as a tiny ball, the electron has such an extremely small mass and indeterminate size that it is better considered simply as a minute electrical charge of insignificant mass. Certainly in electronics it is the effect of the charge which predominates. From the figures given in Figure 1.1 the proton has a mass $(1.67 \times 10^{-27})/(9.1 \times 10^{-31}) = 1835$ times that of the electron. Also because the electron charge *e* is (–) 1.6×10^{-19} coulombs, there are $1/(1.6 \times 10^{-19}) = 6.25 \times 10^{18}$ electron charges per coulomb, hence for one ampere of current 6.25×10^{18} electrons flow per second.

1.3.1.1 Electron Shells

A look into the electron "cloud" of Figure 1.1 shows that the electrons are grouped into concentric *shells.* Up to 7 shells occur, labelled from inner to outer KLMNOPQ although most atoms have less. Each shell can contain a certain number of electrons and with the shells numbered n = 1 to 7, the maximum number of electrons possible for each is $2n^2$, e.g. for the K shell n = 1 and the maximum number of electrons that any atom can have in orbit is 2, for the N shell, n = 4 and the maximum number is $2 \times 4^2 = 32$, as shown in Table 1.1. Shells are not necessarily completely filled. Figure 1.2 shows examples of these facts pictorially, at (i) is the simplest atom, that of hydrogen with its single electron in the only (K) shell which is therefore *incomplete* and at (ii) that of the copper atom with shells K, L, M and N, the first three full but the last with one electron only whereas it has room for 32. The distribution of electrons within the shells for most elements of interest in electronics is also given in Table 1.1. An atom

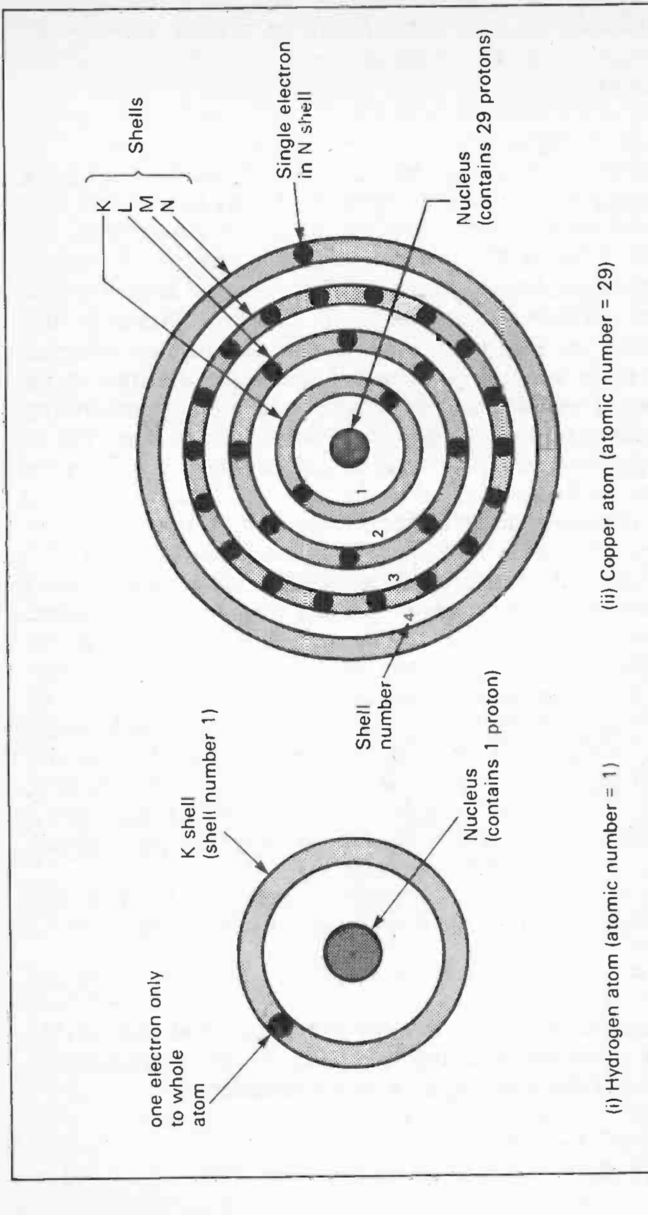


Fig. 1.2 Arrangements of electrons in shells

can therefore be described uniquely by its shell contents (e.g. for copper 2, 8, 18, 1; for silicon 2, 8, 4) for no two atoms are the same.

1.3.1.2 Energy Levels

Each electron is maintained in its own particular orbit (in the same way that a satellite orbits the Earth) by the fact that the centrifugal force due to its motion is balanced by the inward force of the equal and opposite charges of electron and nucleus. Each shell has a different *energy level* which in general increases with the shell number, ultimately the electrons in the outer shell having the highest energies. Because in addition the force binding the electron to its nucleus is weakest for the outermost electrons, these are the ones most easily dislodged from the atom and thus free to take part in *conduction.* The outermost shell is known as the *valency* shell.

In dealing with matters of atomic size, quantity of energy is expressed in *electron-volts* (eV), one eV = 1.6×10^{-19} J and this is the kinetic energy acquired by a free electron when it is accelerated by moving through a p.d. of one volt. Within certain constraints the energy levels of electrons can be changed but only to known fixed levels. An electron may jump from one particular energy level to the next but not to a level in between for it cannot orbit in the gaps between shells, hence the term *forbidden energy gap.* Heat is the general source of energy for electrons but another is from light which nowadays is considered as a stream of elementary particles of energy called *photons*, each containing a certain amount of energy known as a *quantum.*

The energy (E) of a single photon is related to the frequency (ν) of the radiation by the simple formula:

$$E = h\nu \tag{1.10}$$

This was developed by Planck[(A3)] in 1900 and h is *Planck's Constant* with a value of 6.626×10^{-34} joule-seconds. In terms of the wavelength, λ of the radiation:

$$E = \frac{hc}{\lambda} \tag{1.11}$$

where c = velocity of light (3×10^8 m/s).

Energy is only transferred from a photon to an electron if it is of the right value to raise the electron energy from one permissible level to another. When photons are encountered with insufficient energy for this to happen, no radiation is absorbed. The additional energy required to free electrons from their parent atoms so that they can be used as charge carriers is known as the *band-gap energy*. One practical outcome of this process arises when light falls on a semiconductor material and thereby reduces its resistance by freeing electrons, it is known as *photoconduction*.

The energy transfer process is reversible in that by forcing electrons to drop from one permissible energy level to another, they emit photons, this is a basic principle of the *laser*.

To put figures to the process it is better to consider the simplest atom, hydrogen. Its electron is normally at level 1. If 10.2 eV of energy is supplied, the electron energy is raised to level 2. 10.2 V is the *first excitation potential* for hydrogen, it is the p.d. required for energy to be absorbed. Anything lower is not accepted because it would bring the electron into the forbidden energy gap. Adding still more energy will not necessarily raise the electron to an even higher level because the point is reached where the k.e. and hence centrifugal force exceeds the attraction of the electron to the nucleus and the electron is freed. The atom then becomes a positive ion and the voltage required is known as the *first ionization potential*. For hydrogen this voltage is 13.6 and the energy supplied to the electron is 13.6 eV, i.e. $13.6 \times 1.6 \times 10^{-19}$ J $= 2.18 \times 10^{-18}$ J. Naturally other atoms which can be so treated have other potentials because of their differing electron arrangements.

An interesting even though somewhat approximate calculation is that of electron velocity within the atom.

EXAMPLE 1.6:

Given that the separation, d between electron and proton of a hydrogen atom is 0.53×10^{-10} m, calculate the electron velocity.

(We bring forward a well-known formula from Section 2.4

to do this. If not fully understood, it can be accepted here with clarification to follow later.)

The force of attraction varies as

$$\frac{Q_1 Q_2}{4\pi\xi d^2}$$

where Q_1 and Q_2 are two charges. Because they are separated by space:

$$\text{force} = \frac{-e^2}{4\pi\xi_0 d^2}$$

where ξ_0 is the electric constant of free space,

$$= 8.854 \times 10^{-12} \text{ F/m.}$$

Now when a body is held in circular motion, a *centripetal* force towards the centre of the plane of motion acts to prevent the body continuing in a straight line (Newton's First Law – Sect. 1.1.1), i.e. prevents it from flying off at a tangent. By virtue of its energy the body exerts an equal and opposite *centrifugal* force outwards otherwise it would fall into the centre. This can be shown to be equal to mv^2/d where m is the body mass and v its velocity. Hence

$$\frac{e^2}{4\pi\xi_0 d^2} = \frac{mv^2}{d}$$

$$\therefore \quad v = \sqrt{\frac{e^2}{4\pi\xi_0 dm}}$$

$$\text{i.e.} \quad v = \sqrt{\frac{(1.6 \times 10^{-19})^2}{4\pi \times 8.854 \times 10^{-12} \times 0.53 \times 10^{-10} \times 9.1 \times 10^{-31}}}$$

$= 2.18 \times 10^6$ m/s (the mass m is quoted in Fig.1.1), thus resulting in over six thousand million orbits in one millionth of a second!

EXAMPLE 1.7:
Calculate the first ionization potential for hydrogen.

The method is to find when the kinetic energy (E_k) which has the capability of pulling the electron out of orbit is equal to or exceeds the potential energy (E_p) through which it is kept in orbit.

From Equation (1.9) and since from the above example, $v = 2.18 \times 10^6$ m/s

$$E_k = \tfrac{1}{2}mv^2 = \tfrac{1}{2} \times 9.1 \times 10^{-31} \times (2.18 \times 10^6)^2$$

$$= 2.16 \times 10^{-18} \text{ J}$$

$$E_p = F \times s = \frac{e^2}{4\pi\xi_0 d^2} \times r$$

and because the distance d is equal to r

$$\therefore \quad E_p = \frac{e^2}{4\pi\xi_0 d} = \frac{(1.6 \times 10^{-19})^2}{4\pi \times 8.854 \times 10^{-12} \times 0.53 \times 10^{-10}}$$

$$= 4.34 \times 10^{-18} \text{ J}$$

For the electron to be freed therefore the existing k.e. must be increased by

$$(4.34 - 2.16) \times 10^{-18} = 2.18 \times 10^{-18} \text{ J}.$$

$$\therefore \text{ Energy required} = \frac{2.18 \times 10^{-18}}{1.6 \times 10^{-19}} = 13.625 \text{ eV}$$

corresponding to a potential of 13.625 V, which agrees well with the standard figure for hydrogen quoted earlier.

1.4 CURRENT FLOW

From the preceding Section the mechanism of current flow is now apparent. When a p.d. exists due for example to the chemical action of a battery or the mechanical work of a generator, then electrons are released from the valency shells of the atoms of certain materials known as conductors. These free electrons drift under the influence of the p.d. away from its negative and towards the positive pole and the total quantity moving per unit of time is:

$$Q = I.t \text{ coulombs (C)} \qquad (1.12)$$

where I = current in amperes and t = time in seconds.

EXAMPLE 1.8:
A current of 80 mA flows in a circuit for 400 ms. What quantity of electricity and how many electrons are involved?

From Equation (1.12):

$$Q = 0.08 \times 0.4 = 0.032 \text{ C}$$

and from Section 1.3.1:

$$\text{Number of electrons} = \frac{0.032}{1.6 \times 10^{-19}} = 2 \times 10^{17}$$

1.5 SUMMARY OF KEY FORMULAE

a = acceleration (m/s^2)
d, s = distance (m)
G = gravitational constant
I = current (A)
m = mass (kg)
t = time (s)
u = initial velocity (m/s)
v = final velocity (m/s)

QUANTITY	FORMULA	UNIT	SECTION
Acceleration due to gravity	$g_n = 9.80665$	m/s^2	1.1.1
Gravitational constant	$G = 6.670 \times 10^{-11}$	$N.m^2/kg^2$	1.1.1
Electron mass	$m_e = 9.1091 \times 10^{-28}$	g	1.3.1
Electron charge	$e = (-)1.6021 \times 10^{-19}$	C	1.3.1
Electron-volt	$eV = 1.6021 \times 10^{-19}$	J	1.3.1.2
Proton mass	$m_p = 1.6725 \times 10^{-24}$	g	1.3.1
Planck's Constant	$h = 6.626 \times 10^{-34}$	J.s	1.3.1.2
Force	$F = m.a$	N	1.1.1
Force due to gravitation	$F = \frac{G.m_1 m_2}{d^2}$	N	1.1.1
Work	$W = F.s$	J	1.2.1
Distance travelled	$s = ut + \frac{1}{2}at^2$	m	1.2.3.1
Velocity	$v = u + at = \sqrt{u^2 + 2as}$	m/s	1.2.3.1
Kinetic energy	$E_k = \frac{1}{2}mv^2$	J	1.2.3.2
Potential energy	$E_p = F \times s$	J	1.2.3.3
Quantity of charge	$Q = I.t$	C	1.4

2. ELECTROSTATICS

When a voltage is connected across two adjacent conductive terminals or plates an *electrostatic field* is said to exist between them. Diagrammatically this field is shown as lines of *electric flux* each with an arrow pointing in the direction in which an imaginary free *positive* electric charge would move. Because like charges repel and unlike attract such a free charge would move towards the –ve potential and away from the +ve. This conventional way of picturing an electrostatic field is shown in Figure 2.1.

2.1 ELECTRIC FIELD STRENGTH

The *electric field strength* (or *intensity*), E at any point is a measure of the force, F exerted per unit of charge on a charged particle at that point, that is:

$$E = \frac{F}{Q} \text{ newtons per coulomb} \qquad (2.1)$$

where Q is the charge on the particle.

Also, considering two parallel plates d metres apart with V volts impressed across them (for example, as in Fig.2.1) then the electrical field strength can also be expressed by the potential gradient, that is,

$$E = \frac{V}{d} \text{ volts per metre} \qquad (2.2)$$

2.1.1 Potential Difference

If a positive charge is moved in an electric field against the direction of the field (i.e. towards the +ve plate in Fig.2.1) then work is done and the potential energy of the charge

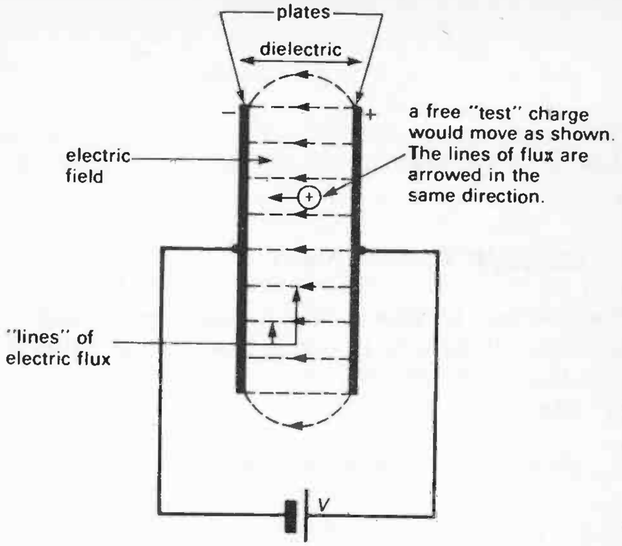


Fig. 2.1 Indicating the direction of an electric field

increases. Alternatively should the charge be moved in the direction of the field then its potential energy is reduced and it appears as an increase in the kinetic energy (Sec.1.2.3.4).

A charge of Q coulombs experiences a force of QE newtons in a field of strength E [Eqn.(2.1)] and if the charge moves a distance d metres the work done is QEd newton-metres.

It is not possible to specify absolutely the potential at a point, only its value relative to some other point, hence we work in units of *potential difference* (p.d.). The p.d. between two points separated by a distance d metres is the work done or change in energy *per unit charge* in moving a charge within an electric field between those points hence:

$$V = \frac{QEd}{Q} \quad \text{or} \quad \frac{Fd}{Q} \text{ newton-metres per coulomb}$$

but 1 newton-metre = 1 joule (Sect.1.2.1)

$$\text{p.d.} = \frac{Fd}{Q} = \frac{W}{Q} \text{ joules per coulomb, i.e. } volts \qquad (2.3)$$

where W is the energy or work involved.

2.2 ELECTRIC FLUX DENSITY

By definition, Q lines of flux radiate from a charge of Q coulombs. If the area through which the flux passes is A sq.m., then *electric flux density:*

$$D = \frac{Q}{A} \text{ coulombs per sq.m. } (\text{C/m}^2) \qquad (2.4)$$

2.3 PERMITTIVITY

Also known as the *dielectric coefficient*, it has a certain resemblance to conductivity in the d.c. circuit. Thus, from conductance $G = I/E$ we would expect permittivity to be equal to

$$\frac{\text{flux density}}{\text{field intensity}}.$$

Permittivity is denoted by ξ hence:

$$\xi = \frac{D}{E} \text{ farads per metre (F/m)} \qquad (2.5)$$

Hence for a given field intensity the flux density created is proportional to the permittivity. ξ is generally termed the *absolute permittivity* and it concerns any *dielectric* used between the plates. The absolute permittivity of any material:

$$\xi = \xi_0 \times \xi_r \qquad (2.6)$$

where ξ_0 is the permittivity of free space (a vacuum) and ξ_r is the *relative* permittivity of the material. ξ_r for air has a value of 1.00059 so generally is taken to have the same value as that for free space. For materials used in electronics ξ_r is greater than 1. ξ_0 has a value of 8.854×10^{-12} F/m (sometimes expressed as $10^{-9}/36\pi$) and remembering that in equations concerned with electrostatics the units of force, charge etc. are chosen arbitrarily, ξ_0 is simply a factor which has to be employed for the equations to make practical sense.

Note that in dielectric materials the electric force of the p.d. causes the electron orbits to distort but not to the extent that electrons are released. If the p.d. reaches such a value that this does occur then the dielectric breaks down and current flows through it.

2.4 COULOMB'S LAW

It was Coulomb(A3) who gave electronics one of its most important and fundamental laws. It is the law governing the force between charges *at rest* and states that for two like charges concentrated at points, the force of repulsion between them (force of attraction when the charges are unlike) is proportional to the product of their charges ($Q_1.Q_2$), inversely proportional to the square of the distance between them and also to the permittivity of the medium. As a formula:

$$F = \frac{Q_1 Q_2}{4\pi\xi d^2} \text{ newtons} \qquad (2.7)$$

(the Q's are in coulombs, d in metres and ξ in farads/metre).

2.5 CAPACITANCE

This is a measure of the ability to store electric charge. The unit of capacitance is the *farad* and a capacitor of this value stores a charge of one coulomb when a p.d. of one volt is present across its terminals. Thus:

$$Q = CV \text{ coulombs} \qquad (2.8)$$

where C is in farads and V in volts.

Hence,

$$C = \frac{Q}{V}$$

Now from Eqn.(2.2)

$$V = Ed$$

$$\therefore \quad C = \frac{Q}{Ed}$$

From Eqn. (2.4)

$$Q = DA$$

$$\therefore \quad C = \frac{DA}{Ed}$$

and from Eqn. (2.5)

$$E = \frac{D}{\xi}$$

$$\therefore \quad C = \frac{D\xi A}{Dd} = \frac{\xi A}{d} = \frac{\xi_0 \xi_r A}{d} \qquad (2.9)$$

This is for two plates only, for multiple-plate capacitors of n plates or N dielectric spaces, because the charge is stored in the dielectric material:

$$C = \frac{\xi_0 \xi_r A.N}{d} \tag{2.10}$$

or $$C = \frac{\xi_0 \xi_r A(n-1)}{d} \tag{2.11}$$

in farads when d (the thickness of the dielectric) is in metres and A is in square metres.

EXAMPLE 2.1:
An air-dielectric radio tuning capacitor has 25 plates each of area 2.6 sq. cm, separated by 0.15 mm. What is its capacitance?

$$A = 2.6 \times 10^{-4}\ \text{m}^2$$

$$d = 0.15 \times 10^{-3}\ \text{m}$$

$$n = 25$$

$$\xi_r = 1$$

From Eqn. (2.11)

$$C = \frac{\xi_0 A(n-1)}{d}$$

$$= \frac{8.854 \times 10^{-12} \times 2.6 \times 10^{-4} \times 24}{0.15 \times 10^{-3}} \times 10^{-12}\ \text{pF}$$

$$= 368\ \text{pF}$$

EXAMPLE 2.2:
A 470 pF mica dielectric capacitor has 5 plates of dimensions 27 x 22 mm and the dielectric relative permittivity is 4.47. What is the thickness of the dielectric?

$$C = 470 \times 10^{-12} \text{ F}$$

$$N = 4$$

$$A = 27 \times 22 \times 10^{-6} \text{ m}^2$$

$$\xi_r = 4.47$$

$$\xi_0 = 8.854 \times 10^{-12}$$

From Eqn. (2.10)

$$C = \frac{\xi_0 \xi_r A.N}{d}$$

$$\therefore \quad d = \frac{\xi_0 \xi_r A.N}{C}$$

$$\therefore \quad d = \frac{8.854 \times 10^{-12} \times 4.47 \times 27 \times 22 \times 10^{-6} \times 4}{470 \times 10^{-12}}$$

$$= 0.0002 \text{ m} = 0.2 \text{ mm}$$

2.5.1 Energy Stored in a Capacitor

Current flowing into capacitor plates creates a p.d. across them which sets up the electric field. This field contains energy which we might picture in terms of the electron orbits being stretched out towards the positive plate, rather like a spring which is always ready to give up the energy used in stretching it by returning to normal. To find the energy equation, consider a capacitor of C farads being charged by a constant current of I amps. Let it reach a voltage V after a time t.

Since $Q = It$ (Sect.1.4) and $Q = CV$ (Sect.2.5),

then $CV = It$ and $V = \dfrac{It}{C}$

Because the p.d. across the capacitor increases linearly with time, the average p.d. over the time interval t is $V/2$ so the average power taken from the supply is $VI/2$ watts. Neglecting losses this power becomes the energy stored in the capacitor.

The total energy, W stored during time t :

$W = \dfrac{VIt}{2}$ joules and total charge $Q = It$ coulombs

$\therefore \quad W = \dfrac{VQ}{2}$ and because $Q = CV$,

$$W = \tfrac{1}{2}CV^2 \text{ joules} \qquad (2.12)$$

EXAMPLE 2.3:

A parallel-plate air-dielectric capacitor has 18 plates each of area 2 sq. cm with a separation of 0.5 mm. It is given a charge of 0.003 coulombs. How much energy is stored?

$\xi_0 = 8.854 \times 10^{-12}$

$\xi_r = 1$

$A = 2 \times 10^{-4}$ m^2

$d = 0.5 \times 10^{-3}$ m

$n = 18$

From Eqn. (2.11)

$$\text{Capacitance, } C = \frac{\xi_0 \xi_r A(n-1)}{d} \text{ farads}$$

$$= \frac{8.854 \times 10^{-12} \times 2 \times 10^{-4} \times 17}{0.5 \times 10^{-3}}$$

$$= 6 \times 10^{-11} \text{ F } (60 \text{ pF})$$

From Eqn. (2.12)

$$\text{Energy stored} = \tfrac{1}{2}CV^2 = \frac{1}{2} \cdot \frac{C^2V^2}{C} = \frac{1}{2} \cdot \frac{Q^2}{C}$$

(since $Q^2 = C^2V^2$)

$$= \frac{1}{2} \cdot \frac{0.003^2}{6 \times 10^{-11}} = \frac{9 \times 10^{-6}}{2 \times 6 \times 10^{-11}} = 7.5 \times 10^4 \text{ joule}$$

2.6 SUMMARY OF KEY FORMULAE

A = area (m^2)

d = distance (m)

n = number of plates

N = number of dielectric spaces

R = resistance (Ω)

t = time (s)

V, v = voltage (V)

W (or E) = work or energy

ξ = permittivity

ξ_0 = permittivity of free space

ξ_r = relative permittivity

QUANTITY	FORMULA	UNIT	SECTION
Permittivity of free space	$\xi_0 = 8.854 \times 10^{-12}$	F/m	2.3
Force between two charges at rest	$F = \frac{Q_1 Q_2}{4\pi\xi d^2}$	N	2.4
Electric field strength	$E = \frac{F}{Q}$	N/C	2.1
	$E = \frac{V}{d}$	V/m	2.1
Electric flux density	$D = \frac{Q}{A}$	C/m^2	2.2
Absolute permittivity	$\xi = \xi_0 \xi_r = \frac{D}{E}$	F/m	2.3
Potential difference	$\text{p.d.} = \frac{Fd}{Q} = \frac{W}{Q}$	V	2.1.1

QUANTITY	FORMULA	UNIT	SECTION
Capacitance	$C = \frac{Q}{V} = \frac{\xi_0 \xi_r A}{d}$	F	2.5
Capacitance of multi-plate capacitor	$C = \frac{\xi_0 \xi_r AN}{d} = \frac{\xi_0 \xi_r A(n-1)}{d}$	F	2.5
Energy stored in a capacitor	$W \text{ (or } E) = ½ \cdot CV^2$	J	2.5.1
Capacitors in series	$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \ldots$	F	Book 1
Two capacitors in series	$C = \frac{C_1 C_2}{C_1 + C_2}$	F	Book 1
Capacitors in parallel	$C = C_1 + C_2 + C_3 + \ldots$	F	Book 1
Capacitive time constant	$v_c = V(1 - e^{(-t/CR)})$	V	Book 1

3. ELECTROMAGNETISM

Whereas in electrostatics there seems to be the comparative simplicity of two types of charge creating forces according to their polarities, it is worth noting that Coulomb's Law for electrostatics does not tell the whole story because as Section 2.4 shows, it deals with charges *at rest.* They never are but if we were to take account of such a refinement the chapter would expand out of all proportion. Electromagnetism cannot be considered in quite such a simple manner for it is considered that magnetic phenomena arise not only from electrons in orbit but also from the fact that they have axial spins. Again we must avoid much of the added complexity by being satisfied with knowing that magnetic effects arise whenever charges move although some little acquaintance first with electron spins and magnetism will not come amiss. This can be gained by observing the spins within, for example, the iron atom which as Table 1.1 shows, has K, L, M and N shells containing 2, 8, 14, 2 electrons respectively. The K, L and N shells all contain even numbers of electrons and in each of these half have spins in one direction, half in the opposite direction so cancelling the magnetic effects. In contrast, of the 14 electrons in shell M, 9 spin in one direction, 5 in the opposite hence the magnetic properties if iron arise from this shell. Indications concerning magnetism are thus gained from knowledge of the directions of electron spins.

In magnetism we call the two different concentrations of magnetic energy, *poles*, North (N) and South (S). They are bound by the invariable rule that like poles repel, unlike attract. As in electrostatics some convention is needed with regard to direction of flux between two poles and this is by imagining that a free N-pole can exist and the direction in which it would move indicates the direction of the so-called "lines of flux". The free pole would therefore move away from a N-pole towards a S. Within the generator of magnetism, as for example the simple bar-magnet shown in Figure 3.1, the flux direction must be S to N. This has its parallel in the direct current circuit where electrons flow in the circuit from the –ve pole of the battery to the +ve but within the battery from +ve to –ve.

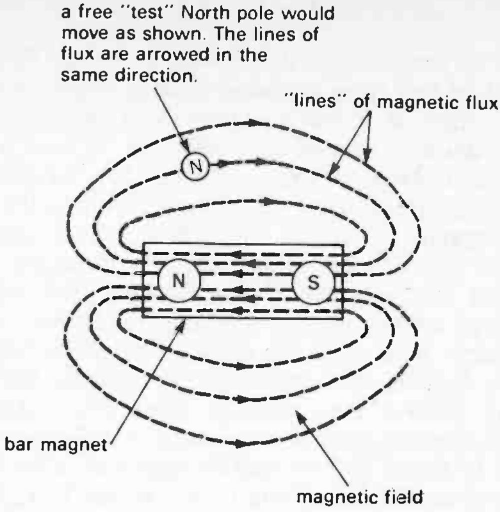


Fig. 3.1 Indicating the direction of a magnetic field

3.1 MAGNETIC FIELDS

Like force, energy, gravity and many other physical phenomena, a magnetic field is an invisible something which we cannot yet fully describe, although we are very aware of what it does and of the rules which govern its behaviour. Magnetic fields are set up by permanent magnets or by charges in motion in one particular direction within a conductor.

3.1.1 Magnetic Flux

For convenience it is considered that the magnetic effects experienced within a magnetic field arise from the presence of a *magnetic flux*. This is denoted by Φ and the S.I. unit is the weber[A3], (Wb). Magnetic flux is set up by a *magnetomotive force* which is discussed in Section 3.2.1.

The magnetic flux per unit area is the *magnetic flux density* (B) and magnetic flux density,

$$B = \frac{\Phi}{A} \qquad (3.1)$$

webers per square metre or tesla[(A3)] (T), A being the area in square metres.

EXAMPLE 3.1:
The pole-pieces of a moving coil loudspeaker magnet are separated by a 0.3 cm gap with a depth of 2.5 cm as shown in Figure 3.2. The diameter of the centre pole-piece is 2 cm with a gap flux density of 1.1 T. What is the total magnetic flux?

Mean diameter of flux path

$$= \frac{2.0 + 2.6}{2} = 2.3 \text{ cm} = 0.023 \text{ m}$$

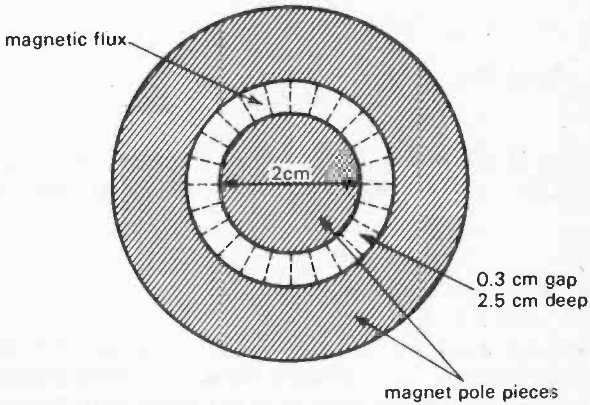


Fig. 3.2 Loudspeaker magnet system

∴ Mean flux path area,

$$A = \pi \times 0.023 \times 0.025 \text{ m}^2 ,$$

then from Eqn. (3.1) the total magentic flux,

$$\Phi = B \times A = 1.1 \times \pi \times 0.023 \times 0.025 = 0.002 \text{ Wb}$$

$$= 2 \text{ mWb}$$

3.2 THE MAGENTIC CIRCUIT

It is perhaps fortunate that we can use the electrical circuit as a guide to what occurs in a magnetic circuit. Flux has some similarities with current although current "flows" whereas flux does not. If therefore we liken flux to current, there must be some "magnetic e.m.f." in a magnetic circuit which is the driving force. This is known as the *magnetomotive force* (m.m.f.). Also the magnetic circuit can be expected to offer resistance to the creation of magnetic flux and this is its *reluctance.* The relationship follows the pattern of Ohm's Law and is therefore:

$$\text{Total flux} = \frac{\text{magnetomotive force}}{\text{total reluctance}}$$

Study of the constituents of this equation follows. It is of great help in understanding many aspects of electromagnetism.

3.2.1 Magnetomotive Force

It has been suggested that movement of charges gives rise to magnetic forces so because under the same external conditions all electrons behave similarly, then the effect must be proportional to the number of electrons flowing at a given point, that is, to the current. Hence the magnetomotive force varies directly as the current.

Accordingly, when current flows through a conductor, lines of magnetic flux are set up around it, if the conductor is wound as an air-cored coil it so happens that the fluxes produced by each turn are additive (Sect. 2.3 of Book 1) and from this it is evident that the magnetomotive force also varies directly as the number of turns (N). Therefore m.m.f.,

$$F = IN \text{ ampere-turns (At)} \tag{3.2}$$

[Note: Because N is simply a number, the SI dimension for F is strictly *amperes* (A). Ampere-turns (At) is used in this Chapter to avoid confusion with current but the SI symbols are added in brackets as a reminder.]

3.2.2 Magnetic Field Strength

From the above, if, for example one ampere of current flows through 100 turns of wire the m.m.f. is $F = IN = 100$ ampere-turns (A). In practice these may be tightly wound or alternatively spaced apart so distributing the m.m.f. over a greater length of coil. The distribution of m.m.f. along the coil is known as the *magnetic field strength*, *magnetic field intensity* or *magnetizing force*, defined as the m.m.f. per unit length of the magnetic circuit and denoted by H :

$$\text{magnetic field strength } H = \frac{\text{m.m.f.}}{\ell} = \frac{F}{\ell} = \frac{IN}{\ell} \tag{3.3}$$

ampere-turns per metre (A/m), where ℓ is the length of the coil in metres.

In the example where $F = 100$ At (A) for a winding 1 cm long,

$$H = \frac{100}{0.01} = 10{,}000 \text{ At/m} \quad \text{(A/m)},$$

for a winding 10 cm long,

$$H = \frac{100}{0.1} = 1{,}000 \text{ At/m} \quad \text{(A/m)}$$

so showing the concentration of H in the shorter coil.

3.2.3 Permeability

When a magnetizing force (H) is set up it is therefore capable of producing a magnetic flux (Φ) in a given material or medium and the flux is proportional to the force (provided that the material does not change). A measurement of the result is by *flux density* (B, Sect.3.1.1) hence $B \propto H$. This is brought into practical terms by rating materials by their permeability (μ) so that:

$$\text{flux density, } B = \mu \times H \text{ teslas (T)} \qquad (3.4)$$

As with permittivity (Sect.2.3) μ is split up into two components, μ_0 for free space (and generally for air also) and μ_r which is the *relative permeability* of a given material. μ is known as the *absolute permeability* and

$$\mu = \mu_0 \mu_r \qquad (3.5)$$

μ_0 has a value of $4\pi \times 10^{-7}$ and because it arises from the ratio B/H is in teslas/ampere-turns per metre [T/(A/m)]. However, because Wb/A is the unit of inductance (the practical aspects of inductance are considered in Book 1), by transposition μ can also be expressed in henrys per metre (H/m).

EXAMPLE 3.2:
A wooden ring (it is non-magnetic, therefore $\mu_r = 1$) of mean diameter 16.67 cm and cross-sectional area 5 cm² has a uniform winding on it of 1000 turns. With a current of 5 A flowing, what is (i) the magnetic field strength, (ii) the flux

density, (iii) the total flux? [An example of such a ring is shown in Figure 3.4(i).]

$$\ell = \pi d = \pi \times 16.67 \text{ cm}$$

$$A = 5 \times 10^{-4} \text{ m}^2$$

$$N = 1000$$

$$I = 5 \text{ A}$$

(i) From Eqn. (3.3), magnetic field strength,

$$H = \frac{IN}{\ell} = \frac{5 \times 1000}{\pi \times 16.67 \times 10^{-2}} = \frac{30{,}000}{\pi}$$

$$= 9549 \text{ At/m} \quad \text{(A/m)}$$

(ii) From Eqns. (3.4) and (3.5), magnetic flux density,

$$B = \mu_0 \mu_r H = 4\pi \times 10^{-7} \times \frac{30{,}000}{\pi}$$

$$= 0.012 \text{ T}$$

(iii) From Eqn. (3.1), total flux,

$$\Phi = B \times A = 0.012 \times 5 \times 10^{-4} \text{ Wb} = 6\ \mu\text{Wb}$$

EXAMPLE 3.3:
884 turns are wound on a core of cross-sectional area 1 cm^2 and relative permeability 30, the length of the winding being 10 cm. What flux is produced with a winding current of 3 A?

$$N = 884$$

$$\ell = 0.1 \text{ m}$$

$$A = 10^{-4} \text{ m}^2$$

$$I = 3 \text{ A}$$

$$\mu_r = 30$$

From Eqn. (3.3)

$$H = \frac{IN}{\ell} = \frac{3 \times 884}{0.1} = 26520 \text{ At/m} \quad \text{(A/m)}$$

From Eqns. (3.4) and (3.5)

$$\mu_0\mu_r = \frac{B}{H} \quad \therefore B = \mu_0\mu_r H = 4\pi \times 10^{-7} \times 30 \times 26520$$

$$= 1.0 \text{ T}$$

From Eqn. (3.1)

$$\Phi = B.A = 1.0 \times 10^{-4} = 0.1 \text{ mWb}$$

3.2.3.1 Variation of Permeability with Magnetizing Force
Because the full magnetic theory is quite complex, we have to be content with explanations which are in fact only generalizations. As such, magnetization is considered as the alignment of minute magnets called *domains* within a *ferromagnetic* material (that with a high value of μ_r). In this process previously unaligned domains (i.e. lying in random directions and therefore exhibiting no overall magnetic properties) are moved into similar directions by a field applied externally just as a compass needle is swung by the Earth's magnetic field. The degree of alignment naturally depends on the strength of the field applied and as the domains move more and more into alignment so the flux density in the material increases. Graphically the effect is illustrated by Figure 3.3(i) from which it is evident that the ratio B/H

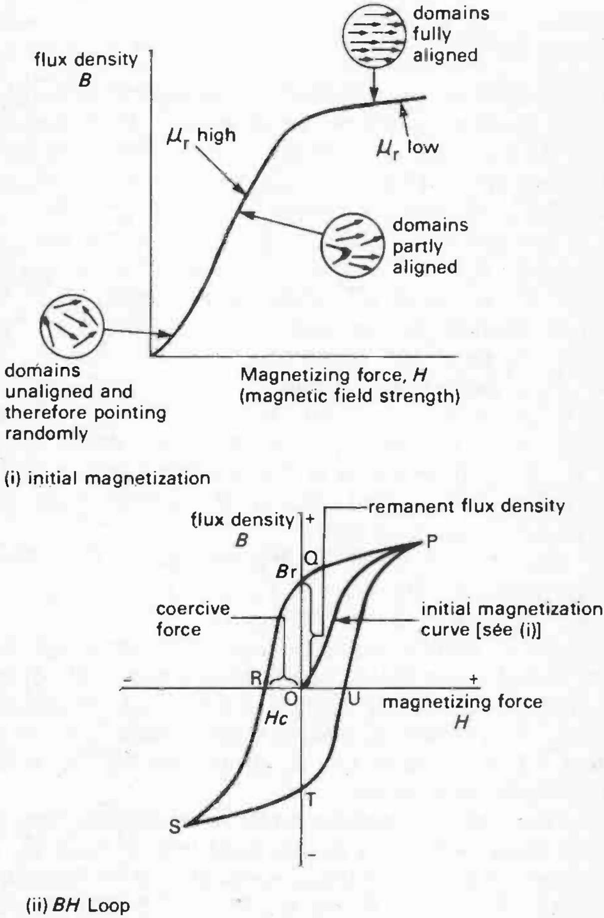


Fig. 3.3 Magnetization curves

varies greatly, i.e. μ_r is far from constant. It can be calculated at any point directly from Equations (3.4) and (3.5):

$$\mu_r = \frac{B}{\mu_0 H} \tag{3.6}$$

If a sample of unmagnetized ferromagnetic material is subjected to an increasing magnetizing force, H, it therefore follows the curve OP as in Figure 3.3(ii) as predicted by (i) of the same figure. If after reaching the point P, H is reduced to zero the curve does not retrace P to O but falls back to Q, that is, the specimen remains magnetized with a *remanent* flux density B_r. H must next be applied in the opposite direction if B_r is to be reduced to zero (point R on the curve) so clearing the specimen of all magnetization. The "negative" value of H to do this is known as the *coercive force*, H_c. Increasing H further in this direction eventually brings magnetic saturation again but with the poles changed over (S). At T, H is again zero with the same value of remanence, B_r but with flux lines in the opposite direction to that above and at U a "positive' value of H_c brings the specimen again to the demagnetized state.

Throughout the cycle of H, it is seen that B lags behind, for this reason the BH curve of Figure 3.3(ii) is known as a *hysteresis loop* (from Greek, coming after).

Permanent magnets not only need to retain a high flux density but also to resist demagnetization hence both B_r and H_c must be high. The material is known as magnetically *hard*. B_r for practical permanent magnets varies from about 0.5 to 1.4 T with a range for H_c from some 40,000 to well above 150,000 At/m (A/m).

Materials used in transformers and electromagnets have the opposite requirement, that is, they need low B_r and H_c so that magnetism is not retained (magnetically *soft*). Commonly used materials range in B_r up to about 0.4 T with H_c only up to a few tens of ampere-turns per metre (A/m).

3.2.4 Reluctance

The remaining parameter required for magnetic circuit calculations is *reluctance*, the resistance offered by the various paths in the circuit to the creation of magnetic flux. Consider the elementary magnetic circuit shown in Figure 3.4(i):

$$\text{magnetomotive force, } F = IN$$

and

$$\text{magnetic field strength, } H = \frac{IN}{\ell}$$

from which

$$F = H\ell \qquad (3.7)$$

The reluctance S must fit into the expression

$$\text{total flux} = \frac{\text{magnetomotive force}}{\text{reluctance}} \qquad \text{(Sect.3.2)}$$

i.e.

$$\text{total flux } \Phi = \frac{F}{S} \qquad (3.8)$$

from this and Eqn. (3.7):

$$S = \frac{H\ell}{\Phi}$$

but from Eqn. (3.1)

$$\Phi = B.A$$

and from Eqn. (3.4)

$$B = \mu H$$

$$\therefore \quad S = \frac{H\ell}{\mu HA} = \frac{\ell}{\mu A}$$

$$\therefore \quad S = \frac{\ell}{\mu_0 \mu_r A} \tag{3.9}$$

and because

$$\text{reluctance} = \frac{\text{magnetomotive force}}{\text{flux}}$$

it is expressed in ampere-turns per weber (A/Wb).

3.2.4.1 Reluctance in Composite Magnetic Circuits

As with electrical circuits in which resistances can be in series or in parallel or in combinations of these, magnetic reluctances exist similarly. In analysis the series circuit is slightly less complicated and it may consist of two or more sections differing in length, area of cross-section or permeability. Neglecting magnetic leakage, the same flux is maintained throughout the circuit and the total reluctance of the complete circuit is the sum of the reluctances of the different parts. Just as in the electrical case where $R = R_1 + R_2 + R_3 + \ldots$ so in the magnetic circuit $S = S_1 + S_2 + S_3 + \ldots$

For a composite magnetic circuit which can be treated as series:

$$\text{total flux } \Phi = \frac{\text{m.m.f.}}{S_1 + S_2 + S_3 + \ldots}$$

$$= \frac{\text{m.m.f.}}{\dfrac{\ell_1}{\mu_0 \mu_{r_1} A_1} + \dfrac{\ell_2}{\mu_0 \mu_{r_2} A_2} + \dfrac{\ell_3}{\mu_0 \mu_{r_3} A_3} + \ldots} \tag{3.10}$$

Commonly found examples are given by relays, magnetic record/replay heads and electromagnets in which the main magnetic path is interrupted by a small air-gap, represented generally as in Figure 3.4(ii). For simplicity it is assumed that all the flux passes straight across the gap instead of bulging out or *fringing* as happens with the longer gaps.

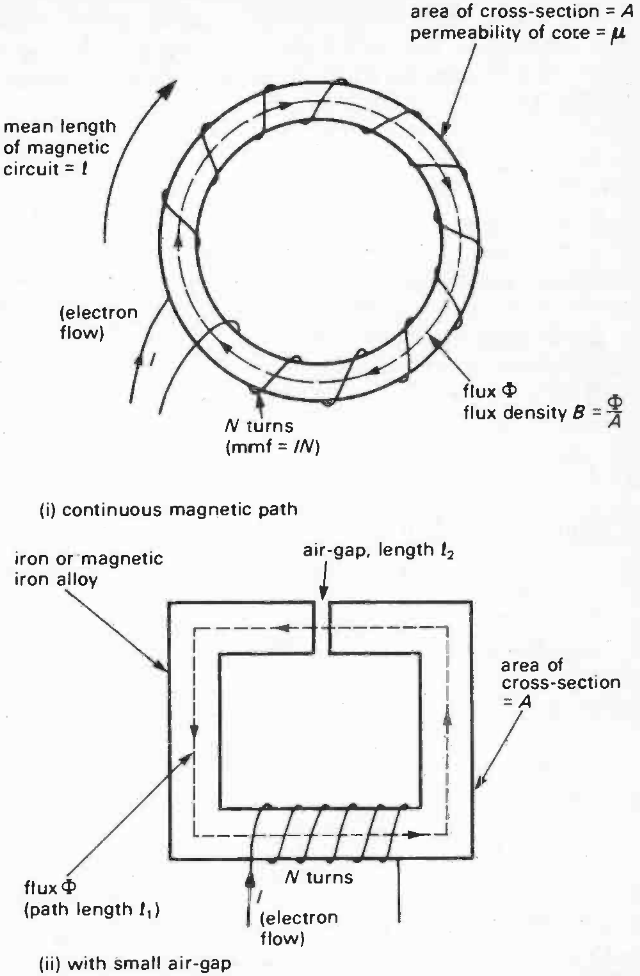


Fig. 3.4 Simple magnetic circuits

EXAMPLE 3.4:

For the magnetic circuit of Figure 3.4(ii) in which $N = 800$, length of magnetic path = 12.4 cm, cross-sectional area = 2 cm^2, relative permeability of iron = 2000, calculate the current required in the winding to produce a flux density in the iron of 1.2 tesla, (i) when no air-gap exists, (ii) when a 2 mm slot is cut.

$N = 800$

$\ell = 0.124$ m

$A = 2 \times 10^{-4}$ m^2

$\mu_r = 2000$

$B = 1.2$ T

(i) no air gap:–

from Eqns (3.3) and (3.6)

$$F = \frac{B\ell}{\mu_0 \mu_r}$$

and from Eqn. (3.2)

$$I = \frac{F}{N}$$

then $$I = \frac{B\ell}{\mu_0 \mu_r N} = \frac{1.2 \times 0.124}{4\pi \times 10^{-7} \times 2000 \times 800} = 74 \text{ mA}$$;

(ii) with the 2 mm air-gap the problem can be solved by calculating the additional current needed to overcome the added reluctance of the air-gap for which $\ell = 2$ mm, $\mu_r = 1$.

Then

$$I = \frac{B\ell}{\mu_0 \mu_r N} = \frac{1.2 \times 2 \times 10^{-3}}{4\pi \times 10^{-7} \times 800} = 2.39 \text{ A}$$

and total current $= 0.074 + 2.39 = 2.46$ A

Thus although the gap length is less than one-sixtieth of that of the iron, the decrease in reluctance on this account is more than outweighed by the considerably lower gap relative permeability.

Alternatively use could be made of

$$\Phi = \frac{F}{S}$$

directly. For example, from Eqn. (3.9) reluctance of iron path (neglecting length of gap),

$$S_i = \frac{\ell}{\mu_0 \mu_r A} = \frac{0.124}{4\pi \times 10^{-7} \times 2000 \times 2 \times 10^{-4}}$$

$$= 2.467 \times 10^5 \text{ At/Wb (A/Wb)}$$

reluctance of air path,

$$S_a = \frac{2 \times 10^{-3}}{4\pi \times 10^{-7} \times 2 \times 10^{-4}}$$

$$= 7.96 \times 10^6 \text{ At/Wb (A/Wb)}$$

total reluctance, $S = S_i + S_a$

$$= 8.21 \times 10^6 \text{ At/Wb (A/Wb)}$$

Also from Eqn. (3.1)

$$\Phi = B \times A = 1.2 \times 2 \times 10^{-4} = 2.4 \times 10^{-4} \text{ Wb}$$

and from Eqns. (3.8) and (3.2)

$$\Phi = \frac{F}{S} = \frac{IN}{S}$$

$$\therefore \quad I = \frac{\Phi S}{N} = \frac{2.4 \times 10^{-4} \times 8.21 \times 10^{6}}{800} = 2.46 \text{ A}$$

slightly more complicated because the cross-sectional area is brought in, only to finally cancel out.

The parallel magnetic circuit follows the rules for resistances in parallel hence for several paths of reluctance S_1, S_2, S_3, . . . it can be shown that

$$\frac{1}{S} = \frac{1}{S_1} + \frac{1}{S_2} + \frac{1}{S_3} + \ldots$$

hence for two reluctances in parallel

$$S = \frac{S_1 S_2}{S_1 + S_2} \tag{3.11}$$

and for three

$$S = \frac{S_1 S_2 S_3}{S_1 S_2 + S_1 S_3 + S_2 S_3} \tag{3.12}$$

The reciprocal of reluctance is known as the *permeance* (symbol Λ),

$$\text{permeance, } \Lambda = \frac{1}{S} = \frac{\mu_0 \mu_r A}{\ell} \text{ Wb/At (Wb/A)} \tag{3.13}$$

3.2.5 Pull of an Electromagnet

From the formulae already developed, the flux density, B for a magnetic circuit containing a winding can be calculated from which a further simple calculation gives the pull of an electromagnet. The force exerted by one over an area of flux A square metres is given by:

$$F = \frac{B^2 A}{2\mu_0} \text{ newtons} \tag{3.14}$$

i.e. approximately by $B^2 A \times 3.98 \times 10^5$ N .

EXAMPLE 3.5:
What weight will an electromagnet support if its core flux density is 1.363 tesla and the circular pole face has a diameter of 13 cm?

$$A = \pi(0.065)^2 \text{ m}^2$$

$$B = 1.363 \text{ T}$$

From Eqn. (3.14)

$$F = \frac{B^2 A}{2\mu_0} = \frac{1.363^2 \times \pi \times 0.065^2}{2 \times 4\pi \times 10^{-7}}$$

$$= 9.81 \times 10^3 \text{ newtons}$$

∴ weight supported (see Sect.1,1,1)

$$= \frac{9.81 \times 10^3}{9.81} \text{ kg} = 1000 \text{ kg or 1 tonne}$$

3.3 ELECTROMAGNETIC INDUCTION

The discovery of electromagnetic induction by Faraday[A3] gave us the amazing power of electricity. He found that when there is relative motion between a conductor and a magnetic field, that is, the conductor "cuts" lines of flux or vice versa, then an e.m.f. is induced in the conductor. The field releases electrons and causes them to move in one particular direction. Faraday's Law of electromagnetic induction states that the magnitude of the e.m.f. induced in a conductor is proportional to the *rate of change* of the magnetic flux linking with it, perhaps visualized in terms of the impetus given to the electrons. In practical terms the e.m.f. induced is one volt when the flux changes at the rate of one weber per second, that is:

$$\text{e.m.f. } e = \frac{\text{total flux cut (Wb)}}{\text{time (secs)}} = \frac{d\Phi}{dt} \qquad (3.15)$$

[($d\Phi/dt$) is the rate of change of flux with time in Wb/s.]

Consider Figure 3.5(i) in which a conductor of length ℓ moves a distance d across the lines of flux of a uniform magnetic field of flux density B at right angles. Then area of flux cut by conductor

$$= \ell \times d$$

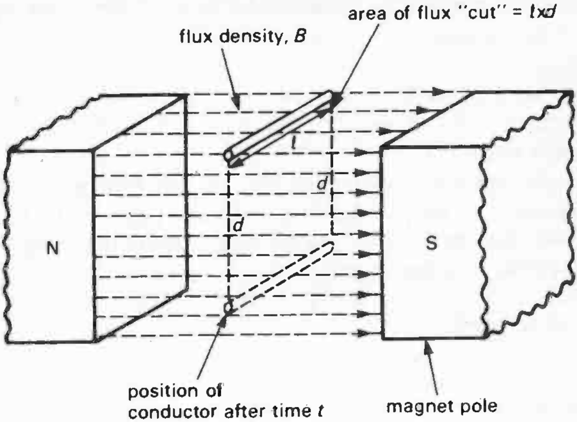
and total flux cut,

$$\Phi = B\ell d$$

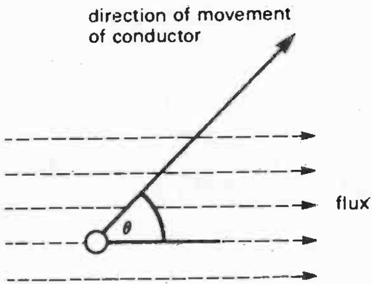
and for constant velocity,

$$v = \frac{d}{t}$$

$$\therefore \quad e = \frac{B\ell d}{t}, \text{ i.e. } e = B\ell v \qquad (3.16)$$



(i) conductor moving in a magnetic field



(ii) conductor moving at an angle to flux lines

Fig. 3.5 Electromagnetic induction

(e is in volts when B is in tesla, ℓ in metres and v in metres/sec.).

It can also be shown that if the conductor moves at an angle θ to the lines of magnetic flux as shown in Figure 3.5(ii), in this case:

$$e = B\ell v \sin\theta \qquad (3.17)$$

EXAMPLE 3.6:
A straight conductor of length 100 cm moves across a magnetic field of flux density 0.8 T but at some angle to the lines of flux. Its velocity is 5 m/s and the e.m.f. generated is measured as 0.2 V. What is the angle?

$e = 0.2$ V

$B = 0.8$ T

$\ell = 0.1$ m

$v = 5$ m/s

$\theta = ?$

From Eqn. (3.17):

$$\sin\theta = \frac{e}{B\ell v} = \frac{0.2}{0.8 \times 0.1 \times 5} = 0.5$$

$$\therefore \quad \theta = \sin^{-1} 0.5 = 30^\circ$$

For a coil of N turns moving in a magnetic field or with magnetic flux moving across the winding; because all turns are in series:

$$\text{Average induced e.m.f., } e = N\frac{d\Phi}{dt} \qquad (3.18)$$

EXAMPLE 3.7:
An air-cored coil has a winding of 100 turns. A flux gener-

ated by a concentric coil changes from 50 to 300 mWb in 100 ms. What e.m.f. is generated in the first winding?

$$\text{Flux change} = \Phi_1 - \Phi_2 = 300 - 50 = 250 \text{ mWb.}$$

From Eqn. (3.18):

$$e = N\frac{d\Phi}{dt} = 100 \times \frac{0.25}{100 \times 10^{-3}} = 250 \text{ V}$$

Using this simple formula it is possible to appreciate how a car ignition coil generates a high voltage (from 4000 up to 20,000 V).

EXAMPLE 3.8:
Suppose a flux of 2 mWb collapses in the secondary winding of a coil of 5000 turns in 1 ms, then from Eqn. (3.18):

$$e = 5000 \times \frac{2 \times 10^{-3}}{1 \times 10^{-3}} = 10{,}000 \text{ volts}$$

3.4 ELECTROMAGNETIC MOTOR ACTION

When current flows in a wire situated in a magnetic field the interaction between the magnetic properties of the moving charges and those of the field itself result in a force on the conductor. If the conductor is at right-angles to the direction of the flux the force is given by:

$$F = BI\ell \text{ newtons} \qquad (3.19)$$

(B in tesla, I in amperes, ℓ in metres).

EXAMPLE 3.9:
Within a dc electric motor the field flux density is 0.8 T. An armature conductor is 50 cm long. If it carries a current of 50 A what is the force on it?

$B = 0.8$ T

$\ell = 0.5$ m

$I = 50$ A

$F = ?$

From Eqn. (3.19), force on conductor,

$$F = BI\ell = 0.8 \times 50 \times 0.5 = 20 \text{ N}.$$

3.4.1 Torque on a Moving Coil

Moving coil measuring instruments are based on the principle of rotation of a current-carrying coil in a magnetic field. Consider the coil shown in Figure 3.6(i) and assume that it is carrying a current of I amperes. From Equation (3.19) the force on each side of the coil is equal to $BI\ell \times N$ where N is the number of conductors, and on both sides together, $2BI\ell N$ with the *moment of force* or *torque* (turning moment):

$$T = 2BI\ell N \times \frac{b}{2} \text{ newton-metres}$$

where b is the breadth of the coil.

(Torque is found by multiplying the magnitude of a force by its perpendicular distance from the axis of rotation.)

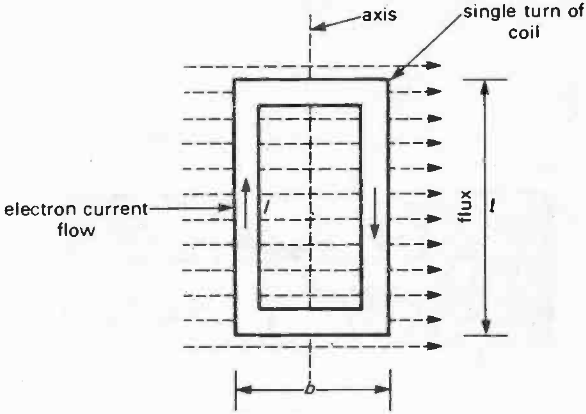
As the coil rotates the torque decreases and at any angle θ as shown in Figure 3.6(ii):

$$T = BI\ell Nb \cos\theta \text{ newton-metres (Nm)}$$

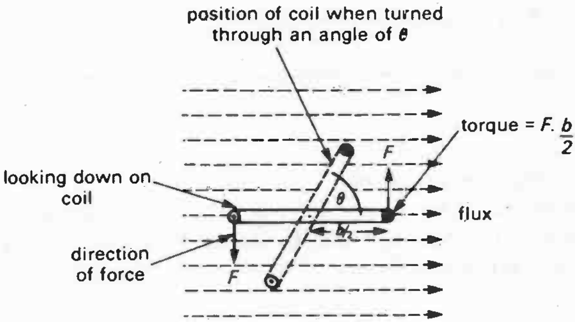
But $b \times \ell$ is the area, A of the coil, hence:

$$T = BINA \cos\theta \text{ Nm} \qquad (3.20)$$

so when the plane of the coil lies in the direction of the flux ($\theta = 0$, $\cos\theta = 1$), T is maximum, falling to zero when $\theta = 90°$.



(i) pivotted coil in a magnetic field



(ii) coil turned through an angle

Fig. 3.6   Torque on a moving coil

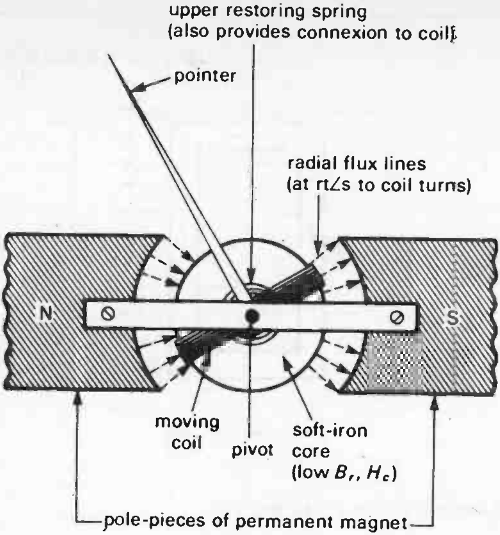


Fig. 3.7 Moving-coil meter movement

3.4.1.1 Moving-Coil Meters

In moving-coil movements the coil is arranged between the magnet pole-pieces with radial flux as in Figure 3.7, hence at all working positions of the coil $\theta = 0$ whereupon

$$T = BINA \text{ Nm}$$

B, N and A are constants for any particular movement, hence $T \propto I$. The controlling torque exerted by the restoring springs is proportional to the pointer deflexion hence the latter is proportional to the coil current, a most useful feature of the moving-coil instrument.

3.5 FLEMING'S RULES

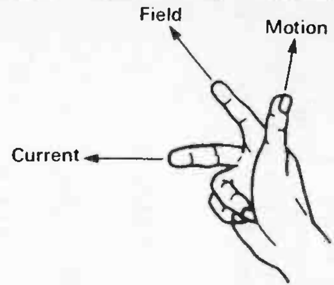
At the workshop bench there is occasionally the need to determine the direction of motion in a motor, electromagnet, meter movement etc. Armed with a cheap pocket compass to sort out magnetic field directions, Fleming's[A3] rules then come to our aid. There are, however, two ways in which current flow is considered. There is the "conventional" in which current is said to flow from positive to negative, born long ago and nowadays considered to be incorrect. In line with the more intimate knowledge of the electron which we have today is the argument that, although positive atoms can move, current is essentially a flow of electrons, therefore from negative to positive (Sect.1.4). Fleming's rules linking the directions of flux, current and motion were formulated on the earlier theory and these are still adhered to by many. For the modern way of thinking we can modify Fleming's rules, without forgetting that the original idea was his. The ultimate choice is the reader's so both sets of rules follow.

Figure 3.8(i) shows how the fingers are extended with thumb, first finger (forefinger or index finger) and second or middle finger mutually at right-angles. Ignore the remaining two fingers. Then always:

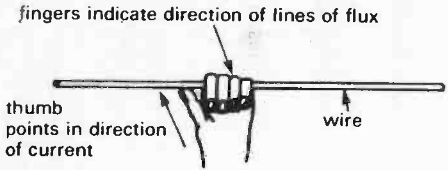
thuMb	indicates direction of Motion
First Finger	indicates direction of Flux
sEcond or Middle Finger	indicates direction of EMF
second or mIddle fInger	indicates direction of current (when cct completed),

the last one added because the direction of a current is taken to be the same as that of the e.m.f. which causes it.

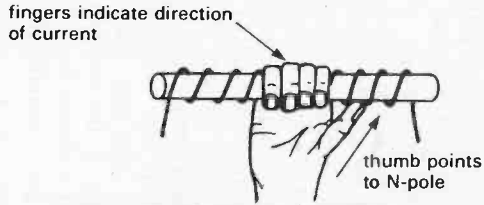
The first part of Table 3.1 sets out which rules are appropriate.



(i) using the right hand with thumb and first two fingers mutually at right-angles.



(ii) using the left hand to indicate relationship between current and flux directions (electron flow)



(iii) using the right hand to indicate relationship between current direction and magnetic polarity (conventional current)

Note: In all cases the opposite hand is used similarly and the fingers give the same indications.

Fig. 3.8 Rules of electromagnetism

Table 3.1 **CHOICE OF HAND IN LEFT- AND RIGHT-HAND RULES**

	CURRENT DIRECTION		
	Electron Flow –ve to +ve	Conventional +ve to –ve	
Generators	Left	Right	Fig.3.8(i)
Motors	Right	Left	Fig.3.8(i)
Mnemonic	"GLaMouR" (generators left, motors right)	"GRuMbLe" (generators rights, motors left)	
Magnetic field around a wire	Left	Right	Fig.3.8(ii)
Coils	Left	Right	Fig.3.8(iii)

3.5.1 Rules Relating Flux and Current

In conjunction with Figure 3.8 the remainder of Table 3.1 shows which hand is used to determine:

the direction of the magnetic lines of flux encircling a wire carrying a current [Fig.3.8(ii)]

the magnetic poles produced by a coil when a current flows through it [Fig.3.8(iii)].

3.6 SUMMARY OF KEY FORMULAE

A = area (m^2)

f = frequency (Hz)

I = current (A)

ℓ = length (m)

N = number of turns

R = resistance (Ω)

t = time (s)

v = velocity (m/s)

θ = angle between conductor and flux

μ_0 = permeability of free space [$4\pi \times 10^{-7}$ T/(At/m)]

μ_r = relative permeability

QUANTITY	FORMULA	UNIT	SECTION
Magnetic field strength	$H = \frac{IN}{\ell}$	At/m (or A/m)	3.2.2
Magnetomotive force	$F = IN = H\ell$	At (or A)	3.2.1
Absolute permeability	$\mu = \mu_0\mu_r$	T/(At/m) or T/(A/m) or H/m	3.2.3
	$\mu = \frac{B}{H}$		3.2.3.1
Reluctance	$S = \frac{\ell}{\mu A}$	At/Wb (or A/Wb)	3.2.4
Magnetic flux	$\Phi = \frac{F}{S}$	Wb	3.2
Composite series circuit	$\Phi = \frac{F}{S_1 + S_2 + S_3 + \ldots}$	Wb	3.2.4.1
Flux density	$B = \frac{\Phi}{A}$	T	3.1.1

QUANTITY	FORMULA	UNIT	SECTION
Permeance	$\Lambda = \frac{\mu_0 \mu_r A}{\ell}$	Wb/At (or Wb/A)	3.2.4.1
Pull of electromagnet	$F = \frac{B^2 A}{2\mu_0}$	N	3.2.5
EMF induced in a conductor	$e = B\ell v \sin\theta$	V	3.3
Force on a conductor	$F = BI\ell$	N	3.4
Torque on a moving-coil	$T = BINA \cos\theta$	Nm	3.4.1
Inductance	$L = \frac{\mu_0 \mu_r N^2 A}{\ell}$	H	Book 1
Inductors in series (not coupled)	$L = L_1 + L_2 + L_3 + \ldots$	H	Book 1
Mutual inductance	$M = \sqrt{L_1 L_2}$	H	Book 1
Two inductors in series (magnetically coupled)	$L = L_1 + L_2 \pm 2M$	H	Book 1

QUANTITY	FORMULA	UNIT	SECTION
Inductors in parallel	$\frac{1}{L} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} + \dots$	H	Book 1
Coupling coefficient	$k = M\sqrt{L_1 L_2}$		Book 1
Transformer flux (p = primary)	$\Phi = \frac{\mu_0 \mu_r N_p I_p A}{\ell}$	Wb	Book 1
Transformer voltage	$e = 4.44 N \Phi_{max} f$	V	Book 1
Inductive time constant	$i = I(1 - e^{(-Rt/L)})$	A	Book 1

4. COMPLEX NUMBERS

When similar events occur within a circuit (for example, voltages or currents reaching their maximum values) but at different *times* then the customary mathematical equation fails to cope. One alternative is the *phasor* diagram in which all time differences are illustrated by their equivalent angles based on $\theta = 2\pi f \times t$ radians. Such a diagram has much to offer visually and can be solved approximately by measurement or accurately by trigonometry. Mathematics does however provide an additional technique by which such circuits can be analysed and although seemingly complicated at first sight it is merely a system which operates only on the rectangular coordinates of phasors and in so doing enables equations to be built up in which the two coordinates are kept apart. Some revision on the facilities of phasor diagrams may be gleaned from Book 1.

4.1 CARTESIAN AND POLAR COORDINATES

Descartes[(A3)] first showed how any point in a plane can be defined by two numbers. The numbers are appropriately known as *cartesian coordinates* and are relative to two axes, in electronics invariably at right angles, and with the x-axis at 3 o'clock and the y axis most frequently at 12 o'clock. These numbers are in fact the x and y values we use when plotting a graph as illustrated in Figure 4.1(i) where P is determined by $x = 4$, $y = 3$ [P = (4,3)].

Any point can also be defined by its *polar coordinates*, the distance the point is from the origin (r), known as the *modulus* (Latin, measure) and the angle (θ) it makes with one of the axes (usually the x), called the *argument* (a mathematical term for the angle on which the calculation of another quantity depends), as shown in (ii).

Comparing (i) and (ii) of the Figure, since P is in the same position relative to the axes, $x = r \cos \theta$, $y = r \sin \theta$, $r = \sqrt{x^2 + y^2}$, $\tan \theta = y/x$ (so if $x = 4$, $y = 3$, then $r = 5$, $\theta = 36.9°$). Conversion from one form to the other is therefore straightforward. The important feature to note is that if the line OP in (ii) is a phasor then it can be *resolved* (i.e. converted) into two separate components, $r \cos \theta$ and $r \sin \theta$ and these two numbers determine the phasor completely when they are relative to a pair of axes at right angles and passing through the same origin.

The coordinates demonstrated in Figure 4.1(i) and (ii) are for positive numbers only and they are said to be in the *first*

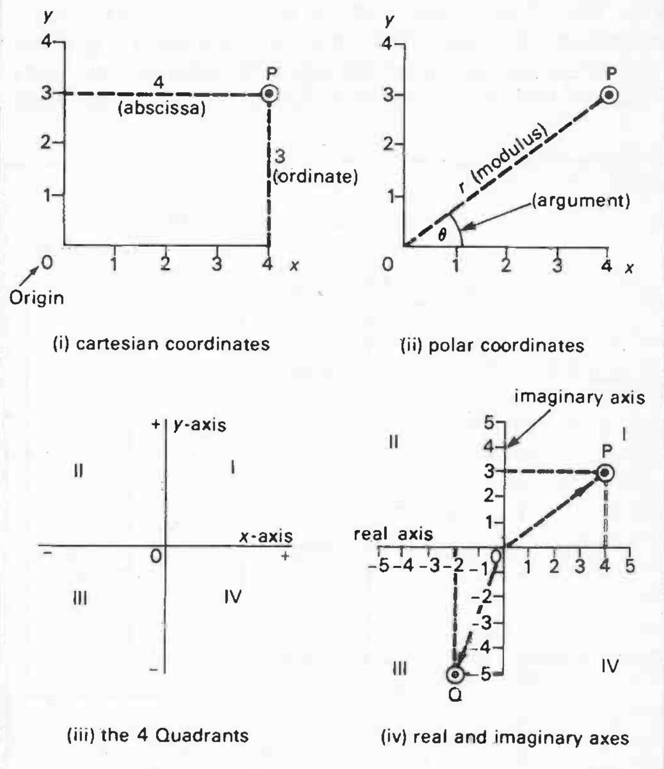
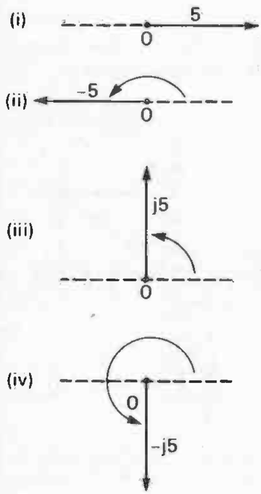


Fig. 4.1 Defining the position of a point on a plane

quadrant. The remaining quadrants follow in an anticlockwise direction as in Figure 4.1(iii). As an example, when both cartesian coordinates are negative, the point they represent lies in the third quadrant.

4.2 THE j-OPERATOR

Figure 4.2(i) shows a phasor of, for example, 5 units at 0° to the x-axis, its polar coordinates are therefore $5\ \angle 0°$. For 0° we use instead an unseen *operator* of +1, i.e. 5 x (+1) but for convenience we usually drop both the + and the 1. However in (ii) the phasor has turned through 180° and now has a value of $5\ \angle 180°$ or 5 x (–1), in effect the operator is –1, but again



polar coordinates	operator(s)	j notation
$5\angle 0°$,	+1	+5
$5\angle 180°$,	-1	-5
$5\angle 90°$,	j	j5
$5\angle 270°$,	j and -1	-j5

Fig. 4.2 Rotating a phasor

we drop the 1. Thus multiplying a phasor by -1 is equivalent to rotating it through 180°. To rotate a phasor through 90° a j-operator is used so that when it is in the position shown in (iii) it is marked as j5 and equally if a j5 phasor is shifted through 180°, this is equivalent to its being multiplied by -1 resulting in $-j5$ as shown in (iv). Thus for 180° a phasor is twice rotated through 90° i.e. multiplied by j and again by j (j^2). This is equal to -1 so $j = \sqrt{-1}$ which appears to be an oddity because it has no real solution.

4.2.1 Real and Imaginary Components

In circuit analysis we forget about x and y axes but use instead the labels *real* and *imaginary*. The exact meaning of the second term may escape us at first but we are not the only ones in trouble, mathematicians first faced the problem many years ago. Because the square of both positive and negative numbers is positive then there can be no real root of a negative number, for example, $x^2 = -1$ has no normal solution. This is highlighted by the standard formula for solving a quadratic equation [see Eqn.(A2.3)] an example being $x^2 - 4x + 13$ which when solved gives $x = 2 \pm \sqrt{-9}$. Accordingly an "operator" was invented which when squared equalled -1 and it was labelled "i". Using the same example, $3i \times 3i = -9$, i.e. $\sqrt{-9} = 3i$. At the time there was much confusion over this breakdown of reality so these numbers were called "imaginary" and quantities with both real and imaginary parts were said to be "complex". But for electronics "imaginary" is not perhaps the best choice of word for capacitance and inductance give rise to such numbers and there is nothing imaginary about them, they are every bit as effective as the so-called real numbers. In electronics we cannot use the mathematician's "i" because there is duplication with current, hence we use "j".

The four quadrants around the real and imaginary axes are illustrated in Figure 4.1(iv) and now point P can be expressed in complex notation as $(4 + j3)$, the 4 is real and the 3 imaginary. Equally point Q is fixed by $(-2 - j5)$ or $-(2 + j5)$.

The lengths of the phasors OP and OQ are easily calculated from their complex numbers for

$$OP = \sqrt{4^2 + 3^2} = 5\,,$$

$$OQ = \sqrt{2^2 + 5^2} = 5.4\,.$$

These are lengths only, the phasor positions in the quadrant are revealed by the angles made with the real axis for

$$\tan\theta = \frac{\text{imaginary component}}{\text{real component}}$$

which for OP makes θ,

$$\tan^{-1}\frac{3}{4} = 36.9°$$

and for OQ makes θ,

$$\tan^{-1}\frac{5}{2} = 68.2°$$

[or from the reference (real, positive) axis, 180 + 68.2 = 248.2°].

Thus it is possible to change from j to polar form for any phasor by use of the general expression:

$$r\angle\theta = \sqrt{a^2 + b^2}\ \tan^{-1}\frac{b}{a} \qquad (4.1)$$

where $r\angle\theta$ is the polar form and a represents its real component and b the imaginary.

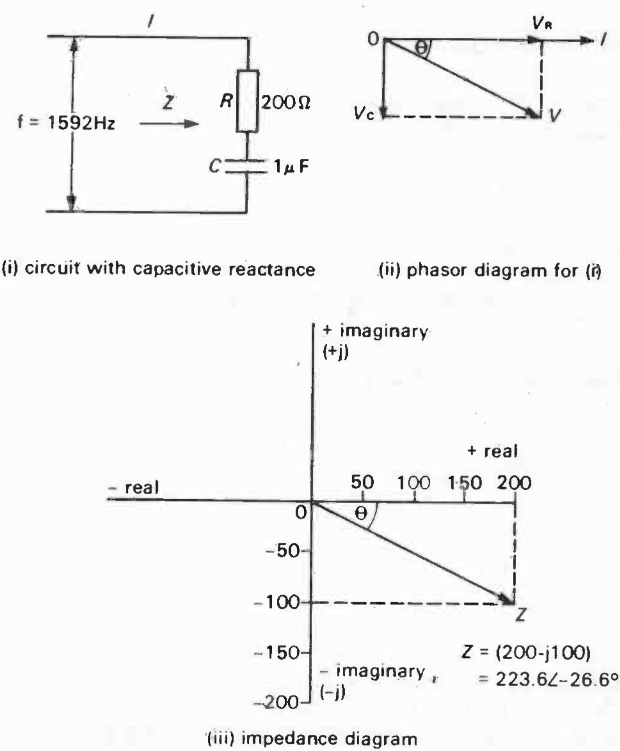
Changing from polar to j follows from the previous general discussion:

$$a = r\cos\theta \qquad b = r\sin\theta \qquad (4.2)$$

EXAMPLE 4.1:
For the circuit of Figure 4.3(i), express the impedance in both j and polar forms.

$$R = 200\ \Omega$$

$$X_c = -\frac{1}{\omega C} = -\frac{10^6}{2\pi \times 1592} = -100\ \Omega$$



(i) circuit with capacitive reactance

(ii) phasor diagram for (i)

(iii) impedance diagram

Fig. 4.3 Impedance of reactive circuit

For revision we first draw the phasor diagram as in Figure 4.3(ii). The current I is common to both R and C and so becomes the reference phasor. V_R $(= IR)$ is in phase with I and $V_C = 1/X_C$ lags on I by $90°$. The resultant phasor for V shows that the current leads the voltage by an angle of less than $90°$ as would be expected for an RC circuit. In j notation:

$$V = V_R - jV_C$$

(V_R is on the real axis and V_C on the imaginary). Dividing throughout by I:

$$Z = R - jX_C$$

leading to an impedance phasor OZ in the fourth quadrant as in Figure 4.3(iii). Putting in circuit values:

$$Z = 200 - j100\ \Omega\ .$$

Also from Eqn. (4.1):

$$|Z| = \sqrt{200^2 + (-100)^2} = 223.6\ \Omega$$

($|Z|$ indicates the modulus of Z).

$$\theta = \tan^{-1} \frac{-100}{200} = -26.6°$$

i.e. in polar form

$$Z = 223.6 \angle -26.6°\ .$$

Finally we can check back by using Eqn. (4.2):

$$R = Z \cos \theta = 223.6 \cos -26.6° = 200\ \Omega$$

$$X_C = Z \sin \theta = 223.6 \sin -26.6° = -100\ \Omega\ .$$

4.2.2 Complex Number Algebra

Complex numbers can be manipulated in many ways to solve what would otherwise be a complicated phasor diagram, never forgetting that the real and imaginary terms represent differences in phase and therefore must be kept apart although eventually they may come together by use of Eqn. (4.1). From this it follows that the two sides of any complex number equation must be equal in both magnitude *and* phase.

4.2.2.1 Rationalization

This is a useful technique available for eliminating the imaginary component from the denominator of a fraction. It is based on the factorization of $a^2 - b^2$ into $(a + b)(a - b)$ from which, with complex numbers, $(a + jb)(a - jb) = a^2 - j^2 b^2 = a^2 + b^2$. Each of the two factors $(a + jb)$ and $(a - jb)$ is said to be the *complex conjugate* of the other, meaning simply that the sign of the imaginary component is reversed. In polar form therefore the complex conjugate of $r\angle\theta$ is $r\angle{-\theta}$. The following example demonstrates the technique.

EXAMPLE 4.2:

Rationalize $\dfrac{1}{4 - j3}$

$$\frac{1}{4 - j3} = \frac{1}{4 - j3} \times \frac{4 + j3}{4 + j3}$$

(numerator and denominator multiplied by complex conjugate)

$$= \frac{4 + j3}{4^2 + 3^2} = \frac{4 + j3}{25} = 0.16 + j0.12$$

If the denominator contains an imaginary term only, then multiply numerator and denominator by j, thus

$$\frac{a + \mathrm{j}b}{\mathrm{j}d} \quad \text{becomes} \quad \frac{-b + \mathrm{j}a}{-d}, \quad \text{i.e.} \quad \frac{b - \mathrm{j}a}{d}.$$

4.2.2.2 Addition and Subtraction

Standard mathematical techniques are used but always keeping real and imaginary parts separate, hence for two phasors $(a + \mathrm{j}b)$ and $(c + \mathrm{j}d)$:

$$(a + \mathrm{j}b) + (c + \mathrm{j}d) = (a + c) + \mathrm{j}(b + d) \tag{4.3}$$

with subtraction:

$$(a + \mathrm{j}b) - (c + \mathrm{j}d) = (a - c) + \mathrm{j}(b - d) \tag{4.4}$$

Parallel reactive circuits fill many of us with disquiet so the following example serves to show how useful the rules so far discussed can be.

EXAMPLE 4.3:

For the circuit of Figure 4.4(i) calculate the current I and its time relationship relative to the voltage.

The current I divides into I_1 and I_2, it is therefore the phasor sum of them.

$$\omega = 2\pi f = 2\pi \times 1592 = 10^4 \text{ rad/s}$$

$$X_L = \omega L = 10^4 \times 1.5 \times 10^{-3} = 15\ \Omega$$

$$\therefore \quad Z_1 = R_1 + \mathrm{j}\omega L = 8 + \mathrm{j}15\ \Omega$$

$$X_C = -\frac{1}{\omega C} = -\frac{10^6}{10^4 \times 20} = -5\ \Omega$$

$$\therefore \quad Z_2 = R_2 - \frac{\mathrm{j}}{\omega C} = 12 - \mathrm{j}5\ \Omega$$

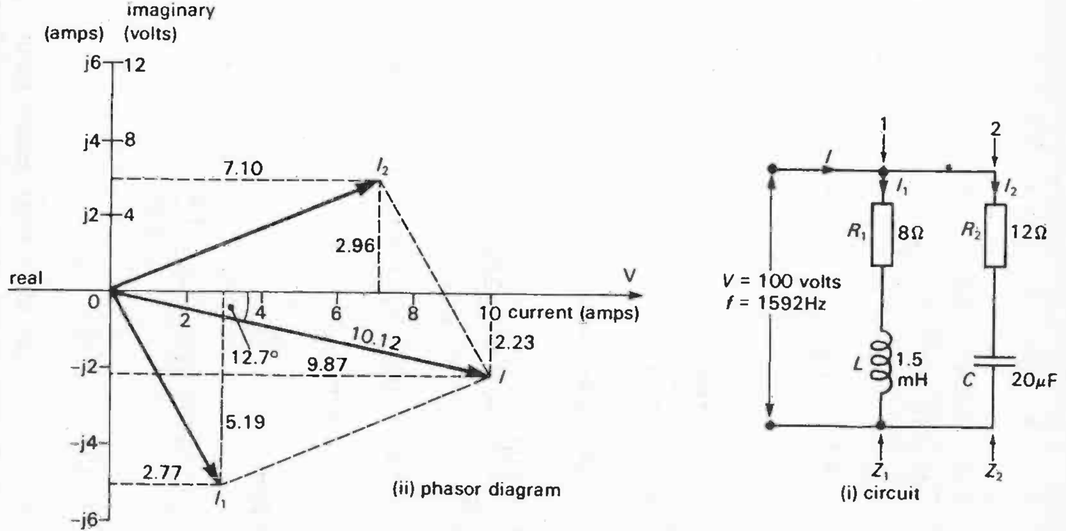


Fig. 4.4 Analysis of a parallel circuit

Hence,

$$I_1 = \frac{V}{Z_1} = \frac{100}{8 + j15} = \frac{100}{8 + j15} \times \frac{8 - j15}{8 - j15} \text{ (rationalizing)}$$

$$= \frac{800 - j1500}{8^2 + 15^2} = \frac{800 - j1500}{289}$$

$$= 2.77 - j5.19 \text{ A}$$

Similarly,

$$I_2 = \frac{V}{Z_2} = \frac{100}{12 - j5}$$

which on rationalizing becomes

$$\frac{1200 + j500}{169} = 7.10 + j2.96 \text{ A}$$

$$\therefore \quad I = I_1 + I_2 = (2.77 - j5.19) + (7.10 + j2.96)$$

$$= 9.87 - j2.23 \text{ A}$$

or in polar form

$$I = \sqrt{9.87^2 + 2.23^2} \tan^{-1} \frac{-2.23}{9.87}$$

$$= 10.12 \angle -12.7^\circ \text{ A} .$$

Figure 4.4(ii) illustrates what these results mean. The current I is obtained in the usual way by completing the parallelogram of sides I_1 and I_2 and drawing the diagonal.

Because I lags 12.7° on V and the wave period

$$T = \frac{1}{f} = \frac{1}{1592},$$

I therefore reaches any point in the cycle

$$\frac{1}{1592} \times \frac{12.7}{360} \times 10^6 \ \mu s,$$

i.e. just over 22 μs later than does the voltage.

Note that addition and subtraction of phasors can only be performed on cartesian coordinates, the polar form must therefore be converted first, hence the preference for j notation.

4.2.2.3 Multiplication and Division

Multiplying together two phasors using j notation gives:

$$(a + jb)(c + jd) = ac + jad + jbc + j^2 bd$$
$$= (ac - bd) + j(ad + bc) \qquad (4.5)$$

Dividing the same two phasors requires one stage of rationalization:

$$\frac{(a + jb)}{(c + jd)} = \frac{(a + jb)(c - jd)}{c^2 + d^2}$$

$$= \frac{(ac + bd) + j(bc - ad)}{c^2 + d^2} \qquad (4.6)$$

Multiplication and division are generally less complicated when using the polar form so because the requirement frequently arises in solving complex equations (as shown in the next example) it is perhaps worth looking at the proof of one of them. Consider Figure 4.5 and ignore the figures at present:

OP_1 and OP_2 are two phasors to be multiplied together. Join P_1 to the point 1 on the real axis (i.e. 1 + j0). Construct the triangle OP_2P to be *similar* (similar triangles have their corresponding angles equal) to $O,1,P_1$ (i.e. so that they have corresponding angles θ_1 and ϕ). Then

$$\frac{OP}{OP_1} = \frac{OP_2}{1}$$

hence

$$OP = OP_1 \times OP_2$$

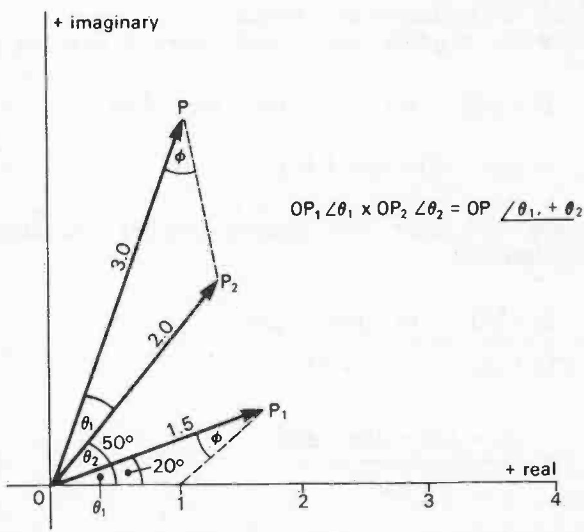


Fig. 4.5 Multiplication of two phasors

and as is evident from the Figure, the angle of the phasor P is equal to the sum of the angles of the phasors P_1 and P_2 $(\theta_1 + \theta_2)$ thus generally:

$$r_1\angle\theta_1 \times r_2\angle\theta_2 = r_1 r_2 \underline{/\theta_1 + \theta_2} \qquad (4.7)$$

the rule therefore being that the moduli are multiplied but the arguments are added.

The numbers given on Figure 4.5 refer to a phasor $1.5\angle 20^\circ$ multiplied by another $2\ \angle 50^\circ$ resulting in a single phasor $3\angle 70^\circ$.

The rule for division of phasors follows from Eqn. (4.7) although it too can be proved geometrically:

$$\frac{r_1\ \angle\theta_1}{r_2\ \angle\theta_2} = \frac{r_1}{r_2}\ \underline{/\theta_1 - \theta_2} \qquad (4.8)$$

The next example develops a different approach to that of Example 4.3.

EXAMPLE 4.4:
For the circuit of Figure 4.4(i) calculate the current I.

$$\omega = 10^4 \text{ rad/s}$$

$$X_L = 15\ \Omega$$

$$X_C = -5\ \Omega$$

$$\therefore \quad |Z_1| = \sqrt{R_1^2 + X_L^2} = \sqrt{8^2 + 15^2} = 17\ \Omega$$

$$\theta_1 = \tan^{-1}\frac{15}{8} = 61.9^\circ.$$

Then

$$I_1 = \frac{V}{Z_1} = \frac{100\angle 0^\circ}{17\angle 61.9^\circ} = 5.88\angle{-61.9^\circ}\text{ A}$$

$$|Z_2| = \sqrt{R_2^2 + X_C^2} = \sqrt{12^2 + 5^2} = 13\ \Omega$$

$$\theta_2 = \tan^{-1} \frac{5}{12} = -22.6^\circ .$$

Then

$$I_2 = \frac{V}{Z_2} = \frac{100 \angle 0^\circ}{13 \angle -22.6^\circ} = 7.69 \angle 22.6^\circ \text{ A}$$

As mentioned above, addition of I_1 and I_2 is carried out via cartesian coordinates, therefore real component of

$$I = I_1 \cos \theta_1 + I_2 \cos \theta_2 = (5.88 \cos - 61.9^\circ)$$

$$+ (7.69 \cos 22.6^\circ) = 2.77 + 7.10 = 9.87 \text{ A}$$

imaginary component of

$$I = I_1 \sin \theta_1 + I_2 \sin \theta_2 = (5.88 \sin - 61.9^\circ)$$

$$+ (7.69 \sin 22.6^\circ) = -5.19 + 2.96 = -2.23 \text{ A}$$

i.e. I is expressed by 9.87 – j2.23 A, or in polar coordinates by

$$\sqrt{9.87^2 + (-2.23)^2} \tan^{-1} \frac{-2.23}{9.87} = 10.12 \angle -12.7^\circ \text{ A}$$

4.2.2.4 Admittance Diagrams

Admittance (Y) is the reciprocal of impedance (Z) and the unit is the siemen (S), hence:

$$Y = \frac{1}{Z} \text{ S}$$

Because Z may be complex, Y may be also and its real and imaginary parts are G, the *conductance* and B, the *susceptance*, hence

$$Y = G + jB = |Y| \angle \phi \tag{4.9}$$

the unit again being the siemen. The susceptance, B is the component of admittance arising from inductance or capacitance. From Eqn. (4.9) and since $Z = R + jX$ then

$$G + jB = \frac{1}{R + jX} = \frac{R - jX}{R^2 + X^2}$$

$$= \frac{R}{R^2 + X^2} - \frac{jX}{R^2 + X^2}$$

and equating real and imaginary parts:

$$G = \frac{R}{R^2 + X^2} \tag{4.10}$$

$$B = \frac{-X}{R^2 + X^2} \tag{4.11}$$

and working the other way round, since

$$R + jX = \frac{1}{G + jB} = \frac{G - jB}{G^2 + B^2},$$

then

$$R = \frac{G}{G^2 + B^2} \tag{4.12}$$

$$X = \frac{-B}{G^2 + B^2} \tag{4.13}$$

thus via Eqns. (4.10) to (4.13) we can convert from the impedance to the admittance form and vice versa.

EXAMPLE 4.5:
For the series circuit of Figure 4.6(i), draw the admittance diagram.

In Figure 4.6(ii) is shown the impedance diagram from which

$$Z = \sqrt{4^2 + 3^2} \tan^{-1} \frac{3}{4} = 5\ \Omega \text{ at an angle of } 36.9° .$$

The equivalent parallel circuit is in (iii) of the Figure and from Eqns. (4.10) and (4.11)

$$G = \frac{R}{R^2 + X^2} = \frac{R}{Z^2} = \frac{4}{25} = 0.16\ \text{S}$$

$$B = \frac{-X}{R^2 + X^2} = \frac{-X}{Z^2} = \frac{-3}{25} = -0.12\ \text{S}$$

The common factor in this circuit is the voltage and for the same current I, of 1 ampere the voltage developed across the equivalent parallel circuit must be the same. The admittance diagram follows in (iv) of the Figure and the scales have been chosen so as to illustrate the similarities of the two diagrams, especially in that the angle of lead of the impedance with regard to our particular reference is identical with the angle of lag of the admittance for

$$\tan^{-1} \frac{-0.6}{0.8} = -36.9° .$$

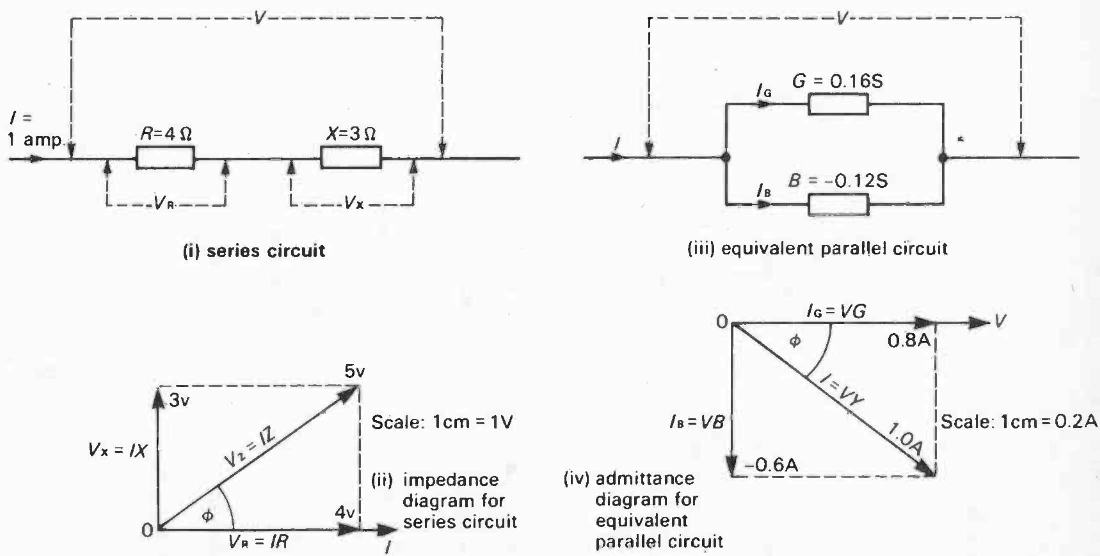


Fig. 4.6 Admittance diagram for a series circuit

The next example also gives experience in switching between impedance and admittance formulae:

EXAMPLE 4.6:
Calculate the current and its phase angle relative to the supply voltage for the circuit of Figure 4.7.

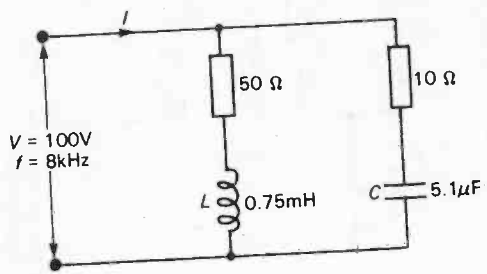


Fig. 4.7 A parallel reactive circuit

$$X_L = 2\pi fL = 2\pi \times 8000 \times 0.75 \times 10^{-3} = 37.7\ \Omega$$

$$X_C = \frac{1}{2\pi fC} = \frac{10^6}{2\pi \times 8000 \times 5.1} = 3.9\ \Omega$$

Inductive arm:–

$$\text{impedance} = 50 + j37.7\ \Omega$$

$$\text{admittance} = \frac{1}{50 + j37.7} = \frac{50 - j37.7}{50^2 + 37.7^2}$$

$$= 0.0128 - j0.0096\ \text{S}$$

Capacitive arm:–

$$\text{impedance} = 10 - j3.9$$

$$\text{admittance} = \frac{1}{10 - j3.9} = \frac{10 + j3.9}{10^2 + 3.9^2}$$

$$= 0.0868 + j0.0339 \text{ S}$$

Total admittance,

$$Y = (0.0868 + 0.0128) + j(0.0339 - 0.0096)$$

$$= 0.0996 + j0.0243 \text{ S}$$

Current,

$$I = VY = 100(0.0996 + j0.0243) = 9.96 + j2.43 \text{ A}$$

$$= \sqrt{9.96^2 + 2.43^2} \ \tan^{-1} \frac{2.43}{9.96} = 10.25 \angle 13.7^\circ \, .$$

This example demonstrates that although we may be reluctant to change over to the admittance concept, it does have advantages where parallel circuits are concerned, especially for those with several parallel paths.

Many more examples of the use of complex algebra are given in subsequent chapters, e.g. Sections 5.5.1, 5.9.1, 12.3.1, 12.3.2.

4.3 SUMMARY OF KEY FORMULAE

a, b, c, d = lengths of cartesian coordinates

r = length of polar coordinate

θ, ϕ = degrees

R = resistance

X = reactance

Z = impedance

OPERATION	FORMULA	UNIT	SECTION
Change from cartesian to polar co-ordinates	$r\angle\theta = \sqrt{a^2 + b^2} \tan^{-1}\frac{b}{a}$		4.2.1
Change from polar to cartesian co-ordinates	$a = r\cos\theta, \quad b = r\sin\theta$		4.2.1
Addition of cartesian co-ordinates	$(a+jb) + (c+jd) = (a+c) + j(b+d)$		4.2.2.2
Subtraction of cartesian co-ordinates	$(a+jb) - (c+jd) = (a-c) + j(b-d)$		4.2.2.2
Multiplication of cartesian co-ordinates	$(a+jb)(c+jd) = (ac-bd) + j(ad+bc)$		4.2.2.3
Division of cartesian co-ordinates	$\frac{(a+jb)}{(c+jd)} = \frac{(ac+bd) + j(bc-ad)}{c^2 + d^2}$		4.2.2.3
Multiplication of polar co-ordinates	$r_1\angle\theta_1 \times r_2\angle\theta_2 = r_1 r_2 \angle\theta_1 + \theta_2$		4.2.2.3
Division of polar co-ordinates	$\frac{r_1\angle\theta_1}{r_2\angle\theta_2} = \frac{r_1}{r_2} \angle\theta_1 - \theta_2$		4.2.2.3

OPERATION	FORMULA	UNIT	SECTION
Admittance	$Y = \frac{1}{Z} = \lvert Y \rvert \angle\phi = G + jB$	S	4.2.2.4
Conductance	$G = \frac{R}{R^2 + X^2}$	S	4.2.2.4
Susceptance	$B = \frac{-X}{R^2 + X^2}$	S	4.2.2.4
Resistance	$R = \frac{G}{G^2 + B^2}$	Ω	4.2.2.4
Reactance	$X = \frac{-B}{G^2 + B^2}$	Ω	4.2.2.4

5. AMPLIFIERS

Most amplifiers, whether of discrete components or I.C.'s are based on the transistor. Accordingly this Chapter commences with a resumé of the fundamentals of semiconductor action.

Semiconductor activity is based on the movement of *charge carriers* within a semiconducting material, typically silicon or germanium. Valency shells and valency electrons are discussed in Section 1.3.1.2 and the atoms of both of these elements have four electrons in the outer (valence) shell – they are *tetravalent*. Given energy from heat some valence electrons break free so the conductivity increases with temperature. The minimum energy required to free a valence electron is 1.1 eV for silicon but only 0.7 eV for germanium (see also Sect.1.3.1.2). In the pure crystalline semiconductor material the atoms are locked in a lattice with the valence electrons of adjacent atoms combining to form *covalent bonds*. Each atom has four such bonds with its four neighbours and so in a way has 8 valence electrons in orbit but with each electron serving two atoms. Because of this tight bond arrangement the pure material has poor conductivity.

Conductivity is improved by *doping* which is the addition to the material of a tiny amount of a trivalent (3 valency electrons) or pentavalent (5) element. The atoms of these elements are capable of fitting themselves into the lattice structure by forming covalent bonds with the existing atoms. For a pentavalent atom therefore, since only four valence electrons are required for bonding, one is thrown spare so the conductivity is increased. With a trivalent atom there is one electron too few and a *hole* is said to exist because one of the covalent bonds is incomplete. A hole is considered to be equivalent to a positively charged particle, the magnitude of the charge being equal to that of an electron (Sect.1.3.1) and in semiconductor theory it is regarded as a positive current carrier.

Pentavalent elements which provide free electrons within the crystal are known as donors and silicon or germanium so doped is called n-type. Phosphorus, arsenic and antimony are examples of elements which are used (see Table 1.1 in

Sect.1.3). On the other hand trivalent elements such as boron, gallium and indium are known as acceptors because they take electrons from the crystal which is then known as p-type. When n-type and p-type materials are formed in intimate contact a *p-n junction* is formed and this is the centre of activity in semiconductor diodes and transistors. Because there is an excess of electrons in the n-type and holes in the p-type, they attract each other across the junction and a certain amount of recombination takes place. Accordingly the n-type takes on a slightly positive charge with the p-type negative, as shown pictorially in Figure 5.1(i). There is therefore a potential gradient across the junction which is represented in the Figure by a battery, the magnitude of the potential becoming such as to inhibit further diffusion. This *potential barrier* gives the p-n junction its special characteristics for if an external battery is connected across the junction aiding the barrier potential (reverse bias), no current flows but if the external battery opposes the barrier potential (forward bias), current flows, its magnitude being controlled by the voltage applied. Two p-n junctions close together within a p-type or n-type semiconducting material creates a transistor, either p-n-p or n-p-n as shown in Figure 5.1(ii). The three layers are labelled *collector, base* and *emitter* as also shown.

A reminder about arrow direction on semiconductor symbols may be appropriate. The arrows point in the direction of "hole current" flow which is not the same as "conventional current" as discussed in Section 3.5. In fact when an external battery is connected to a p-n junction in the forward direction it can be considered as delivering electrons from its negative terminal to fill the arriving holes while receiving electrons at its positive terminal, this being exactly as we see electron flow from – to + round a circuit.

In operation the emitter injects charges into the emitter-base junction which is usually forward biassed and although the collector-base junction is reverse-biassed, the charges are swept through the very thin base into the collector. There is therefore a through current but what is of major significance is that it is controlled by the potential on the base. The results of this action are made visible through the many characteristic

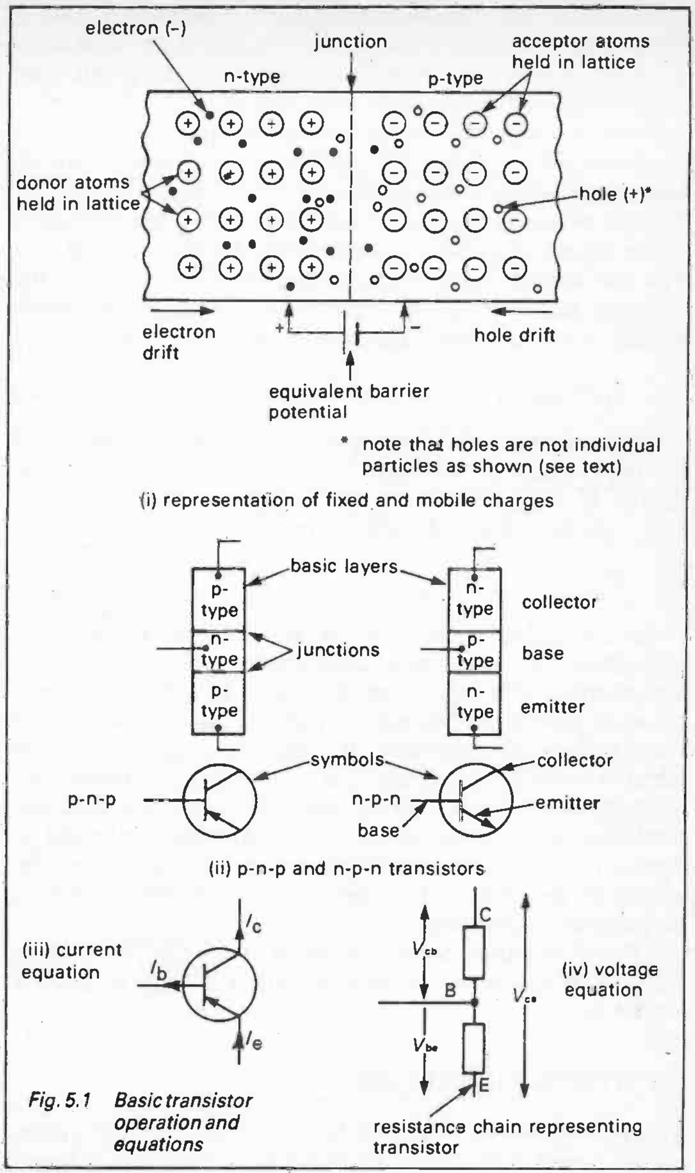


Fig. 5.1 Basic transistor operation and equations

curves published by manufacturers. Alternatively simple voltage/current measurements can be made in the workshop. The manipulation of such data enables us to predict the performance of any transistor in any circuit.

Perhaps the most likely need of calculation arises with the straightforward 3-terminal junction transistor shown schematically and with symbols in Figure 5.1(ii). The simplest formula of all, applicable to both p-n-p and n-p-n transistors is the *transistor current equation*, developed from the fact that the emitter current divides between the base and the collector as shown typically for a p-n-p transistor in Figure 5.1(iii). Ignoring current directions therefore:

$$I_e = I_b + I_c \tag{5.1}$$

Also irrespective of the way in which it is connected or of the type, under steady conditions the transistor can be considered as a tapped resistance as in Figure 5.1(iv). V_{cb} and V_{be} are in series and their sum is equal to V_{ce}, hence:

$$V_{ce} = V_{cb} + V_{be} \tag{5.2}$$

Several analysis techniques are available, those used most extensively in small-signal applications are the T and π equivalent circuits and a third, known as the *hybrid parameter network*. The latter does not consider the physical make-up of the transistor as a network of resistances, capacitances and generators but instead treats it as a *black-box* and determines what happens at one pair of terminals when the electrical conditions are changed at the other. Because the hybrid or *h-parameters* are the most easily measured and are usually quoted in published data, these are the ones developed here for practical calculations.

[Tuned amplifiers are not included in this Chapter because the general theory of coupled circuits is already in Book 1 (Sect.4.7.]

5.1 HYBRID PARAMETERS

Always denoted by the small letter h for "hybrid", these are so named because the various parameters have different dimensions as will become evident later. A reminder of the black-box approach is perhaps useful first.

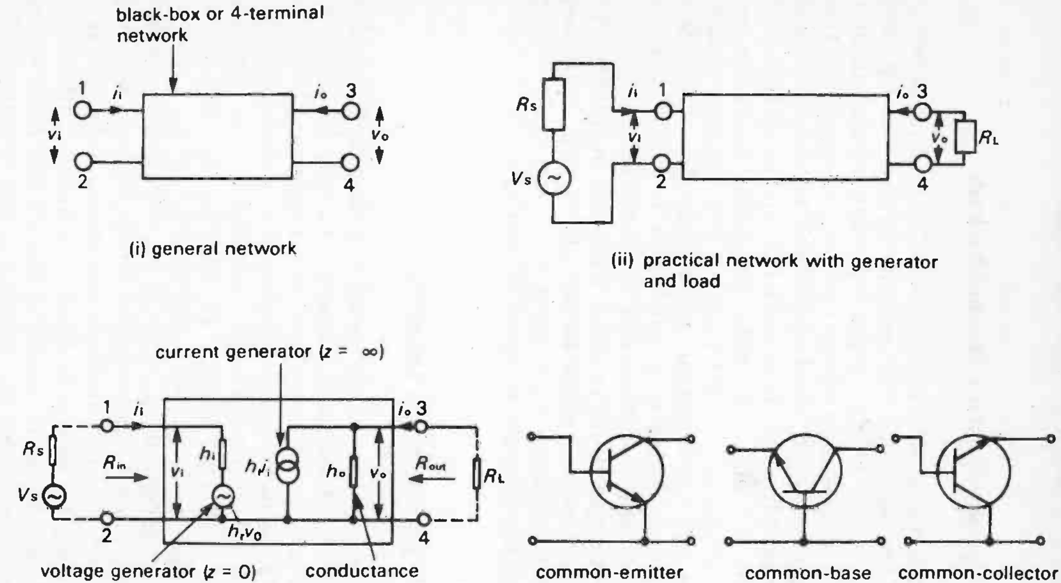
5.1.1 *h*-parameters of a 4-Terminal Network

The black-box as shown in Figure 5.2(i) is so called because we are less concerned with what it contains than with what it does as can be determined by simple measurements of voltage and current on its input and output terminals. These are considered to be "small signals" and so whatever is in the box operates linearly. In practice the currents i_i and i_0 are unable to flow unless the input and output circuits are completed, accordingly Figure 5.2(ii) shows the box operating with an input drive and an output load. The a.c. measurements are conducted at a certain test frequency and from these the network performance is determined, for example, the input resistance (at terminals 1 and 2) is v_i/i_i, the output resistance is v_0/i_0 and the current gain is i_0/i_i. The lower case letters indicate small signal a.c. conditions. Voltage gains and conductances are measured similarly.

5.1.2 *h*- parameters of a Transistor

Certain parameters set out below are particularly suited to the analysis of transistor circuits. Each is calculated for specific terminating conditions and a subscript to the h indicates the particular parameter with a second subscript if required for common-emitter, common-base or common-collector [see Fig.5.2(iv)]. As an example, h_{fe} is the current gain (h_f) in common-emitter configuration.

For a transistor the contents of the black-box are revealed in Figure 5.2(iii). The equivalent circuit shown is effectively in two parts (i) an input circuit containing a voltage generator representing the voltage feedback of a transistor (the output current also flows in the input circuit) and (ii) a current generator representing the transistor action operating with a parallel output resistance which for convenience in these calculations is expressed as a conductance.



(i) general network

(ii) practical network with generator and load

(iii) equivalent circuit of transistor

(iv) the 3 basic transistor configurations (n-p-n transistor)

Fig. 5.2 4-terminal networks

Accordingly the four h-parameters used to describe the small-signal performance of a transistor are:

(i) input impedance, h_i , measured with terminals 3 and 4 [Fig.5.2(iii)] short-circuited to a.c. so that $v_0 = 0$:

$$h_i = \frac{v_i}{i_i} \ \Omega \tag{5.3}$$

(ii) current gain or forward current transfer ratio, h_f , measured as above:

$$h_f = \frac{i_0}{i_i} \text{ (no dimension, simply a ratio)} \tag{5.4}$$

(iii) output conductance, h_0 , measured with terminals 1 and 2 open-circuited so that $i_i = 0$. The output resistance $= v_0/i_0$, therefore

$$R_0 = \frac{i_0}{v_0} \ \text{S} \tag{5.5}$$

(iv) voltage feedback ratio or reverse voltage transfer ratio, h_r , measured as above:

$$h_r = \frac{v_i}{v_0} \text{ (no dimension)} \tag{5.6}$$

Manufacturers' data frequently includes curves of each of these at various values of collector current.

5.1.2.1 Evaluation

Although h-parameters may be derived directly from a.c. measurements, it is also possible to calculate them from transistor static (d.c.) characteristics. A typical circuit designed for a range of measurements on, for example, an n-p-n transistor in common-base is shown in Figure 5.3, in this all voltages and currents are indicated at once. Supply voltages are adjusted by R_1 and R_2 .

EXAMPLE 5.1:
The following measurements were made on a transistor connected in common-emitter at a constant collector voltage of 1 V.

V_{be} (V)	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
I_b (mA)	0.03	0.10	0.22	0.38	0.55	0.80	1.08	1.40	1.70	2.06

Determine the input resistance, h_{ie} over the straight portion of the characteristic.

The input characteristic is drawn in Figure 5.4(i) from Eqn. (5.3)

$$h_i \text{ for the general circuit } = \frac{v_i}{i_i},$$

so in this practical case,

$$h_{ie} = \frac{\delta V_{be}}{\delta I_b}$$

(where δ means "a small change in"), i.e.

$$h_{ie} = \frac{(1.1 - 0.9)\text{ V}}{(1.4 - 0.8) \times 10^{-3}\text{ A}}$$

$$= \frac{0.2}{0.6 \times 10^{-3}} = 333\ \Omega$$

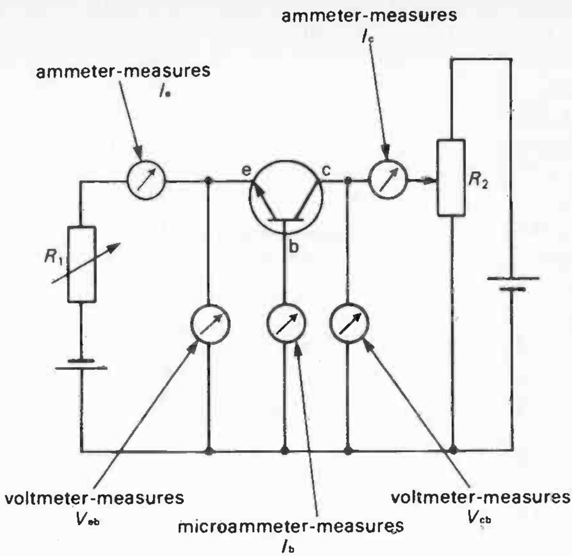


Fig. 5.3 Static characteristic measuring circuit

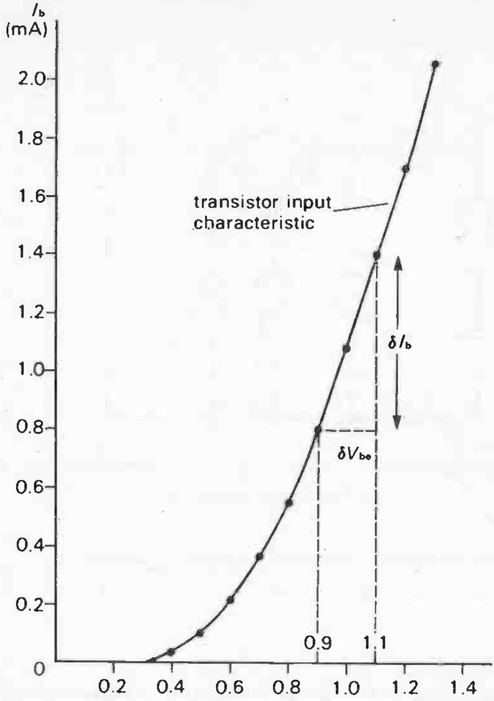
EXAMPLE 5.2:
From a manufacturer's data the output characteristics of an n-p-n transistor in common-emitter are as presented in Figure 5.4(ii). From these calculate h_{fe} at $V_{ce} = 1$ V and h_{oe} at $I_b = 10\,\mu A$.

From Eqn. (5.4)

$$h_f \text{ for the general circuit } = \frac{i_0}{i_i},$$

therefore in this practical case, at $V_{ce} = 1.0$ V :

$$h_{fe} = \frac{\delta I_c}{\delta I_b} = \frac{(3.3 - 2.45)\,\text{mA}}{(10 - 7.5)\,\mu\text{A}}$$



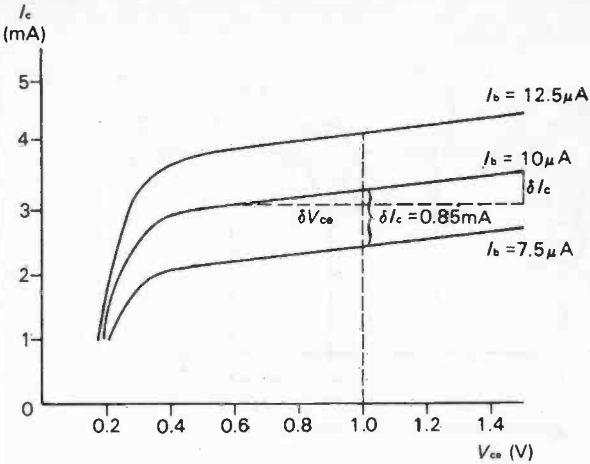
(i) input resistance h_{ie}

Fig.5.4 Measurement of h-parameters

$$= \frac{0.85 \text{ mA}}{2.5 \times 10^{-3} \text{ mA}} = 340$$

From Eqn. (5.5)

$$h_0 \text{ for the general circuit } = \frac{i_0}{v_0} \text{, therefore:}$$



(ii) current gain, h_{fe} and output conductance, h_{re}

Fig. 5.4 (contd)

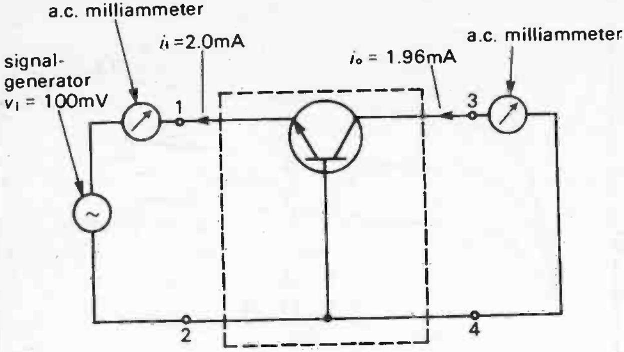
$$\therefore \quad h_{oe} = \frac{\delta\ I_c}{\delta\ V_{ce}}.$$

The graph at $I_b = 10\ \mu A$ is straight above $V_{ce} = 0.6$ V

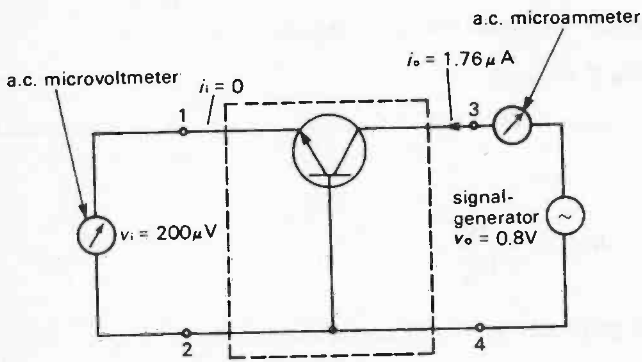
$$\therefore \quad h_{oe} = \frac{(3.7 - 3.1)\ \text{mA}}{(1.6 - 0.6)\ \text{V}} = \frac{0.6 \times 10^{-3}\ \text{A}}{1\ \text{V}} = 600\ \mu\text{S}$$

(microsiemens – on some data sheets expressed as μA/V).

The above two examples show how the *h*-parameters are derived from transistor static characteristics. The alternative method of direct measurement requires moderately simple test gear using pure-tone signals. Taking Figure 5.4(i) as an



(a) measurements for h_{ib} and h_{fb}



(b) measurements for h_{ob} and h_{rb}

(iii) measurement of h- parameters using signal-generator and test meters

Fig. 5.4 (contd)

example, an a.c. signal of δV_{be} is applied between base and emitter of the transistor resulting in an input current of δI_b , the d.c. bias being set so that the transistor operates linearly (see Sect. 5.4). The signal frequency is chosen according to the projected use of the transistor, e.g. for low-frequency applications, 1000 Hz. The following example illustrates this and the basic practical arrangements required.

EXAMPLE 5.3:
In Figure 5.4(iii) a transistor connected in common-base (biassing and supply arrangements are omitted) has a 1000 Hz signal generator connected as shown. Calculate the *h*-parameters from the measurements (r.m.s.) given.

From Eqns. (5.3) to (5.6):

(i) input resistance,

$$h_{ib} = \frac{v_i}{i_i} (v_o = 0) = \frac{100 \times 10^{-3}}{2 \times 10^{-3}} = 50\ \Omega$$

(ii) current gain,

$$h_{fb} = \frac{i_o}{i_i} (v_o = 0) = \frac{1.96 \times 10^{-3}}{2.0 \times 10^{-3}} = 0.98$$

(iii) output conductance,

$$h_{ob} = \frac{i_o}{v_o} (i_i = 0) = \frac{1.76 \times 10^{-6}}{0.8} = 2.2\ \mu S$$

(iv) voltage feedback ratio,

$$h_{rb} = \frac{v_i}{v_o} (i_i = 0) = \frac{200 \times 10^{-6}}{0.8} = 2.5 \times 10^{-4}$$

the simplicity of measurement and of subsequent calculation is well demonstrated.

5.1.3 The Amplifier Equations

Given the h-parameters of a bipolar transistor, the amplifier design process can begin. Here amplifiers are considered generally but excluding those for very high frequencies for which internal capacitances cannot be ignored. In the basic design of an amplifier the performance and operating conditions are specified mainly by the input and output resistances (R_{in} and R_{out}) and the gain required. The gain can be quoted in terms of current (K_i), voltage (K_v) or power (K_p). The appropriate general circuit is that of Figure 5.2(iii).

With four h-parameters and three methods of connecting a transistor, many formulae are needed to cover the whole process hence a few formulae are developed for experience with the remainder merely quoted. Transistor stage design is unlikely to be exacting, nor can it be considering the variability of the transistors themselves, hence the designer usually relies on negative feedback to stabilize gain (Sect. 5.9). Thus we will find that, given certain conditions for the values of the source and load resistances, R_S and R_L, some simplification is possible. Also because many manufacturers quote h-parameters for common-emitter only, relying on the fact that there are conversion formulae for the other configurations, we will conduct our proofs entirely with common-emitter, but eventually leading to a full range of useful and practical formulae.

5.1.3.1 Current Gain

From Figure 5.2(iii), i_0 is the sum of the currents in the generator and the transistor output conductance, h_o, therefore:

$$i_0 = h_{fe}i_i + h_{oe}v_0$$

the voltage across R_L opposes v_0 hence $v_0 = -i_0R_L$

$$\therefore \quad i_0 = h_{fe}i_i - h_{oe}i_0R_L$$

$$\therefore \quad i_0(1 + h_{oe}R_L) = h_{fe}i_i$$

hence Current Gain,

$$K_i = \frac{i_0}{i_i} = \frac{h_{fe}}{1 + h_{oe}R_L} \tag{5.7}$$

This is the exact equation. However a useful approximation based on the fact that in practice $h_{oe}R_L$ is likely to be small (e.g. from Example 5.5 which follows, $h_{oe}R_L = 0.054$) leads to:

$$K_i = h_{fe} \text{ (approx.)} \tag{5.8}$$

5.1.3.2 Input Resistance

The voltage looking into terminals 1 and 2 of Figure 5.2(iii) is the sum of that across h_i plus that of the voltage generator, $h_r v_0$ hence for common-emitter:

$$v_i = h_{ie}i_i + h_{re}v_0$$

also as shown in the preceding section,

$$v_0 = -i_0 R_L \quad \text{and} \quad i_0 = K_i i_i$$

$$\therefore \quad v_0 = -i_i K_i R_L$$

hence

$$v_i = h_{ie}i_i + h_{re}(-i_i K_i R_L) = i_i(h_{ie} - h_{re}K_i R_L)$$

$$\therefore \quad R_{in} = \frac{v_i}{i_i} = h_{ie} - h_{re}K_i R_L \tag{5.9}$$

and substituting for K_i

$$R_{in} = h_{ie} - \frac{h_{fe}h_{re}R_L}{1 + h_{oe}R_L} \tag{5.10}$$

This is the exact formula showing that as a working amplifier the stage input resistance is lower than the input resistance of the transistor itself (h_{ie}). However, the degree of reduction is dependent on h_{re} which for most transistors is very small (10^{-3} to 10^{-4}) so as in the preceding section, if $h_{oe}R_L$ is also small, the second term can be neglected for approximate calculations, hence

$$R_{in} = h_{ie} \text{ (approx.)} \tag{5.11}$$

EXAMPLE 5.4:
A transistor is to be operated as a common-emitter amplifier with a 3.3 kΩ load. Its h-parameters are quoted at the chosen operating conditions as $h_{ie} = 4.5$ kΩ, $h_{fe} = 300$, $h_{oe} = 20\ \mu$S, $h_{re} = 2 \times 10^{-4}$. What are the current gain and input resistance?

From Eqn. (5.7), Current gain,

$$K_i = \frac{h_{fe}}{1 + h_{oe}R_L} = \frac{300}{1 + (20 \times 10^{-6} \times 3300)}$$

$$= \frac{300}{1.066} = 281$$

Using the approximate formula [Eqn.(5.8)], the answer is 300, 7% high.

From Eqn. (5.10), Input resistance,

$$R_{in} = h_{ie} - \frac{h_{fe}h_{re}R_L}{1 + h_{oe}R_L}$$

$$= 4500 - \frac{300 \times 2 \times 10^{-4} \times 3300}{1.066}$$

$$= 4500 - 185.7 = 4314\ \Omega$$

If using the approximate formula [Eqn.(5.11)], we simply ignore the fraction, giving an answer of 4500, some 4.3% high.

5.1.3.3 *Output Resistance*

In the preceding Section it is shown that for the input circuit of Figure 5.2(iii):

$$v_i = h_{ie}i_i + h_{re}v_o$$

But i_i also flows through R_S and so generates an equal and opposite voltage, hence:

$$-i_iR_S = h_{ie}i_i + h_{re}v_o$$

$$\therefore \quad i_i = -\frac{h_{re}v_o}{h_{ie} + R_S}$$

From Section 5.1.3.1,

$$i_o = h_{fe}i_i + h_{oe}v_o$$

$$\therefore \quad i_o = \frac{-h_{fe}h_{re}v_o}{h_{ie} + R_S} + h_{oe}v_o$$

$$= v_o\left(h_{oe} - \frac{h_{fe}h_{re}}{h_{ie} + R_S}\right)$$

$$\therefore \quad i_o = v_o\left(\frac{h_{oe}(h_{ie} + R_S) - h_{fe}h_{re}}{h_{ie} + R_S}\right)$$

$\therefore$ Output resistance,

$$R_{out} = \frac{v_o}{i_o} = \frac{h_{ie} + R_S}{h_{oe}(h_{ie} + R_S) - h_{fe}h_{re}} \qquad (5.12)$$

hence the output resistance depends to a certain extent on the source resistance. If, as is frequently the case, R_S is very much greater than h_{ie} and again h_{re} is very small so that $h_{fe}h_{re}$ can be neglected:

$$R_{out} = \frac{R_S}{h_{oe}R_S} = \frac{1}{h_{oe}} = \text{(approx.)} \qquad (5.13)$$

EXAMPLE 5.5:
A transistor connected in common-emitter with a collector load of 2.7 kΩ has the following *h*-parameters, $h_{ie} = 2.5$ kΩ, $h_{fe} = 200$, $h_{oe} = 20\ \mu S$, $h_{re} = 1.5 \times 10^{-4}$. Its input is connected to a 600 Ω line. What are the input and output resistances?

From Eqn. (5.10):

$$R_{in} = h_{ie} - \frac{h_{fe}h_{re}R_L}{1 + h_{oe}R_L}$$

$$= 2500 - \frac{200 \times 1.5 \times 10^{-4} \times 2700}{1 + (20 \times 10^{-6} \times 2700)}$$

$$= 2500 - 76.85 = 2423\ \Omega$$

From Eqn. (5.12):

$$R_{out} = \frac{h_{ie} + R_S}{h_{oe}(h_{ie} + R_S) - h_{fe}h_{re}}$$

$$= \frac{2500 + 600}{(20 \times 10^{-6})(2500 + 600) - 200 \times 1.5 \times 10^{-4}}$$

$$= \frac{3100}{0.032} = 96.88 \text{ k}\Omega \, .$$

Combining the output resistance of the transistor with its load (they are effectively in parallel) gives a net value of 2.63 kΩ. Its own input resistance is about 2.4 kΩ, hence two or more similar stages are well matched, a useful feature of the common-emitter amplifier.

5.1.3.4 Voltage and Power Gains

In Figure 5.2(iii), $v_i = i_i R_{in}$, $v_0 = -i_0 R_L$

∴ Voltage gain,

$$K_v = \frac{v_0}{v_i} = \frac{-i_0 R_L}{i_i R_{in}} = \frac{-K_i R_L}{R_{in}} \tag{5.14}$$

or, by substituting for K_i and R_{in} , (Eqns. 5.7 and 5.10):

$$K_v = -\frac{h_{fe} R_L}{h_{ie}(1 + h_{oe} R_L) - h_{fe} h_{re} R_L} \tag{5.15}$$

the minus sign reminding us that through a common-emitter stage there is a 180° phase change.

By using the argument that h_{oe} and h_{re} are usually very small:

$$K_v = -\frac{h_{fe} R_L}{h_{ie}} \text{ (approx.)} \tag{5.16}$$

The power gain, K_p follows for it can be shown to be the product of K_i and K_v , so, substituting for K_v [Eqn. (5.14)] and ignoring the minus sign because such phase differences are of no concern in power calculations:

$$K_p = K_i \times \frac{K_i R_L}{R_{in}} = \frac{K_i^2 R_L}{R_{in}} \qquad (5.17)$$

Now a little practice with decibels never comes amiss but first some revision:

No. of decibels $= 10 \log \frac{P_2}{P_1}$ and if

$$P_1 = I_1^2 R_1 \text{ and } P_2 = I_2^2 R_2 ,$$

$$\text{No. of dB} = 10 \log \frac{I_2^2 R_2}{I_1^2 R_1} = 10 \log \left(\frac{I_2}{I_1}\right)^2 \times \frac{R_2}{R_1}$$

$$= 20 \log \frac{I_2}{I_1} + 10 \log \frac{R_2}{R_1}$$

Similarly no. of dB

$$= 10 \log \frac{E_2^2}{E_1^2} \times \frac{R_1}{R_2} = 20 \log \frac{E_2}{E_1} + 10 \log \frac{R_1}{R_2}$$

EXAMPLE 5.6:
A single-stage transistor amplifier with a collector load of 4.7 kΩ has the following h-parameters, $h_{ie} = 1.5$ kΩ, $h_{fe} = 100$, $h_{oe} = 40\ \mu S$, $h_{re} = 5 \times 10^{-4}$. Calculate the gain in decibels via the current, voltage and power gains.

From Eqn. (5.7)

$$K_i = \frac{h_{fe}}{1 + h_{oe} R_L} = \frac{100}{1.188} = 84.175$$

From Eqn. (5.10)

$$R_{in} = h_{ie} - \frac{h_{fe}h_{re}R_L}{1 + h_{oe}R_L} = 1500 - \frac{235}{1.188}$$

$$= 1500 - 197.8 = 1302.2\ \Omega$$

From Eqn. (5.14)

$$K_v = \frac{K_i R_L}{R_{in}} = \frac{84.175 \times 4700}{1302.2} = 303.81$$

From Eqn. (5.17)

$$K_p = K_i^2 \frac{R_L}{R_{in}} = 84.175^2 \times \frac{4700}{1302.2} = 25{,}573$$

Then:

(dB) gain calculated via

$$K_i = 20 \log 84.175 + 10 \log \frac{4700}{1302} = 38.50 + 5.575$$

$$= 44.08 \text{ dB} ;$$

gain calculated via

$$K_v = 20 \log 303.81 + 10 \log \frac{1302}{4700} = 49.65 - 5.575$$

$$= 44.08 \text{ dB} ;$$

gain calculated via

$$K_p = 10 \log 25{,}573 = 44.08 \text{ dB} .$$

5.1.4 *h*-parameter Conversions

The common-base and common-collector exact amplifier equations follow the pattern of those developed above for common-emitter, the only difference being that the subscript e is dropped and is replaced by b or c as appropriate. For example:

$$K_i \text{ for common-emitter} = \frac{h_{fe}}{1 + h_{oe}R_L} \quad \text{[Eqn.(5.7)]}$$

$$K_i \text{ for common-base} = \frac{h_{fb}}{1 + h_{ob}R_L}$$

$$K_i \text{ for common-collector} = \frac{h_{fc}}{1 + h_{oc}R_L}$$

and K_v , K_p , R_{in} and R_{out} can be stated similarly. The approximations do not follow the same pattern however so are given later. A difficulty already mentioned is that we may wish to calculate, say, K_i for a common-base circuit but, as is quite likely, have only *h*-parameters for common-emitter available. We get over this problem by using a set of conversion equations from common-emitter to the other two as follows:

To common-base:

$$h_{ib} = \frac{h_{ie}}{1 + h_{fe}} \quad (5.18)$$

$$h_{fb} = \frac{-h_{fe}}{1 + h_{fe}} \quad (5.19)$$

$$h_{ob} = \frac{h_{oe}}{1 + h_{fe}} \quad (5.20)$$

$$h_{rb} = \frac{h_{ie}h_{oe}}{1 + h_{fe}} \tag{5.21}$$

To common-collector:

$$h_{ic} = h_{ie} \tag{5.22}$$

$$h_{fc} = -(1 + h_{fe}) \tag{5.23}$$

$$h_{oc} = h_{oe} \tag{5.24}$$

$$h_{rc} = \frac{1}{1 + h_{re}} \tag{5.25}$$

We are now in a position to calculate amplifier performance for a transistor in any mode and the following example is chosen as a reminder of the main features. For these we use the h-parameters of a practical transistor measured at I_c = 2.0 mA, V_{ce} = 5 V, they are h_{ie} = 4.5 kΩ, h_{fe} = 330, h_{oe} = 30 μS, h_{re} = 2.0×10^{-4}.

EXAMPLE 5.7:
For the transistor above, calculate K_i, K_v, R_{in}, R_{out} for the common-emitter (c-e) and common-base (c-b) configurations. The source resistance in both cases is 1000 Ω and load resistance, c-e = 3.3 kΩ, c-b = 200 kΩ.

Convert the h-parameters first:

$$h_{ib} = \frac{h_{ie}}{1 + h_{fe}} = \frac{4500}{331} = 13.6\ \Omega \qquad \text{[Eqn.(5.18)]}$$

$$h_{fb} = \frac{-h_{fe}}{1 + h_{fe}} = \frac{-330}{331} = -0.997 \qquad \text{[Eqn.(5.19)]}$$

$$h_{ob} = \frac{h_{oe}}{1 + h_{fe}} = \frac{30 \times 10^{-6}}{331} = 0.091\ \mu S \text{[Eqn.(5.20)]}$$

$$h_{rb} = \frac{h_{ie}h_{oe}}{1 + h_{fe}} = \frac{4500 \times 30 \times 10^{-6}}{331} \qquad \text{[Eqn.(5.21)]}$$

$$= 4.08 \times 10^{-4}$$

Then from Eqn. (5.7)

$$K_i \text{ (c-e)} = \frac{h_{fe}}{1 + h_{oe}R_L} = \frac{330}{1 + (30 \times 10^{-6} \times 3300)}$$

$$= 300$$

$$K_i \text{ (c-b)} = \frac{h_{fb}}{1 + h_{ob}R_L} = \frac{-0.997}{1 + (0.091 \times 10^{-6} \times 2 \times 10^5)}$$

$$= -0.979$$

i.e. current gain in c-e is high, in c-b slightly less than 1. From Eqn.(5.9):

$$R_{in} \text{ (c-e)} = h_{ie} - h_{re}K_iR_L$$

$$= 4500 - (2.0 \times 10^{-4} \times 300 \times 3300) = 4302\ \Omega$$

$$R_{in} \text{ (c-b)} = h_{ib} - h_{rb}K_iR_L$$

$$= 13.6 - (4.08 \times 10^{-4} \times -0.979 \times 2 \times 10^5)$$

$$= 93.5\ \Omega$$

i.e. input resistance is moderately high for c-e but low for c-b (but note that even this arises mostly from $h_{rb}R_L$, i.e. the reflection of the load resistance back through the transistor). From Eqn. (5.12):

$$R_{out}\text{(c-e)} = \frac{h_{ie} + R_S}{h_{oe}(h_{ie} + R_S) - h_{fe}h_{re}}$$

$$= \frac{4500 + 1000}{30 \times 10^{-6}(5500) - 330 \times 2 \times 10^{-4}}$$

$$= 55.6\ \text{k}\Omega$$

$$R_{out}\text{(c-b)} = \frac{h_{ib} + R_S}{h_{ob}(h_{ib} + R_S) - h_{fb}h_{rb}}$$

$$= \frac{13.6 + 1000}{0.091 \times 10^{-6}(1013.6) + 0.997 \times 4.08 \times 10^{-4}}$$

$$= 2.03\ \text{M}\Omega$$

i.e. both output resistances are high, especially in the common-base mode. From Eqn. (5.14):

$$K_v\text{ (c-e)} = \frac{-K_iR_L}{R_{in}} = \frac{-300 \times 3300}{4302} = -230$$

$$K_v\text{ (c-b)} = \frac{-K_iR_L}{R_{in}} = \frac{0.979 \times 2 \times 10^5}{93.5} = 2094$$

c-e therefore provides good voltage gain with a phase reversal, c-b has a high voltage gain even though the current gain is less than unity. The formula shows this to be due to the input and output currents operating within very different resistances ($2 \times 10^5\ \Omega$ compared with $93.5\ \Omega$).

Repeating the whole process for c-c amplifiers with say a 1.2 kΩ load also demonstrates the general features of common-collector. From Eqns. (5.22) to (5.25)

$$h_{ic} = 4500\ \Omega\ ,$$

$$h_{fc} = -331\ ,$$

$$h_{oc} = 30 \times 10^{-6}\ \text{S}\ ,$$

$$h_{rc} = 0.9998$$

and substituting in the general equations for K_i , K_v , R_{in} and R_{out} gives:

K_i (c-c) $= -319.5$ . . . . high

K_v (c-c) $= 0.989$ . . . . ≈ 1

R_{in} (c-c) $= 387.8$ kΩ . . . . high

R_{out} (c-c) $= 16.6\ \Omega$ . . . . very low

thus we see the features of the *emitter follower*, a voltage gain of about unity ("following" the input voltage) with low output resistance but high input resistance, hence its use for resistance conversion or as a *buffer amplifier*.

Frequently transistor stage design can tolerate some loss of accuracy so it is propitious to round off this Section with a table of approximate relationships. The h-parameters for common-emitter only are given since these are the most likely to be available (Table 5.1).

5.2 FIELD-EFFECT TRANSISTORS

The f.e.t. differs from its forerunner, the bipolar transistor mainly in that the current path is through one type of semiconductor material only, i.e. wholly through n-type or through p-type. The path is known as a *channel* and the connexions

Table 5.1 *h*-PARAMETERS – FORMULAE FOR APPROXIMATE DESIGN

	Common-Emitter	Common-Base	Common-Collector
Current gain, K_i	h_{fe}	$\frac{-h_{fe}}{1 + h_{fe}}$	$-(1 + h_{fe})$
Voltage gain, K_v	$\frac{-h_{fe}R_L}{h_{ie}}$	$\frac{h_{fe}R_L}{h_{ie}}$	1
Input resistance, R_{in}	h_{ie}	$\frac{h_{ie}}{1 + h_{fe}}$	$h_{ie} + (1 + h_{fe})R_L$
Output resistance, R_{out}	$\frac{1}{h_{oe}}$	$\frac{1 + h_{fe}}{h_{oe}}$	$\frac{h_{ie} + R_S}{1 + h_{fe}}$

to it are labelled *source* and *drain*. Control is via the *gate* and effectively the gate potential controls the channel conductivity either by reverse biassing a p-n junction which surrounds the channel (*junction gate f.e.t.* – j.u.g.f.e.t.) or alternatively by creating an electrostatic field from an insulated gate. In this case the gate has no direct connexion with the channel, accordingly the device is known as an *insulated gate f.e.t.* (i.g.f.e.t.), or perhaps more generally as a *metal-oxide semiconductor f.e.t.* (m.o.s.f.e.t. or m.o.s.t.). The graphical symbols for the j.u.g.f.e.t. are shown in Figure 5.5(i).

From the analysis point of view because the gate input circuit is that of a reverse-biassed p-n junction, it has a sufficiently high input resistance (up to 10^4 MΩ) that the input circuit can be represented as having a disconnexion as shown in Figure 5.5(ii). At the higher frequencies however we must talk in terms of input impedance rather than resistance because gate capacitance although usually no more than 5 pF, can for some devices reach 15 pF or more.

Figure 5.5(ii) is for the *common-source* amplifier, the configuration most frequently used and for which information is usually readily available from manufacturers. In such data y-parameters are often quoted instead of h-parameters because they suit f.e.t.'s better. They are easily measured and all have the nature of admittances, thus any y-parameter multiplied by a voltage calculates a current. We are perhaps less accustomed to working in admittances rather than impedances but it does have the advantage that admittances in parallel are directly additive whereas impedances are not, that is, for Z_1 and Z_2 in parallel, the net impedance, Z is derived from

$$\frac{1}{Z} = \frac{1}{Z_1} + \frac{1}{Z_2}$$

but for the corresponding admittances, simply

$$Y = Y_1 + Y_2 \, .$$

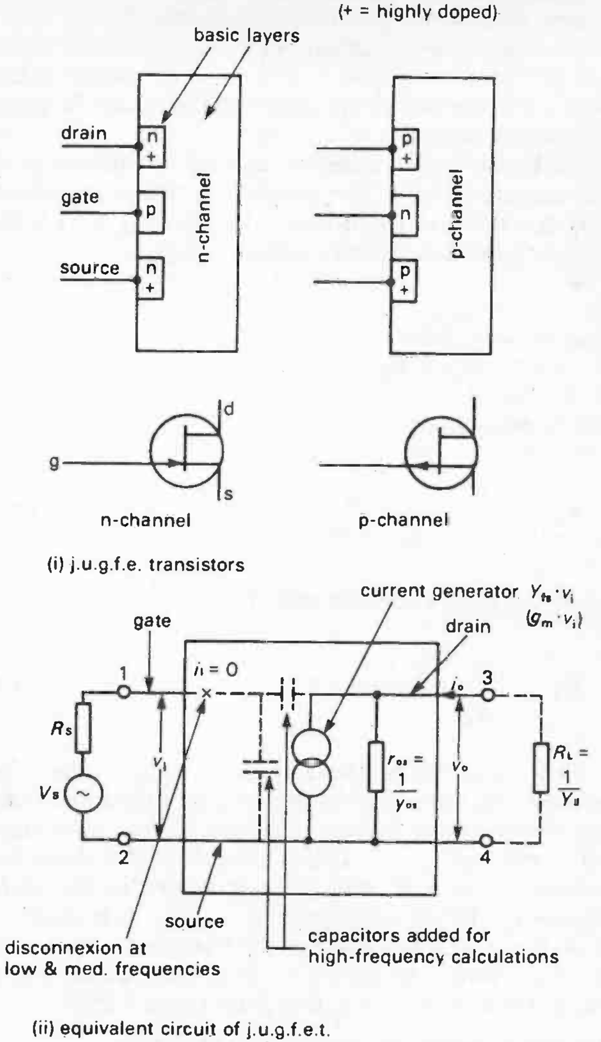


Fig. 5.5 The field-effect transistor

On this subject a reminder may be helpful:

admittance is the reciprocal of impedance;

conductance in a d.c. circuit is the reciprocal of resistance. For an a.c. circuit we need to refer back to Section 4.2.2.4. In both cases the unit is the siemen (S) but quite frequently μA/V or mA/V is used.

From Figure 5.5(ii) therefore, the net admittance of the output admittance y_{os} (the subscript s indicates common-source) and the load admittance, Y_L is $(y_{os} + Y_L)$ so if the current generator delivers a current $-y_{fs}v_i$

$$v_o = -\frac{y_{fs}v_i}{y_{os} + Y_L}$$

so that, Voltage gain,

$$K_v = \frac{v_o}{v_i} = \frac{-y_{fs}}{y_{os} + Y_L} \qquad (5.26)$$

and if y_{os} is small compared with Y_L,

$$K_v = \frac{-y_{fs}}{Y_L} \text{ (approx.)} \qquad (5.27)$$

y_{fs} is known as the *transfer admittance*, y_{os} as the *output admittance*, the minus sign as before, indicating a 180° phase change from input to output. The symbol g_m is frequently found in place of y_{fs}. It is a symbol handed down from thermionic valves and with these is known as the *mutual conductance*. Whether labelled y_{fs} or g_m it is simply the ratio of the current generated in the transistor output circuit to the input voltage producing it. With g_m calculations are mainly in terms of resistance, thus from Figure 5.5(ii):

net resistance connected across current generator

$$= \frac{r_{os}R_L}{r_{os} + R_L} \ \Omega$$

so that

$$v_0 = -g_m v_i \times \frac{r_{os} R_L}{r_{os} + R_L}$$

and

$$K_v = \frac{v_0}{v_i} = -g_m \times \frac{r_{os} R_L}{r_{os} + R_L} \qquad (5.28)$$

and if R is small compared with r_{os},

$$K_v = -g_m R_L \text{ (approx.)} \qquad (5.29)$$

EXAMPLE 5.8:
An f.e.t. used as a voltage amplifier has the following y-parameters, $y_{fs} = 5 \times 10^{-3}$ S (5 mA/V), $y_{os} = 25 \times 10^{-6}$ S (25 μA/V). Calculate the voltage gains with load resistances of 20 kΩ and 2kΩ.

When $R_L = 20\ \text{k}\Omega$, $Y_L = \dfrac{1}{20 \times 10^3} = 5 \times 10^{-5}$ S

From Eqn. (5.26)

$$K_v = \frac{-y_{fs}}{y_{os} + Y_L} = -\frac{5 \times 10^{-3}}{(25 \times 10^{-6}) + (5 \times 10^{-5})}$$

$$= -\frac{5 \times 10^{-3}}{(2.5 \times 10^{-5}) + (5 \times 10^{-5})} = -66.67$$

clearly y_{os} is not sufficiently small compared with Y_L for the approximate formula to apply.

When $R_L = 2\ \text{k}\Omega$, $Y_L = \dfrac{1}{2000} = 5 \times 10^{-4}$ S

$$\therefore \quad K_v = -\frac{5 \times 10^{-3}}{(25 \times 10^{-6}) + (5 \times 10^{-4})}$$

$$= -\frac{5 \times 10^{-3}}{(0.25 \times 10^{-4}) + (5 \times 10^{-4})} = -9.52$$

in this case y_{os} is small compared with Y_L hence the approximate formula can apply [Eqn.(5.27)]

$$K_v = \frac{-y_{fs}}{Y_L} \text{ (approx.)} = \frac{-5 \times 10^{-3}}{5 \times 10^{-4}} = -10$$

which is only 5% high.

Working instead in resistances and recalling that

$$g_m = y_{fs},$$

$$r_{os} = \frac{1}{y_{os}} = \frac{1}{25 \times 10^{-6}} = 40{,}000\ \Omega$$

From Eqn. (5.28)

When $R_L = 20\ \text{k}\Omega$, $K_v = -g_m \dfrac{r_{os}R_L}{r_{os} + R_L}$

$$= \frac{-5 \times 10^{-3} \times 40{,}000 \times 20{,}000}{40{,}000 + 20{,}000} = -66.67$$

When $R_L = 2\ \text{k}\Omega$,

$$K_v = \frac{-5 \times 10^{-3} \times 40{,}000 \times 2000}{42{,}000} = -9.52$$

The common-source f.e.t. amplifier therefore has extremely high input resistance with useful voltage amplification.

5.3 LOAD LINES

In the examples given in the preceding discussion of *h*-parameters the load resistance is invariably stated. To design a practical circuit therefore the first decision following choice of the transistor is that of the load value to be used, that is, the R_L which appears in Figures 5.2 and 5.5. A load is necessary because the variations in collector current must be detected in such a way as to drive a subsequent stage or other circuit. A transformer does this admirably but is normally out of the question because of bulk and cost, accordingly we revert when possible to the simple resistor which detects current variations by development of a voltage. This has a secondary effect however, for considering the transistor stage in Figure 5.6(i), it is evident that when a current I_c flows through R_L the collector voltage:

$$V_{ce} = V_{cc} - I_c R_L \tag{5.30}$$

hence most characteristic curves supplied by manufacturers cannot be used directly because a constant value of V_{ce} is assumed. To overcome this difficulty the technique of adding a *load line* to the output characteristics is available.

If a high impedance is connected to terminal 3 in Figure 5.6(i), the static and signal currents both operate in the same load, R_L and a single load line is sufficient. Consider the typical set of output characteristics shown at the right-hand side of Figure 5.7 and the tentative load line for 300 Ω. At say, I_c = 30 mA, from Ohm's Law the voltage dropped across 300 Ω is 9 V, hence if V_{cc} = 15 V, $V_{ce} = 15 - 9 = 6$ V and this value is read from the V_{ce} axis, the load line is in fact working out Equation (5.30) for us. In this particular case if I_c could swing from 0 to 50 mA, V_{ce} would move from 15 down to zero volts. We now have a better picture of what happens in Figure 5.6(i) for a certain value of R_L. Other load lines can be drawn as required once the supply voltage V_{cc} is determined. The method is simple, a line is drawn from the supply voltage value on the V_{ce} axis to the point on the I_c axis where theoretically all the supply voltage is

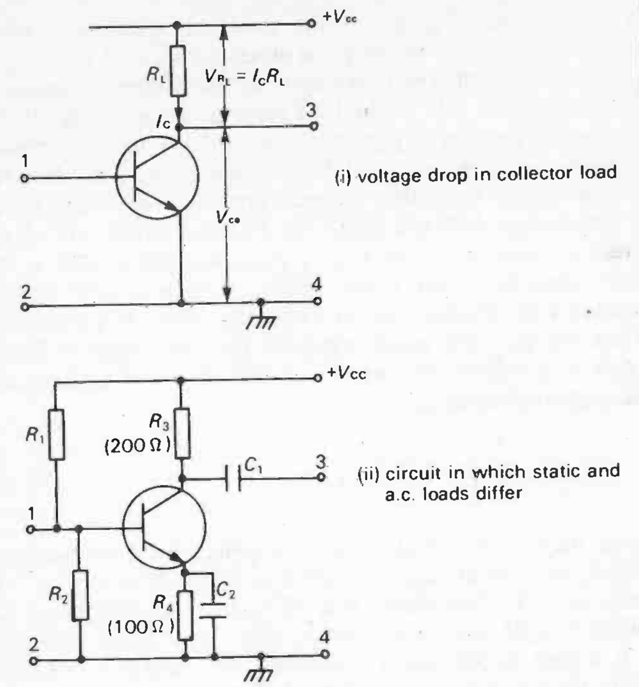


Fig. 5.6 Loading transistors

dropped across the load, i.e. in Figure 5.7 at

$$I_c = \frac{V_{cc}}{R_L} = \frac{15}{300}\ \mathrm{A} = 50\ \mathrm{mA}\ .$$

An informative but not essential addition is to draw the *dynamic current transfer characteristic* so that the relationship between I_b and I_c for the particular load line is also displayed. This is shown to the left in Figure 5.7, all the

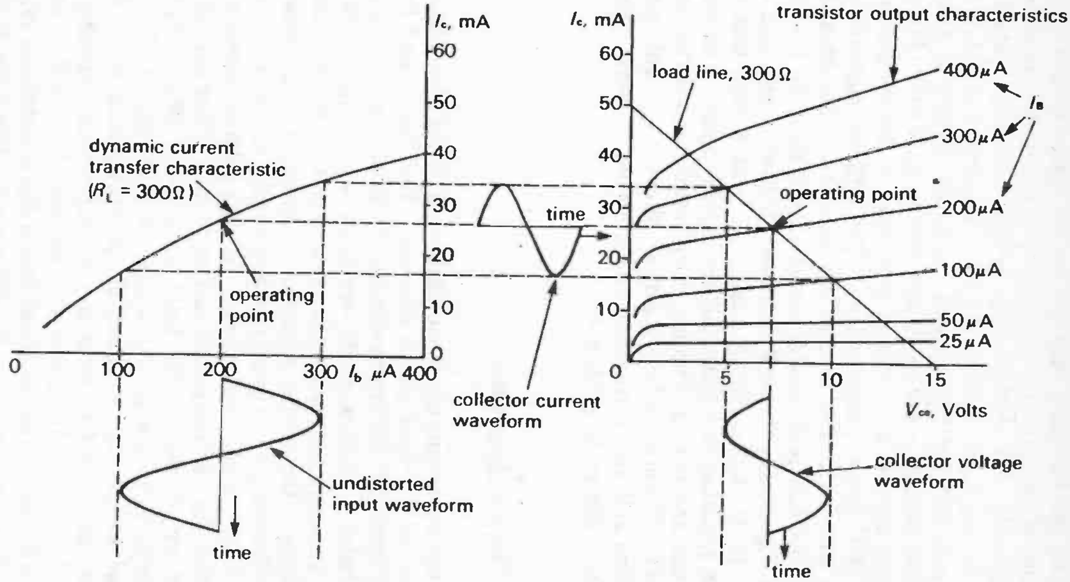


Fig. 5.7 Load line construction

information for it being obtained from the points of intersection of the load line with the output characteristics. The curvature of the transfer characteristic becomes evident so demonstrating the non-linearity which is always present. The *operating point* is chosen to be at the centre of the most linear part and the input signal amplitude limited to avoid excursions beyond. The small sine waves added to the Figure show how the input waveform progresses to the output and that a 180° phase change occurs. The collector current and voltage waveforms also indicate the presence of non-linearity for the first half-cycle is of lower amplitude than the second.

The Figure also shows that for this transistor and load, at $V_{CC} = 15$ V the operating point occurs at a base bias of 200 μA resulting in a quiescent (no signal) collector current of 26 mA and the d.c. biassing circuit must provide this (see Sect.5.4). Because collector current flows under quiescent conditions and also throughout the cycle of the input signal the bias is referred to as *Class A*.

5.3.1 The A.C. Load Line

When the static and signal currents work in different values of load, an *a.c. load line* is required. The "d.c." load line as in Figure 5.7 cannot be dispensed with however because it is needed to determine the operating point. Figure 5.6(ii) shows a circuit for which the requirement of the two load lines arises. The circuit employs a commonly used biassing system (biassing is discussed later) in which R_4 and C_2 are part. Under static conditions C_2 has infinite impedance therefore R_4 is unaffected and so is part of the collector-emitter circuit making the total static load $(R_3 + R_4)$. Conversely at frequencies at which the reactance of C_2 is low and hence R_4 is effectively short-circuited, then R_3 becomes the load (for simplicity we assume that nothing is connected to terminals 3 and 4). Suppose $R_3 = 200\ \Omega$ and $R_4 = 100\ \Omega$. The d.c. load line is for a resistance value $(R_3 + R_4)$, i.e. 300 Ω so for the same transistor Figure 5.7 applies. This has been re-drawn in Figure 5.8. For a.c. R_3 only is effective, i.e. 200 Ω.

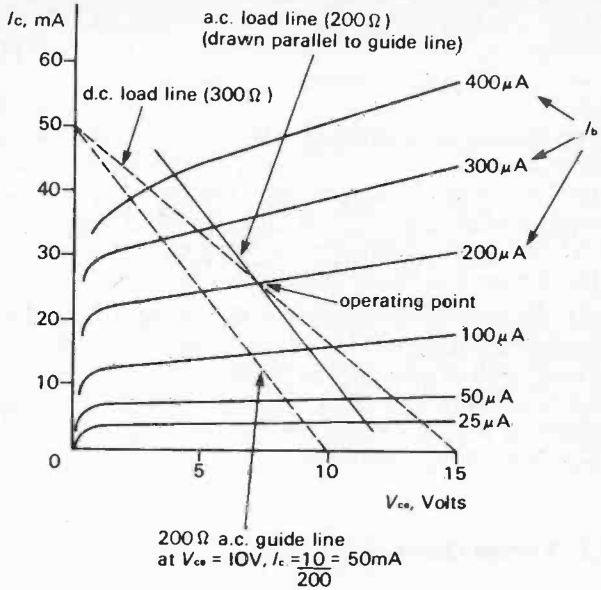


Fig. 5.8 Adding an a.c. load line

Remembering that the a.c. load line must pass through the operating point and that it is its *slope* which matters, a guide line is first constructed at any convenient position on the graph (all lines of the same resistance have the same slope). Next the load line is drawn parallel to the guide line passing through the operating point already positioned by the d.c. load line as shown in Figure 5.8.

5.3.2 Current Gain of Bipolar Transistor

The fact that Figure 5.7 shows graphically both the input waveform and also that of the corresponding collector current leads directly to an estimation of current gain thus

$$K_i = \frac{I_{c(max)} - I_{c(min)}}{I_{b(max)} - I_{b(min)}} \tag{5.31}$$

and for the transistor of Figure 5.7,

$$K_i = \frac{(34 - 16)\text{mA}}{(300 - 100)\mu\text{A}} = \frac{18 \times 10^3\ \mu\text{A}}{200\ \mu\text{A}} = 90\ .$$

No example is included for this Section nor for the preceding one because this would merely be a repetition of the work done in constructing Figures 5.7 and 5.8 but with a different set of output characteristics. More practice however is gained in the next Section which more or less repeats the technique but for a different class of transistor.

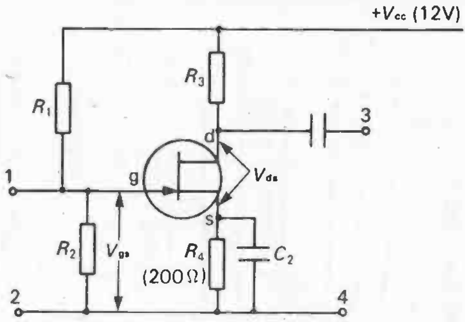
5.3.3 Voltage Gain of F.E.T.

The very high input resistance of an f.e.t. means that normally the device is used as a voltage amplifier. Figure 5.9 shows a single stage f.e.t. amplifier, the circuit is not unlike that of Figure 5.6 for the bipolar transistor. R_3 is the a.c. load but $(R_3 + R_4)$ form the d.c. load. R_1 and R_2 are biassing resistors which are considered in Section 5.4.

An a.c. input voltage applied between gate and source results in a drain voltage swing according to the value of the load and

$$K_v = \frac{V_{ds(max)} - V_{ds(min)}}{V_{gs(max)} - V_{gs(min)}} \tag{5.32}$$

EXAMPLE 5.9:
Figure 5.9(ii) shows at the right the output characteristics of an n-channel f.e.t. It is connected in common-source as in Figure 5.9(i) with a supply voltage of 12. By means of the

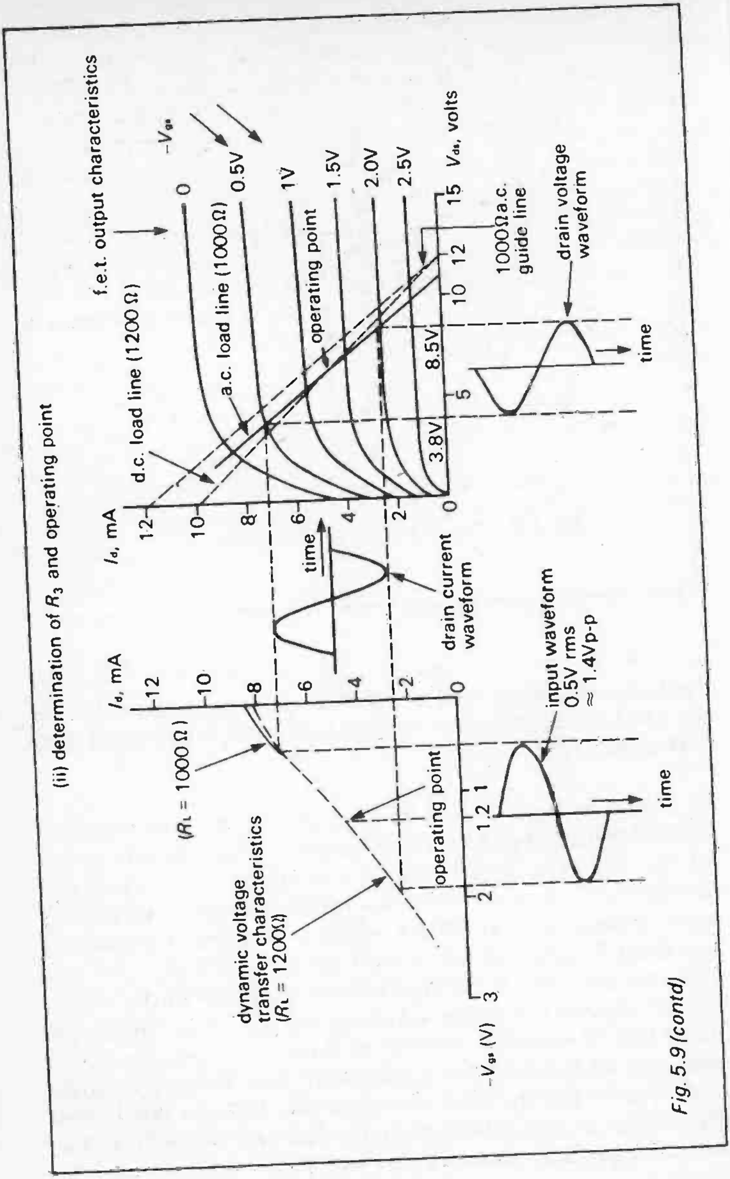


(i) common-source circuit

Fig. 5.9 Single stage f.e.t. amplifier

load line technique find a suitable operating point, value for R_3 and the approximate voltage gain. The input signal does not exceed 0.5 V r.m.s.

Not knowing the value of the d.c. load, a provisional load line must first be positioned by instinct, being especially aware of the requirement of minimum distortion. The bottom end is fixed at V_{ds} = 12 V and a reasonable point on the I_d scale appears to be at 10 mA. Such a load line therefore has a resistance value of (12 V)/(10 mA) = 1200 Ω. The line is drawn and from the points where it cuts each of the output characteristics, the dynamic voltage transfer characteristic for 1200 Ω is projected to the left as shown. It is not absolutely necessary to produce this characteristic but it does aid clarity and helps in placing the operating point. If R_4 = 200 Ω and $(R_3 + R_4)$ = 1200 Ω [Fig.5.9(i)], then R_3 = 1000 Ω. The operating point is now chosen, say at $V_{gs} = -1.2$ V.



(ii) determination of R_3 and operating point
f.e.t. output characteristics
$-V_{gs}$
0
0.5V
1V
1.5V
2.0V
2.5V
V_{ds}, volts
d.c. load line (1200 Ω)
a.c. load line (1000 Ω)
operating point
I_d, mA
12
10
8
6
4
2
0
3.8V
5
8.5V
10
12
15
1000 Ω a.c. guide line
drain voltage waveform
time
drain current waveform
$(R_L = 1000\,\Omega)$
dynamic voltage transfer characteristics $(R_L = 1200\,\Omega)$
operating point
1
1.2
2
3
$-V_{gs}$ (V)
input waveform
0.5V rms
≈ 1.4Vp-p

Fig. 5.9 (contd)

Drawing the a.c. load line for 1000 Ω parallel to a 1000 Ω guide line between I_d = 12 mA and V_{ds} = 12 V (or I_d = 8 ma , V_{ds} = 8 V etc.) enables an a.c. transfer characteristic to be drawn to the left. It is evident that the d.c. and a.c. transfer characteristics are almost identical except at low values of $-V_{gs}$ so indicating that it may often be adequate to produce the static (d.c.) load line only.

The maximum input signal is 0.5 V r.m.s., i.e. $0.5 \times \sqrt{2} \times 2 \approx 1.4$ V peak–peak, this is shown graphically as the input waveform in the Figure and by projecting across via the transfer characteristic to the a.c. load line the drain voltage is seen to swing from 3.8 to 8.5 V, therefore:

$$\text{Voltage gain, } K_v = \frac{8.5 - 3.8}{1.4} = 3.36$$

and there is 180° phase change from input to output.

5.4 BIAS AND STABILIZATION

Because of the wide variations in characteristics between transistors of the same type and of the effects of temperature, special precautions are necessary to assure stability of the operating point.

5.4.1 Leakage Current

Energy from heat causes some covalent bonds to break down and release valence electrons (Sect.1.3.1.2) so creating an unwanted flow of current in addition to and unaffected by the main operative current. The unwanted current is known as the *collector leakage current* and for a common-base connected transistor it is designated I_{CBO} , the B of the subscript indicating common-base and the O showing that the third electrode (the emitter in this case) is open-circuited. For the common-emitter connexion the leakage current is designated I_{CEO} , in this case the base is on open-circuit. In operation therefore, in common-emitter, the total collector current, I_C is derived from the amplified input current, I_B plus the collector leakage current, i.e.

$$I_C = h_{FE}I_B + I_{CEO} \tag{5.33}$$

where h_{FE} is the d.c. current gain, sometimes referred to as the *static forward current transfer ratio* (h_{fe} is the *small signal* current gain but usually h_{FE} and h_{fe} for a given transistor are reasonably similar). Manufacturers' data is likely to include a figure for h_{FE}.

Similarly for common-base:

$$I_C = h_{FB}I_E + I_{CBO} \tag{5.34}$$

Considering the common-emitter connexion, I_{CBO} arises in the base and flows across the collector-base junction into the collector circuit. By normal transistor action such a current in the base also gives rise to a current of $h_{FE}I_{CBO}$ in the same collector circuit. Hence the total leakage current in the output circuit for common-emitter:

$$I_{CEO} = I_{CBO} + h_{FE}I_{CBO}$$

$$\therefore \quad I_{CEO} = (h_{FE} + 1)I_{CBO} \tag{5.35}$$

If the temperature of the collector-base junction rises, I_{CBO} increases and the greater power dissipated in the junction causes the temperature to rise further, so accelerating the breakdown of covalent bonds. The effect is cumulative and without restraint the physically minute transistor junction can be destroyed, the process is known as *thermal runaway*.

5.4.2 Stability Factor

Stabilization is accomplished through a form of d.c. negative feedback which counteracts any occurrence tending to change the steady value of the collector current. It is useful to quantify stability so that various methods can be compared and this can be done by the *stability factor*. There are various ways of defining it, one in common use being by the ratio of I_C when stabilized to its value unstabilized. This results in a general formula, sufficiently accurate for most calculations:

$$\text{Stability Factor}, S = \frac{1}{1 + h_{FE}F} \qquad (5.36)$$

where F is the fraction of collector current which is fed back to the base. As might be expected the formula is not unlike that for a.c. negative feedback (Sect.5.9) and evidently the greater the values of h_{FE} and F, the lower the stability factor so a small value of S is an indication of good stability. When $F = 0$, i.e. no fraction of I_C is fed back, then $S = 1$, representing no stabilization at all. This is the case in the circuit of Figure 5.10(i) in which the base bias is derived directly from the supply and a look at the range of collector currents obtainable with a commercial transistor used in this circuit clearly demonstrates the need of stabilization.

EXAMPLE 5.10:
An n-p-n transistor of a type for which h_{FE} is quoted as 200 – 800 has a leakage current, I_{CBO} of 15 μA. It is used in the circuit of Figure 5.10(i) where $V_{cc} = 9$ V, $R_b = 330$ kΩ, $R_c = 2000\ \Omega$ and $V_{BE} = 620$ mV. What is the collector current of a transistor of h_{FE} 200 ? In the unlikely but possible event of the transistor being changed for one of h_{FE} 800, to what value would the collector current rise?

$$\text{Voltage across } R_b = 9 - 0.62 = 8.38 \text{ V}$$

$$\therefore \quad \text{base current}, I_B = \frac{8.38}{330 \times 10^3} \times 10^6 \mu\text{A} = 25.4\ \mu\text{A}$$

From Eqns. (5.33) and (5.35):

$$I_C = h_{FE}I_B + (h_{FE} + 1)I_{CBO}$$

so when $h_{FE} = 200$

$$I_C = (200 \times 25.4) + (201 \times 15)\ \mu A = 8.1\ mA\ .$$

Also when $h_{FE} = 800$

$$I_C = (800 \times 25.4) + (801 \times 15)\ \mu A = 32.3\ mA\ .$$

[Note that there is very little loss of accuracy if $(1 + h_{FE})$ is taken as h_{FE} , then $I_C \approx h_{FE}(I_B + I_{CBO})$.]

The results clearly indicate that changing a transistor for one of the same type in such a circuit as in Figure 5.10(i) could completely invalidate the principle of fixing the operating point.

5.4.2.1 Biassing Circuits

We can best appreciate the design features for d.c. stabilization by analysing both a simple arrangement and a more complicated but nevertheless popular one. This will also give some indication of the range of values of the stability factor. Figure 5.10(ii) shows the simple method in which the base current is supplied via R_b from the collector. Any rise in collector current reduces the collector voltage hence the current through R_b . The drop in base current reduces the collector current so tending to compensate for the original change. To use Equation (5.36) for the stability factor we first need to calculate F . The circuit of Figure 5.10(ii) can be re-drawn as shown and it is evident that the current I_C is made up by the currents through R_b and R_c in parallel, hence:

$$\text{Current through } R_b = I_C \times \frac{R_c}{R_b + R_c} \quad \text{(Book 1, Sect.3.2.4)}$$

therefore the fraction of current fed back:

$$F = \frac{R_c}{R_b + R_c}$$

which we note is in terms of the circuit resistor values only.

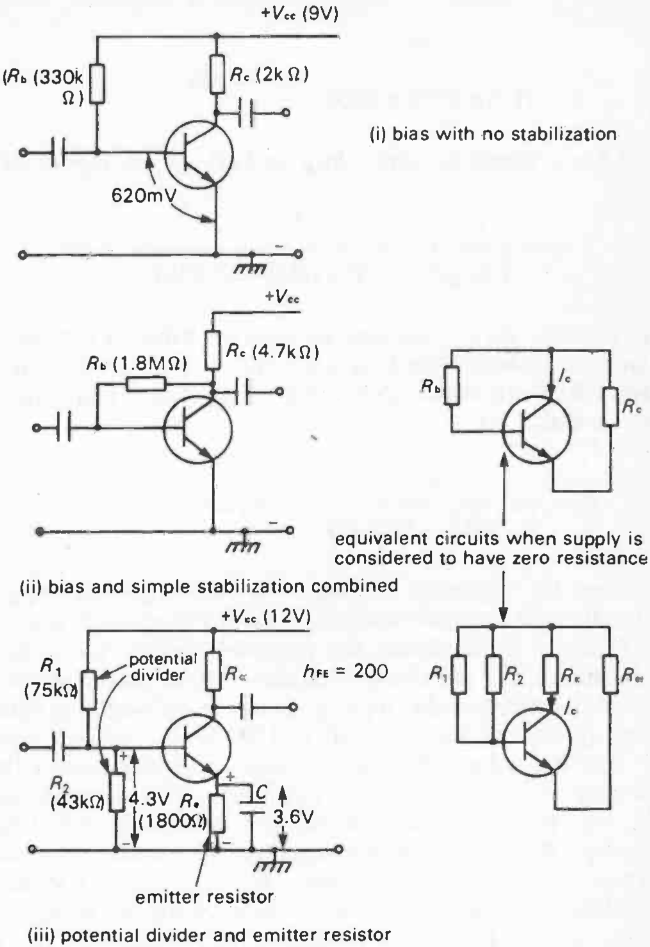


Fig. 5.10 Bias and stabilization circuits

EXAMPLE 5.11:
Calculate the stability factor for the circuit of Figure 5.10(ii) for values of h_{FE} of 200 and 800, using the resistor values shown.

$$F = \frac{4700}{(1.8 \times 10^6) + 4700} = 0.0026$$

and for a transistor with $h_{FE} = 200$, from Eqn.(5.36):

$$S = \frac{1}{1 + h_{FE}F} = \frac{1}{1 + (200 \times 0.0026)} = 0.66$$

i.e. collector current changes are not even halved, e.g. a rise of 1 mA in an unstabilized circuit would only be reduced to a rise of 0.66 mA with stabilization as in Figure 5.10(ii). With $h_{FE} = 800$:

$$S = \frac{1}{1 + (800 \times 0.0026)} = 0.32$$

showing the improved stability with high values of h_{FE}, a result to be expected simply because the feedback is greater.

Figure 5.10(iii) shows the potential divider and emitter resistor bias and stabilization circuit. Design might proceed as follows assuming that from the load line the operating point gives $I_C = 2$ mA for a d.c. load of 6100 Ω. R_e is first chosen so that the voltage developed across it is several times the working base-emitter voltage (say, for a silicon transistor, 0.7 V). It is a compromise because making R_e too large reduces R_c and hence the signal voltage developed at the output. Assume a voltage across R_e of 3.6 V. We seldom get things right the first time but adjustments can always be made later. Then, d.c. voltage developed across R_e = $2 \times 10^{-3} \times 1800 = 3.6$. For a transistor of $h_{FE} = 200$,

$$I_B = \frac{2}{200} \text{ mA} = 10\ \mu\text{A}.$$

Let $(R_1 + R_2)$ have a drain of at least 10 times the base current so that changes in the latter have an insignificant effect on the base potential, i.e. the potential divider current is $10 \times 10\,\mu A = 0.1$ mA. Then

$$(R_1 + R_2) = \frac{12}{0.1 \times 10^{-3}} = 120\text{ k}\Omega\,.$$

The voltage across R_e makes the base *negative* to the emitter by 3.6 V. To make it *positive* to the emitter by 0.7 V therefore the voltage drop across R_2 must be $3.6 + 0.7 =$ 4.3 V [see Figure for polarities and note that this is for an n-p-n transistor, for p-n-p all polarities are reversed].

Therefore:

$$R_2 = \frac{4.3}{0.1 \times 10^{-3}} = 43\text{ k}\Omega\,,$$

this happens to be a preferred value and accordingly

$$R_1 = 120 - 43\text{ k}\Omega = 77\text{ k}\Omega\,,\ \text{say } 75\text{ k}\Omega\,,$$

the nearest preferred value.

The equivalent circuit of Figure 5.10(iii) shows that, the two resistors R_1 and R_2 are effectively in parallel in the base circuit so that:

$$F = \frac{R_e}{R + R_e} \qquad \text{where } R = \frac{R_1 R_2}{R_1 + R_2}\,.$$

EXAMPLE 5.12:
Calculate the stability factor for the circuit of Figure 5.10(iii).

$$R = \frac{R_1 R_2}{R_1 + R_2} = \frac{75 \times 43}{75 + 43} = 27.33\text{ k}\Omega$$

$$F = \frac{R_e}{R + R_e} = \frac{1800}{27330 + 1800} = 0.06$$

therefore

$$S = \frac{1}{1 + h_{FE}F} = \frac{1}{1 + (200 \times 0.06)} = 0.077$$

thus with this more complex circuit the collector current variations are reduced to 0.077 of their unstabilized values, that is, to about one thirteenth. With $h_{FE} = 800$, $S = 0.02$, there is a reduction to about one fiftieth. $S = 0.02$ is a good stability factor although values down to 0.01 or less are obtainable.

To avoid signal negative feedback by R_e, a large capacitor, C of reactance some one-tenth of the resistance of R_e, at the lowest frequency to be amplified is added as shown.

5.5 MULTISTAGE AMPLIFIER BANDWIDTH

Although for many purposes IC amplifiers are taking over, the standard RC-coupled amplifier still has its part to play. A simple examination of the factors affecting the shape of the gain/frequency characteristics of such amplifiers is therefore useful.

Input and output impedance calculations for the single stage are discussed in Section 5.1.3, good design requires that reasonable matching between stages is obtained. Common-base cannot provide this because the high output impedance of one stage would be connected to the low input impedance of the next. Common-collector stages have a voltage gain of a little less than 1 hence there is no point in having a multi-stage arrangement. Common-emitter amplifiers on the other hand, as mentioned in Section 5.1.3.3, are admirably suited to multi-staging because their output and input impedances are of similar order.

5.5.1 The Resistance-Capacitance Coupled Amplifier

Figure 5.11(i) shows the elements of a common-emitter stage (bias arrangements omitted) with coupling capacitor C_c blocking off d.c. to the next stage, represented by R_2 and C_2 in parallel. In (ii) is shown a first simplification in which the d.c. supply (V_{cc}) is assumed to have zero impedance to a.c. In (iii) the h-parameter equivalent circuit is shown, as in Figure 5.2(iii) but further simplified by neglecting the small effects of h_{re} and h_{oe}. By assuming that over certain frequency bands the effects of some of the capacitances are negligible, a reasonable prediction of the gain/frequency characteristics can be obtained.

5.5.1.1 Low Frequency Response

At low frequencies the effect of C_2 can be neglected, hence from Figure 5.11(iii), ignoring C_2 and noting that R_L and the network C_cR_2 are in parallel across the current generator:

$$\text{Impedance of } C_cR_2 \text{ network} = R_2 - \frac{j}{\omega C_c}$$

[for complex (j) numbers see Chapter 4].

The current $-h_{fe}i_i$ divides between the two parallel paths hence:

$$\text{current in } R_2 = -h_{fe}i_i \frac{R_L}{R_L + \left(R_2 - \dfrac{j}{\omega C_c}\right)}$$

so that

$$v_o = -h_{fe}i_i \frac{R_L}{R_L + R_2 - \dfrac{j}{\omega C_c}} \times R_2$$

and since

$$v_i = i_i h_{ie}$$

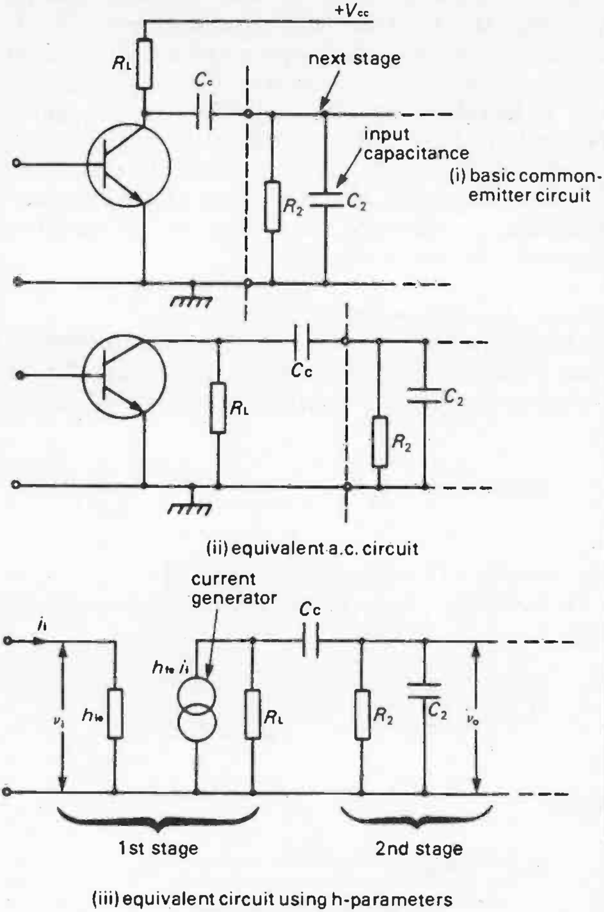


Fig. 5.11 RC coupling of transistor stages

Voltage gain,

$$A_v = \frac{v_o}{v_i} = \frac{-h_{fe}}{h_{ie}} \times \frac{R_L R_2}{\left(R_L + R_2 - \frac{j}{\omega C'_c}\right)} \qquad (5.37)$$

Recalling from Chapter 4 that rationalizing an expression having a −j term in the denominator results in a +j term in the numerator (Sect.4.2.2.1) shows that the expected 180° lag of v_o on v_i is in effect reduced by the presence of the capacitor C_c and that the degree of reduction depends on ω. At low frequencies therefore the amplifier phase shift is less than 180°, and the lower the frequency, the less the shift. This is discussed in more detail later.

5.5.1.2 Medium Frequency Response

In practical circuits C_c is likely to have a value of several microfarads thus whereas at low audio frequencies its reactance is appreciable (e.g. for 10 μF, 80 Ω at 200 Hz), at medium frequencies the reactance has fallen sufficiently for it to be neglected. C_2 is likely to be of the order of picofarads and although as frequency rises its reactance falls, at medium frequencies it still is unlikely to pose a problem. Not having to consider these two capacitances certainly makes the voltage gain equation more manageable for we now only need to remove the j term from Equation (5.37). Therefore, at medium frequencies:

$$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{(R_L + R_2)} \qquad (5.38)$$

and theoretically the phase change is exactly 180°.

5.5.1.3 High Frequency Response

It follows that at the higher frequencies the reactance of C_c is low enough to be negligible but that the effect of the reactance of C_2 is significant. Let the combined impedance of the parallel network R_2C_2 be Z [Fig.5.11(iii)], so using complex notation:

$$\frac{1}{Z} = \frac{1}{R_2} - \frac{\omega C_2}{j} = \frac{j - \omega C_2 R_2}{jR_2}$$

$$\therefore \quad Z = \frac{jR_2}{j - \omega C_2 R_2} = \frac{R_2}{1 + j\omega C_2 R_2}$$

(multiplying numerator and denominator by −j).

$$\text{Current in } R_2 C_2 = -\frac{h_{fe} i_i R_L}{(R_L + Z)}$$

hence

$$v_o = -\frac{h_{fe} i_i R_L}{(R_L + Z)} \times Z$$

and since

$$v_i = i_i h_{ie}$$

$$A_v = \frac{v_o}{v_i} = \frac{-h_{fe}}{h_{ie}} \times \frac{R_L Z}{(R_L + Z)}$$

substituting for Z,

$$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L \left(\dfrac{R_2}{1 + j\omega C_2 R_2}\right)}{R_L + \left(\dfrac{R_2}{1 + j\omega C_2 R_2}\right)}$$

i.e. at high frequencies:

$$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{R_L + R_2 + j\omega C_2 R_L R_2} \qquad (5.39)$$

A $-j$ in the denominator of the low frequency expression is shown to result in a reduction of the transistor 180° phase-change. A $+j$ in the denominator of Equation (5.39) thus indicates that at the higher frequencies the overall phase change is *greater* than 180°.

EXAMPLE 5.13:
An amplifier stage with a transistor of $h_{fe} = 270$, $h_{ie} = 4.5$ kΩ and load resistance 4.7 kΩ is coupled to the next stage by a 2 μF capacitor. The input circuit of the second stage is equivalent to 1.3 kΩ shunted by 1 nF. Draw the approximate gain and phase/frequency characteristics.

In Figure 5.11(iii):

$$R_L = 4700\ \Omega$$

$$R_2 = 1300\ \Omega$$

$$C_c = 2 \times 10^{-6}\ \text{F}$$

$$C_2 = 10^{-9}\ \text{F}$$

$$\frac{h_{fe}}{h_{ie}} = \frac{270}{4500} = 0.06$$

Maximum response is realized at medium frequencies, thus from Equation (5.38):

$$A_{v(max)} = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{(R_L + R_2)}$$

$$= -0.06 \times \frac{4700 \times 1300}{6000} = -61.1$$

We have no idea yet as to the frequencies at which the "low" and "high" formulae apply so trials may be necessary, for example at 100 Hz, from Equation (5.37):

$$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{R_L + R_2 - \dfrac{j}{\omega C_c}}$$

$$= -\frac{0.06 \times 4700 \times 1300}{6000 - \dfrac{j}{\omega C_c}}$$

$$= -\left(\frac{366600}{6000 - j796}\right) = -\left(\frac{366600 \angle 0^\circ}{6053 \angle -7.6^\circ}\right)$$

$$= -(60.6 \angle 7.6^\circ) = 60.6 \angle 187.6^\circ$$

so indicating that the voltage gain is 60.6 with the output voltage v_o *leading* on the input v_i by 187.6° or equally with v_o *lagging* on v_i by (360 − 187.6) = 172.4°. Technically both are correct but with the feeling that there cannot be an output until after an input has been applied, it is perhaps more acceptable to talk finally in terms of the amount v_o lags on v_i rather than the other way round. Hence

$$A_v = 60.6 \angle 172.4^\circ, \; v_o \text{ lagging on } v_i,$$

and evidently at this frequency the gain is just beginning to fall with the phase shifting to below 180°. Equation (5.37) therefore applies up to about 100 Hz.

The calculations necessary to produce the curves are somewhat tedious but less so with a scientific calculator or computer programmed to handle the square root of the sum of two squares easily. From the above trial calculation a formula for A_v with f as the variable may be employed, e.g.

$$A_v = -\frac{366600}{6000 - j\dfrac{79577}{f}}$$

with suggested table headings as follows. Two examples of the calculations are given.

At the high frequency end the gain begins to fall off above about 20 kHz so a table listing the value of

$$A_v = -\frac{366600}{6000 + j(0.0384f)}$$

enables this end of the curve to be drawn. Between 100 Hz an 20 kHz frequency has no effect because there is no j component in Equation (5.38), a straight line at $A_v = 61.1$ between these two frequency points is therefore sufficient.

The complete characteristics are given in Figure 5.12 and it is obvious from this example that with Equations (5.37) to (5.39) we have at our disposal a practical means of predicting gain performance.

5.5.1.4 Bandwidth

The *bandwidth* of an amplifier is most frequently specified as the range of frequencies between the 3 dB points, that is, the upper and lower frequencies at which the response is 3 dB below the maximum. Figure 5.12 shows that normally the response of a single stage is maximum at medium frequencies [Eqn.(5.38)] and if we can determine the upper and lower 3 dB points from the other two equations [(5.37) and (5.39)], then for any given amplifier we can calculate its bandwidth.

From Equation (5.37) it is evident that when

$$(R_L + R_2) = \frac{1}{\omega C_c},$$

f (Hz) (1)	$\frac{79577}{f}$ (2)	$\sqrt{[6000^2 + (\text{Col.2})^2]}$ (3)	$A_V = \frac{366600}{(\text{Col.3})}$ (4)	$\tan^{-1}\frac{(\text{Col.2})}{6000}$ (degrees) (5)	Phase shift = 180 + (Col.5) (degrees) (6)	Phase shift expressed as a lagging angle (degrees)
100	795.8	6053	60.6	7.6	187.6	172
50	1592	6208	59.1	14.9	194.9	165
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

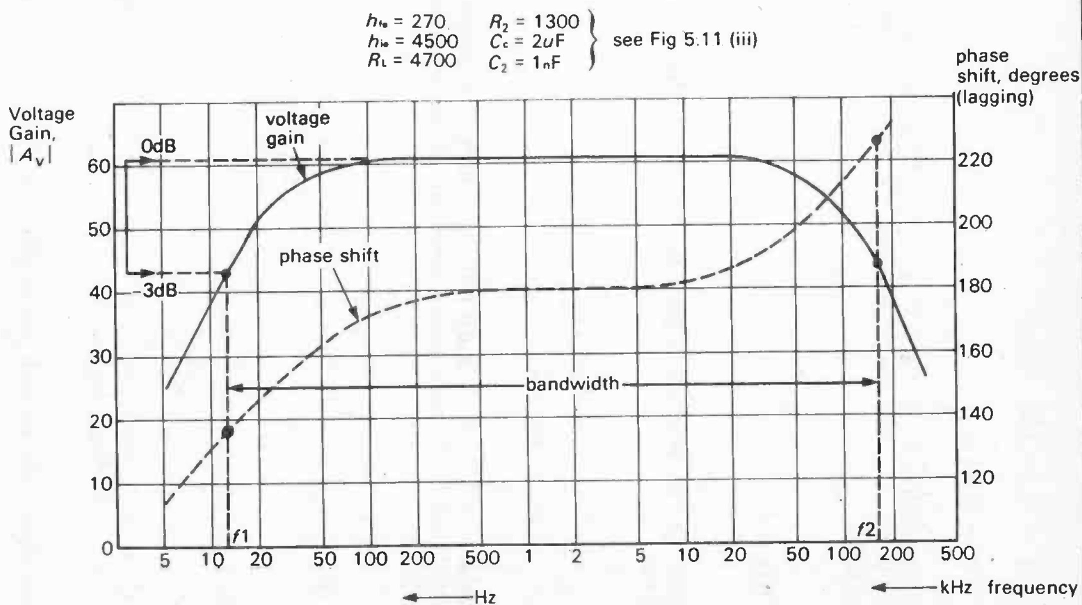


Fig. 5.12 Gain/frequency characteristic for RC coupled amplifier

this part of the denominator can be rewritten as

$$(R_L + R_2) - j(R_L + R_2),$$

i.e. $(R_L + R_2)(1 - j)$.

Now the modulus of $1 - j$ (i.e. $1 - j1$) is (from Sect.4.2.1) $\sqrt{1^2 + 1^2} = \sqrt{2}$ and the angle is $\tan^{-1}(-1/1) = -45°$, so that when

$$(R_L + R_2) = \frac{1}{\omega C_c},$$

Equation (5.37) becomes Equation (5.38) multiplied by $1/\sqrt{2}$, i.e.

$$A_v = \frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{R_L + R_2} \times \frac{1}{\sqrt{2}}$$

and because $1/\sqrt{2}$ is -3 dB (20 log 0.707), it is evident that the lower 3 dB point occurs when the above condition applies, i.e.

$$(R_L + R_2) = \frac{1}{\omega C_c} \quad \text{or} \quad \omega = \frac{1}{C_c(R_L + R_2)}$$

∴ lower 3 dB frequency,

$$f_1 = \frac{1}{2\pi C_c(R_L + R_2)} \tag{5.40}$$

and the phase-shift through the amplifier is $(180 - 45)° = 135°$.

The upper 3 dB frequency f_2 follows and again we have to find the condition which reduces Equation (5.39) to

$1/\sqrt{2}$ of the maximum value. This occurs when

$$\frac{R_L R_2}{R_L + R_2} = \frac{1}{\omega C_2}\,.$$

Hence, substituting in the equation:

$$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{(R_L + R_2) + j\dfrac{(R_L + R_2)}{R_L R_2} \times R_L R_2}$$

$$= -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{(R_L + R_2)(1 + j)}$$

$$= -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{(R_L + R_2)} \times \frac{1}{\sqrt{2}}$$

with an angle $\tan^{-1} 1 = 45°$, so the response is 3 dB down when

$$\frac{R_L R_2}{(R_L + R_2)} = \frac{1}{2\pi f_2 C_2}$$

i.e. the upper 3 dB frequency,

$$f_2 = \frac{R_L + R_2}{2\pi C_2 R_L R_2} \tag{5.41}$$

and the phase-shift through the amplifier is $(180 + 45) = 225°$.

EXAMPLE 5.14:
Calculate the bandwidth of the amplifier considered in the previous example (5.13).

$$R_L = 4700\ \Omega\,,$$

$$R_2 = 1300\ \Omega\,,$$

$$C_c = 2 \times 10^{-6}\ \mathrm{F}\,,$$

$$C_2 = 10^{-9}\ \mathrm{F}\,.$$

From Equation (5.40),

$$f_1 = \frac{1}{2\pi C_c(R_L + R_2)} = \frac{1}{2\pi \times 2 \times 10^{-6} \times 6000}$$

$$= 13.3\ \mathrm{Hz}$$

From Equation (5.41),

$$f_2 = \frac{R_L + R_2}{2\pi C_2 R_L R_2} = \frac{6000}{2\pi \times 10^{-9} \times 4700 \times 1300}\ \mathrm{Hz}$$

$$= 156.3\ \mathrm{kHz}$$

The amplifier bandwidth thus extends from 13.3 Hz to 156.3 kHz, i.e. over some 156 kHz. This is shown on Figure 5.12 from which it is also evident that at the lower 3 dB point the phase-shift through the amplifier is 135° and at the upper 3 dB point, 225° .

EXAMPLE 5.15:
A transistor working into a load of 3.3 kΩ is capacitance coupled to the next stage which has an input resistance of 600 Ω . What is the smallest preferred value of coupling capacitor which can be used if the response (3 dB) of the amplifier is to extend down to 90 Hz?

$$R_L = 3300\ \Omega\,,$$

$$R_2 = 600\ \Omega\,,$$

$$C_c = ?$$

From Equation (5.40),

$$f_1 = \frac{1}{2\pi C_c(R_L + R_2)}$$

$$\therefore \quad C_c = \frac{1}{2\pi f_1 (R_L + R_2)} = \frac{10^6}{2\pi \times 90 \times 3900}\ \mu F$$

$$= 0.45\ \mu F\ .$$

The nearest preferred value must be equal to or greater than 0.45 μF , e.g. 0.47 μF.

5.6 POWER AMPLIFIERS

As with small-signal amplifiers, with power amplification the integrated circuit has much to offer in efficiency, small size and low cost. They are freely available for up to moderate power outputs (say, 20 W) but above this the discrete component is usually more suitable. Despite the increasing use of IC's it is instructive to examine some of the basic features of power amplifiers especially the limitations on transistor power dissipation, on conversion efficiencies and distortion.

5.6.1 Transistor Operating Area

"Area" in this case refers to that part of the output characteristics within which the transistor can be safely operated. Dissipation of power in the junctions raises the temperature and as shown in Section 5.4.1 there is a limit. Hence design aims at ensuring that the junction power dissipation is never excessive. The single figure of maximum power dissipation, P_{tot} quoted by the manufacturer is all we need to determine the main boundary line and for common-emitter (now using capital letter subscripts to denote d.c. values):

$$P_{tot} = V_{CE}I_C + V_{BE}I_B \qquad (5.42)$$

but because $V_{BE}I_B$ is usually insignificant compared with $V_{CE}I_C$, we can generally work with

$$P_{tot} \approx V_{CE}I_C .$$

The power dissipated in a transistor is maximum under no-signal conditions for then none is being passed to the output. Thus because the output characteristics are drawn from d.c. values it is only necessary to draw the curve through the points where $V_{CE} \times I_C = P_{tot}$. This is shown on a typical set of characteristics in Figure 5.13 between points b and c (ignore at present the load line and other figures). The remainder of the safe operating area is given by the manufacturer in the form of $I_{C(max)}$ (a – b) and $V_{CE(max)}$ (c – d) for the particular transistor. Any load line must lie within the safe area otherwise overheating of the device may occur, possibly leading to *second breakdown*, a condition in which the resistance falls rapidly so leading to destruction unless an external current limiter is in use. Note from the Figure that for the particular maximum power quoted, a mounting base temperature is given, should this temperature be exceeded under working conditions, the dissipation is further limited.

EXAMPLE 5.16:
The output characteristics of a power transistor are given in Figure 5.13. For a mounting base temperature not exceeding 60°C the maximum collector dissipation is quoted as 1.5 W. $I_{C(max)}$ is 150 mA and $V_{CE(max)}$, 38 V. Determine the safe operating area. The transistor is coupled to an 8 Ω loudspeaker via an output transformer (primary winding resistance negligible). If the d.c. supply is 20 V, draw a suitable load line for maximum power output and calculate this for a sinusoidal input signal swing of 2 mA. Also calculate the output transformer turns ratio.

Referring to Figure 5.13 the section bc of the safe operating area is plotted from calculation of the currents at various voltages to produce 1.5 W of power. Thus from

$$I_C = \frac{1.5}{V_{CE}} \times 1000 \text{ mA},$$

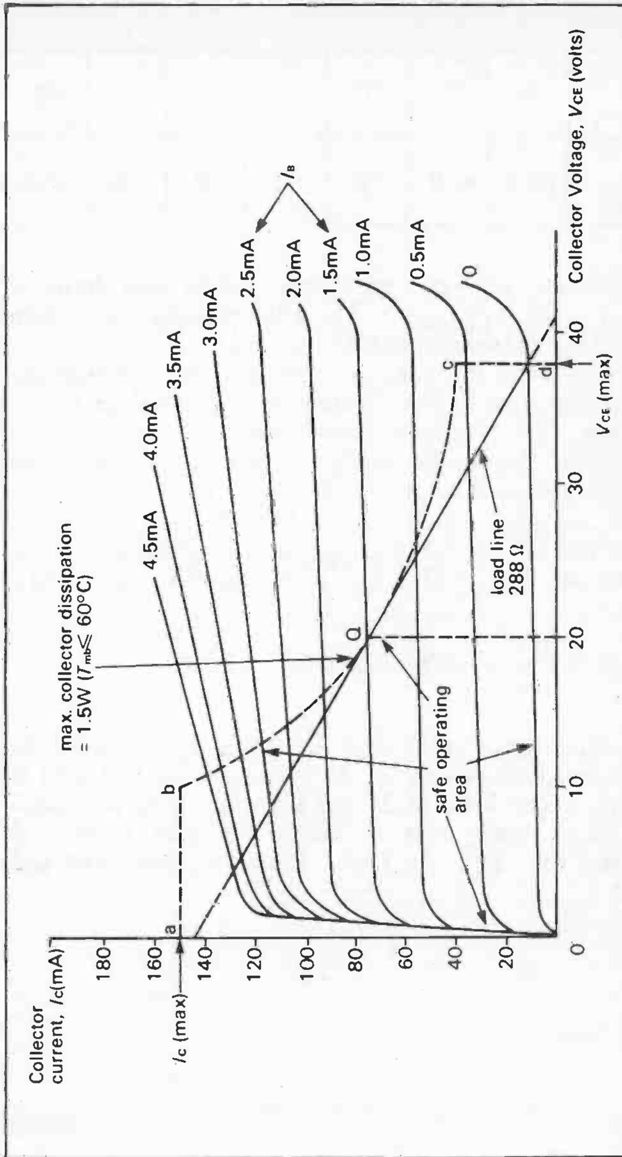


Fig. 5.13 Power transistor safe operating area

the following table defines the curve:

V_{CE} (volts)	38	35	30	25	20	15	10
I_C (mA)	39.5	42.9	50	60	75	100	150

Sections ab and cd are simply straight lines drawn at $I_{C(max)}$ and $V_{CE(max)}$. The safe operating area is then enclosed by the two axes and the line abcd.

If $V_{CC} = 20$ V the d.c. load line is in fact a vertical line drawn from $V_{CE} = 20$ V because there is no voltage drop in the load. For maximum power output the load line can theoretically touch the maximum dissipation curve, bc (although for practical reasons it may be preferable to work slightly within). Hence an a.c. load line can be drawn, passing through the 1.5 mA I_B curve at point Q, cutting the axes at $V_{CE} = 41.5$ V, $I_C = 144$ mA, i.e. having a value,

$$R = \frac{41.5}{144} \times 1000 = 288\ \Omega .$$

An input signal swing of 2 mA drives I_B from 0.5 to 2.5 mA and from the curves, V_{CE} swings from 31.3 to 11 V and I_C swings between 35 and 106 mA. These are peak–peak values, they must be halved for peak values and multiplied by $1/\sqrt{2}$ for r.m.s. Therefore power developed in load

$$= \frac{(V_{max} - V_{min})}{2\sqrt{2}} \times \frac{(I_{max} - I_{min})}{2\sqrt{2}}$$

i.e. load power

$$P = \frac{(V_{max} - V_{min})(I_{max} - I_{min})}{8} \qquad (5.43)$$

$$= \frac{(31.3 - 11)(106 - 35)}{8} = 180 \text{ mW}$$

(this is not necessarily the absolute maximum because the load line is fitted by eye, slight adjustment may bring some improvement).

Let n = output transformer turns ratio, then

$$\frac{R_p}{R_s} = n^2 \qquad \text{(Book 1, Sect.3.5)}$$

where R_p and R_s are the primary and secondary load resistances and since $R_s = 8\ \Omega$ and $R_p = 288\ \Omega$,

$$\frac{288}{8} = n^2 = 36 \quad \therefore n = 6$$

5.6.2 Conversion Efficiency

This is a figure of merit for conversion of d.c. power into signal power by a transistor. It is usually denoted by η (Greek, eta) where:

$$\eta = \frac{\text{signal power output}}{\text{d.c. power input}}$$

$$= \frac{(V_{max} - V_{min})(I_{max} - I_{min})}{8 V_q I_q} \qquad (5.44)$$

where V_q and I_q are the quiescent V_{CE} and I_C values. The expression is multiplied by 100 to present the result as a percentage.

EXAMPLE 5.17:
From Figure 5.13 what are the conversion efficiencies with input signal swings of 2 and 3 mA?

Input signal swing of 2 mA:

from Example 5.16, power into load = 180 mW, and because the load line touches the 1.5 W dissipation curve at the quiescent operating point Q, the d.c. power input = 1.5 W. Therefore from Equation (5.44):

$$\text{Conversion efficiency, } \eta = \frac{180 \times 10^{-3}}{1.5} \times 100 = 12\%$$

Input signal swing of 3 mA:

from the points of intersection of the load line with the $I_B = 0$ and $I_B = 3$ mA characteristics:

$$V_{max} = 38 \text{ V}, \quad V_{min} = 9 \text{ V},$$

$$I_{max} = 114 \text{ mA}, \quad I_{min} = 11 \text{ mA}$$

∴ Power into load

$$= \frac{(38 - 9)(114 - 11)}{8} \text{ mW} = 373 \text{ mW}$$

and

$$\eta = \frac{373 \times 10^{-3}}{1.5} \times 100\% = 25\%$$

which shows that swinging the transistor over a greater part of its load line increases the a.c. power output and hence the efficiency. At the limit when the peak value of the output voltage is equal to the supply voltage, V_{CC} and the peak output current is equal to the quiescent (d.c.) value, I_q, then:

$$V_{max} = 2V_{CC}, \quad V_{min} = 0,$$

$$I_{max} = 2I_q, \quad I_{min} = 0$$

(note that in this particular case of a load which creates no d.c. voltage drop, $V_{CC} = V_q$), thus

$$\eta = \frac{2I_q \times 2V_{CC}}{8 \times V_{CC}I_q} \times 100 = 50\%$$

so a Class A amplifier (collector current flows throughout the input cycle) can never be more than 50% efficient. The example demonstrates that efficiencies considerably lower are to be expected.

5.6.3 Distortion

The inevitable distortion arising from the crowding of the curves at the higher values of I_B in Figure 5.13 is brought into focus by plotting the dynamic current transfer characteristic for the particular load line. This is done in Figure 5.14 and the curvature is unmistakable. Assuming a steady bias of 2 mA, the bottom sine wave represents an input signal and the top wave the result in collector current swing.

When a pure wave is distorted, harmonics are generated and in Book 1, Section 5.2 there is the general conclusion that in an asymmetrical waveform ($i_1 \neq i_2$ as in Fig.5.14) it is most likely that even harmonics predominate. In the case of Figure 5.14 this is sufficiently true for us to estimate the second harmonic content only on the assumption that the third and above are negligible. If the base current is $I_B \sin \omega t$, then the general expression which takes account of the slope and non-linearity of the characteristic can be used (see Appendix A2.1.4.1), i.e.:

$$I_C = a(I_B \sin \omega t) + b(I_B \sin \omega t)^2 + c(I_B \sin \omega t)^3 + \ldots$$

where a, b, c etc. are constants depending on the shape of the curve. Taking the first two terms only and expanding using Equation (A2.42):

$$\begin{aligned} I_C &\approx aI_B \sin \omega t + \frac{bI_B{}^2}{2}(1 - \cos 2\omega t) \\ &\approx \frac{bI_B{}^2}{2} + aI_B \sin \omega t - \frac{bI_B{}^2}{2} \cos 2\omega t \end{aligned} \qquad (5.45)$$

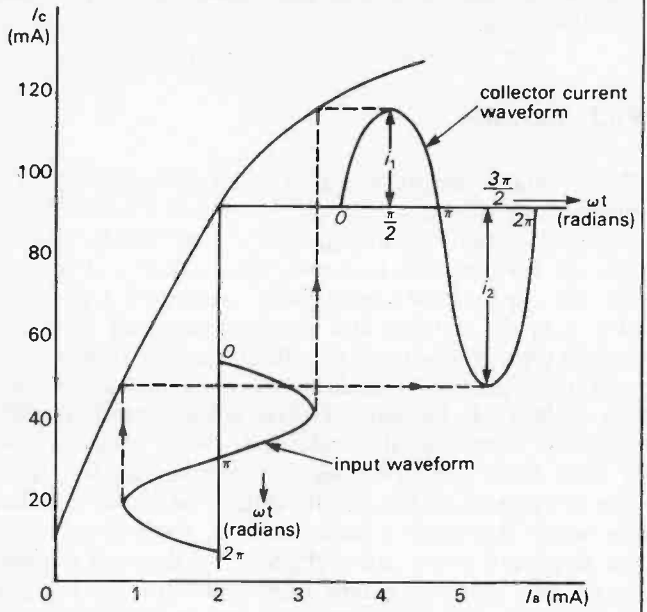


Fig. 5.14 Dynamic current transfer characteristic for Fig. 5.13 (288Ω load line)

$(bI_B{}^2)/2$ is independent of ω and is therefore a d.c. component which is to be expected because in fact a certain amount of rectification has taken place.

The magnitude of i_1 is given when the wave of I_C has moved through $\pi/2$ radians (90°), i.e. $\omega t = \pi/2$, hence $\sin \omega t = 1$, $\cos 2\omega t = -1$:

$$\therefore \quad i_1 = \frac{bI_B{}^2}{2} + aI_B + \frac{bI_B{}^2}{2} = aI_B + bI_B{}^2 .$$

Similarly, at $\omega t = 3\pi/2$, $\sin \omega t = -1$, $\cos 2\omega t = -1$ and if i_1 is +ve, i_2 is −ve ,

$$\therefore \quad -i_2 = \frac{bI_B{}^2}{2} - aI_B + \frac{bI_B{}^2}{2} = -aI_B + bI_B{}^2$$

$$\therefore \quad i_2 = aI_B - bI_B{}^2 .$$

Then:

$$i_1 + i_2 = 2aI_B , \text{ i.e. } \frac{i_1 + i_2}{2} = aI_B$$

which from Equation (5.45) is the amplitude of the fundamental. Also:

$$i_1 - i_2 = 2bI_B{}^2 , \text{ i.e. } \frac{i_1 - i_2}{4} = \frac{bI_B{}^2}{2}$$

which from Equation (5.45) is the amplitude of the second harmonic. Accordingly we now have the fundamental and second harmonic amplitudes expressed in terms of the collector waveform currents. The efficacy of the circuit is estimated by comparing the two to give the percentage distortion, i.e. second harmonic distortion is equal to:

$$\frac{\text{amplitude of second harmonic}}{\text{amplitude of fundamental}} \times 100\% .$$

In terms of i_1 and i_2 this becomes:

$$\frac{\dfrac{i_1 - i_2}{4}}{\dfrac{i_1 + i_2}{2}} \times 100\%$$

i.e. second harmonic distortion is equal to:

$$\frac{i_1 - i_2}{2(i_1 + i_2)} \times 100\% \qquad (5.46)$$

Should i_2 exceed i_1 (as in Fig.5.14) a negative result occurs. This can be ignored because all that matters is the *relationship* between them. For Figure 5.14 $i_1 = 116 - 92 = 24$ mA , $i_2 = 92 - 48 = 44$ mA , hence from Equation (5.46) the second harmonic distortion is equal to:

$$\frac{20}{2(68)} \times 100 = 14.7\%$$

generally an intolerable amount but then it must be admitted that for ease of illustration, Figure 5.14 has been deliberately drawn to make it so!

EXAMPLE 5.18:
Figure 5.7 shows the collector current waveform with a sinusoidal input to a certain transistor. The quiescent current is 26 mA and the swing is from 34 down to 16 mA. What is the percentage second harmonic distortion?

$$i_1 = 34 - 26 = 8 \text{ mA}$$

$$i_2 = 26 - 16 = 10 \text{ mA}$$

From Equation (5.46) percentage second harmonic distortion is equal to:

$$\frac{2}{2 \times 18} \times 100 = 5.6\% .$$

It now seems evident that as i_1 approaches i_2 so the fraction in Equation (5.46) gets smaller, eventually reaching zero when $i_1 = i_2$ and one might assume that no distortion is present. This is true with the second harmonic which is considered here because it is the most likely outcome from a transistor. With only the third harmonic present, i_1 is always equal to i_2 irrespective of the degree of distortion (as shown in Fig.5.10 of Book 1) and this analysis needs to go a step further.

Having examined distortion in a single-ended power output stage, it is perhaps proper to remind ourselves that a push-pull arrangement effects a considerable reduction. Negative feedback also reduces distortion and this is considered in Section 5.9.

5.7 WIDEBAND AMPLIFIERS

The basic factors which limit bandwidth are discussed in Section 5.5 so a wideband amplifier is simply one designed to reduce their effects. Such amplifiers are often called upon to transmit pulses (e.g. the synchronizing pulses in television receivers) but never can they do this with perfection because theoretically a rectangular waveform contains harmonics to infinity. So far we have been accustomed to looking at amplifier performance on a frequency basis but this fails to give a clear picture of how a pulse outline is changed. This Section therefore looks at some of the links between the frequency response of an amplifier and what happens to a transmitted pulse. The main deviations in pulse shape from the ideal appear later in Figure 6.8 and we first recall from Section 5.5.1.4 that the bandwidth of an amplifier is the range of frequencies between the lower 3 dB point, f_1 and the upper, f_2 .

5.7.1 Pulse Sag

Pulse sag is caused by a poor amplifier low-frequency response. In Figure 5.11(iii) at low frequencies, C_2 has little effect so the main frequency sensitive component is C_c , shunted on the two sides by R_L and R_2 . A pulse applied to the circuit quickly charges C_c but instead of the charge remaining throughout the duration of the pulse, it decays according to the time-constant, $\tau = C_c(R_L + R_2)$. Thus while the input pulse is at its maximum and steady, the output pulse level is falling. The output pulse sag [Fig.6.8(iv)] is therefore proportional to the pulse duration but inversely proportional to the time constant for the greater this is, the less the output voltage falls within a given time, i.e.:

$$\text{pulse sag} \propto \frac{t_p}{\tau}$$

where t_p is the pulse duration. Many approximations need to be made to produce a simple formula so the following proof has no pretence to accuracy. We consider a square wave because this has the simplification built in that

$$t_p = \frac{1}{2f}$$

where f is the wave frequency. Hence, provided that the sag is reasonably small and the pulse rise-time (Sect.5.7.2) is zero:

$$\text{percentage sag} \approx \frac{t_p}{\tau} \times 100 = \frac{100}{2f\tau},$$

or in terms of the components of Figure 5.11(iii):

$$\% \text{ sag} \approx \frac{100}{2f \times C_c(R_L + R_2)}$$

which from Equation (5.40) gives:

$$\% \text{ sag} \approx \frac{\pi f_1 \times 100}{f} \qquad (5.47)$$

EXAMPLE 5.19:
The maximum permissible sag of a square wave at the output

of a wideband amplifier is 4%. If the lower 3 dB frequency of the amplifier is 2 Hz, what is the lowest frequency which can be amplified?

$$f_1 = 2\text{ Hz} \qquad \%\,\text{sag} = 4$$

From Equation (5.47):

$$f = \frac{\pi \times f_1 \times 100}{\%\,\text{sag}} = \frac{\pi \times 2 \times 100}{4} = 157\text{ Hz}$$

5.7.2 Pulse Rise-Time

Fourier(A3) teaches us that waves other than sine comprise a fundamental plus harmonics (odd, even or both) to infinity and as the harmonic increases in number so it decreases in amplitude. For example, a symmetrical square wave is expressed by:

$$e = \frac{4E}{\pi}\left(\sin \omega t + \frac{\sin 3\omega t}{3} + \frac{\sin 5\,\omega t}{5} + \ldots\right) \quad (5.48)$$

where E is the maximum value and $\omega = 2\pi f$.

Now if the harmonics are progressively cut off from the highest numbers downwards, the square wave degenerates into a sine wave, in effect its vertical sides begin to slope inwards. Hence a fall in amplifier bandwidth increases pulse *rise-time* [Fig.6.8(iii)].

Again referring to Figure 5.11(iii), at high frequencies C_c can be neglected but C_2 is effective. Therefore the rise in v_0 when the leading edge of a pulse reaches the second stage is exponential and controlled mainly by the time constant,

$$\tau = \frac{R_L R_2}{R_L + R_2} \quad \text{and} \quad v_0 = V_p\,(1 - e^{(-t/\tau)})$$

where V_p is the maximum pulse level.

The rise-time occurs between $v_0 = 0.1V_p$ and $v_0 = 0.9V_p$ [Fig.6.8(iii)]. When

$$v_0 = 0.1V_p, \quad 0.1V_p = V_p(1 - e^{(-t/\tau)})$$

$$\therefore \quad 0.9 = e^{(-t/\tau)}$$

$$\therefore \quad \log_e 0.9 = -t, \quad \text{i.e. } t = 0.105\tau .$$

Similarly, when

$$v_0 = 0.9V_p \qquad t = 2.303\tau$$

Hence rise-time, $t_r = (2.303 - 0.105)\tau = 2.2\tau$.

Now, from Equation (5.41) the upper 3 dB frequency,

$$f_2 = \frac{1}{2\pi \times \tau}$$

and because for most wideband amplifiers the bandwidth and f_2 are of the same order:

$$f_2 \times t_r = \frac{2.2\tau}{2\pi \times \tau} = 0.35, \quad \text{i.e. } t_r \approx \frac{0.35}{f_2} \text{ secs} \qquad (5.49)$$

EXAMPLE 5.20:
On test with a 50 Hz square wave a wideband amplifier was found to produce 1.5% sag and a rise-time of 175 ns. Determine the approximate bandwidth.

$$f = 50 \text{ Hz}$$

$$t_r = 175 \times 10^{-9} \text{ s}$$

$$\text{Bandwidth} = (f_2 - f_1)$$

From Equation (5.47):

$$f_1 = \frac{f \times \% \text{ sag}}{\pi} = \frac{50 \times 1.5}{\pi} = 24 \text{ Hz}$$

From Equation (5.49):

$$f_2 = \frac{0.35}{t_r} = \frac{0.35}{175 \times 10^{-9}} = 2.0 \text{ MHz}$$

The amplifier bandwidth therefore extends from approximately 24 Hz to 2 MHz.

This shows that tests using square waves enable the high and low frequency performances of an amplifier to be assessed, hence also the bandwidth.

5.8 AMPLIFIER NOISE

Thermal noise, which is generated within any resistance is considered in Section 7.1. Such noise is naturally present in any amplifier and its value at the output depends mainly on the gain following the point of generation, thus the input circuit is usually the most troublesome. Thermal noise also arises in the base resistance of transistors and in addition *shot* noise is generated by the random movement of electrons across a potential barrier and is most noticeable in transistors in the base current. *Partition* noise adds to the above, it is due to the random division of the emitter current between base and collector.

5.9 NEGATIVE FEEDBACK

Negative feedback, available in exchange for amplifier gain, is of such benefit that the majority of amplifiers incorporate it in one or more of its many forms. In deriving the appropriate formulae less effort is required by assuming that the phase-change through the amplifier is exactly 180°, but we should not assume this having already produced Figure 5.12 in which the phase shift is 180° only over the middle range of frequencies. Accordingly we must start by accepting that the amplifier gain may need to be expressed in both modulus and argument (Sect.4.1) and that the feedback path may also be complex. We first recall that quoting an amplifier gain in the general form $A\angle\theta$ means that the output is A times the input and the two differ in phase by the angle θ.

5.9.1 The General Formulae

Consider an example of a feedback system as shown in Figure 5.15. The feedback network accepts the output voltage v_0 and injects the fraction $\beta\angle\phi$ of it in series with the input circuit to oppose the input signal. The angle ϕ allows for the fact that β could have a reactive component and hence cause a phase-shift in the feedback voltage. Then:

$$v_0 = A\angle\theta \times v_i$$

and v_0 at terminals 1 and 2 of the feedback network results in a voltage $\beta\angle\phi \times v_0$ at terminals 3 and 4. v_i is therefore the sum of v_s and $\beta\angle\phi \times v_0$. Hence

$$v_0 = A\angle\theta\,(v_s + \beta\angle\phi \times v_0)$$

$$= A \times v_s\angle\theta + \beta A \times v_0\underline{/\theta + \phi}$$

$$\therefore \quad v_0\,(1 - \beta A\underline{/\theta + \phi}) = A \times v_s\angle\theta$$

and gain with feedback,

$$A_f = \frac{v_0}{v_s} = \frac{A\angle\theta}{1 - \beta A\underline{/\theta + \phi}} \qquad (5.50)$$

and in the particular case where $\theta = 180°$ and $\phi = 0°$, since

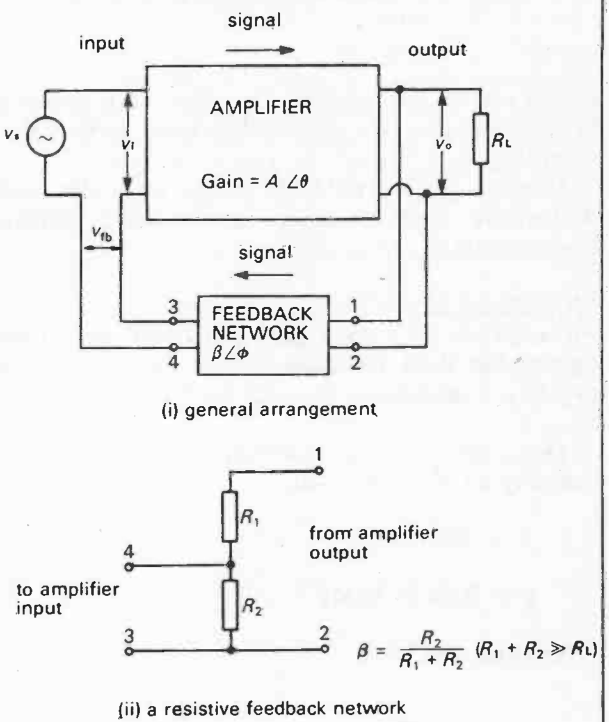


Fig. 5.15 A negative feedback system

$A \angle 180^\circ = -A$,

$$A_f = (-)\ \frac{A}{1 + \beta A} \tag{5.51}$$

the minus sign remaining as a reminder of the polarity reversal.

It is quite usual to meet Equation (5.51) in the form

$$\frac{A}{1-\beta A},$$

in this case β is made negative. The rule is simply that the denominator must be greater than unity so that the gain is *reduced.*

Usually β has no angle except when the feedback is deliberately used to shape the amplifier gain/frequency characteristic.

EXAMPLE 5.21:
An amplifier of voltage gain $800\angle 180^\circ$ at a certain frequency has 0.5% feedback applied via a resistive network ($\phi = 0^\circ$). Calculate the gain with feedback.

Obviously a case for Equation (5.51) but to improve our dexterity, (5.50) is also used.

$$A = 800\angle 180^\circ,$$

$$\beta = 0.5\% = 0.005$$

From Equation (5.51):

$$A_f = -\frac{A}{1+\beta A} = -\frac{800}{1+0.005 \times 800}$$

$$= \frac{-800}{5} = -160$$

From Equation (5.50):

$$A_f = \frac{A\angle\theta}{1-\beta A\underline{/\theta+\phi}} = \frac{800\angle 180^\circ}{1-0.005 \times 800\angle 180^\circ}$$

$$= \frac{800\angle 180^\circ}{1 - 4(\cos 180^\circ + j\sin 180^\circ)} = \frac{800\angle 180^\circ}{1 - 4(-1 + j0)}$$

$$= \frac{800\angle 180^\circ}{5 - j0} = \frac{800\angle 180^\circ}{5\angle 0^\circ} = 160\angle 180^\circ,$$

the two answers being effectively the same.

EXAMPLE 5.22:
At 20 Hz the amplifier of Figure 5.12 has a gain of 50.9 $\angle 214^\circ$. What is the gain if 1% of non-reactive feedback is added?

For this, only Equation (5.50) can be used:

$$A_f = \frac{A\angle\theta}{1 - \beta A\underline{/\theta + \phi}} = \frac{50.9\angle 214^\circ}{1 - 0.01 \times 50.9\angle 214^\circ}$$

$$= \frac{50.9\angle 214^\circ}{1 - 0.509(\cos 214^\circ + j\sin 214^\circ)}$$

$$= \frac{50.9\angle 214^\circ}{1 - 0.509(-0.83 - j0.56)} = \frac{50.9\angle 214^\circ}{1.42 + j0.29}$$

$$= \frac{50.9\angle 214^\circ}{\sqrt{1.42^2 + j0.29^2}\ \tan^{-1}\dfrac{0.29}{1.42}}$$

$$= \frac{50.9\angle 214^\circ}{1.45\angle 11.5^\circ} = 35.1\angle 202.5^\circ$$

(or expressed as output voltage lagging on input voltage, 35.1 $\angle 157.5^\circ$).

It is instructive to repeat the above exercise over the full frequency range of the amplifier. This is done in Figure 5.16 and two other features of negative feedback become apparent:

(i) the bandwidth is increased – in this example from approximately (13 Hz → 156 kHz) to (8 Hz → 250 kHz), a useful way of minimizing the effects of RC coupling and inherent shunt capacitances;

(ii) the shift of phase away from 180° at low and high frequencies is less.

5.9.2 Gain Stability

Semiconductor components usually have wide manufacturing tolerances, for example h_{fe} for a commonly used transistor, between 240 and 500. A basic amplifier therefore could have a gain very different from that intended. Also once constructed, the amplifier gain may vary with both time and temperature due to changes within the components themselves. Negative feedback reduces these effects for taking Equation (5.51), if $\beta A \gg 1$, then the gain with feedback:

$$A_f \approx -\frac{1}{\beta} \tag{5.52}$$

The gain is therefore almost independent of anything other than the value of β. This is most acceptable but it does mean that the basic amplifier must have high gain so that a useful gain remains when feedback is applied. This is one principle on which the operational amplifier rests.

EXAMPLE 5.23:
An amplifier has an initial gain of 5000. 2% negative feedback is applied. What is the percentage reduction in gain if changing a certain transistor reduces the initial gain by 10%?

$$A_1 = 5000$$

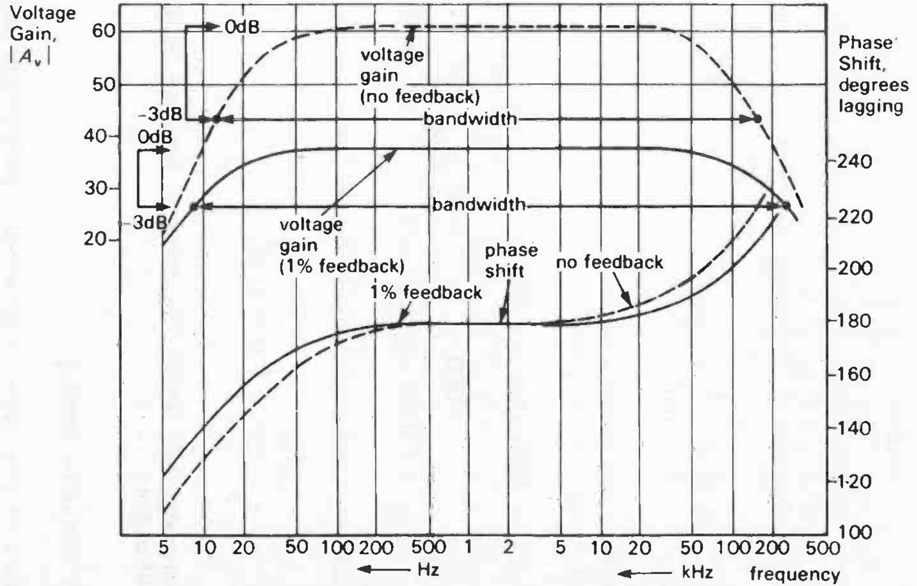


Fig. 5.16 Effect of adding 1% negative feedback to amplifier of Fig. 5.12

$$\beta = 2\% = 0.02$$

$$\therefore \quad \beta A_1 = 100$$

From Equation (5.51):

$$A_{1f} = \frac{5000}{1 + 100} = 49.5$$

[also using Equation (5.52) because $\beta A_1 \gg 1$,

$$A_{1f} \approx \frac{1}{\beta} = \frac{1}{0.02} = 50,$$

the difference is therefore minimal].

After change of transistor,

$$A_2 = 5000 \times 0.9 = 4500$$

$$\therefore \quad A_{2f} = \frac{4500}{1 + (0.02 \times 4500)} = \frac{4500}{91} = 49.45$$

Therefore, percentage change in gain is equal to:

$$\frac{49.5 - 49.45}{49.5} \times 100 = 0.1\%$$

evidence beyond doubt that negative feedback stabilizes amplifier gain.

5.9.3 Amplifier Stability

Figure 5.12 demonstrates that amplifier phase-change deviates from the central 180° line at both low and high frequencies. As an example, at 200 kHz the phase is already some 50° away from the 180. It may be possible therefore that a shift of phase will ultimately be sufficient to change negative feedback into positive. Recourse is sometimes made to the simple shunt capacitor across the output to reduce the h.f. gain but this may have undesirable repercussions elsewhere. H. Nyquist, an American scientist, analysed the general stability problem and although in its entirety his work is beyond the scope of this book, we may find it rewarding to come to terms with the basic idea.

Consider the input circuit of Figure 5.15 which for convenience is repeated at (i) on Figure 5.17. For negative feedback v_i must be less than v_s and for positive feedback, greater. The idea is to determine the relationship between v_i and v_s from phasor diagrams. At medium frequencies, if β has zero angle, the phasor diagram is as in Figure 5.17(ii), a straight line in which the phasor for v_s is hidden because it lies along the same line as v_{fb} and v_i. v_i is drawn as the reference phasor and is given a value of 1 unit. Then:

$$v_{fb} = \beta v_o \qquad \text{but } v_o = A \angle\theta \times v_i$$

$$\therefore \quad v_{fb} = \beta A \angle\theta \times v_i \text{, i.e. } \beta A \angle\theta \text{ units.}$$

Again using the amplifier of Figure 5.12 as an example, because $A = 61$ and $\beta = 0.01$, the phasor v_{fb} in (ii) of the Figure has a length of 0.61 units. v_s can therefore be measured on the diagram or alternatively, calculated. Evidently in this diagram $v_s > v_i$ so the feedback is negative, that is v_{fb} opposes v_s in such a way that the result, v_i is something less.

As frequency changes, so does the phasor diagram and we know from Figure 5.12 that considering the h.f. response for example, as frequency rises A ($|A_v|$ in the Figure) falls and θ increases. Thus at 200 kHz Figure 5.17(iii) shows that βA has fallen to 0.38 and θ increased to 232°. The remaining phasor diagrams are self-explanatory and are a sufficient indication that by holding v_i constant and plotting $\beta A \angle\theta$ ($= v_{fb}$) at various frequencies the ratio between v_i and v_s can be judged. In fact, because our only concern is that $v_s > v_i$, all that is required is to draw a circle using v_i as radius [as shown at (iv) for f = 10 MHz] and if v_{fb} enters the circle at any frequency, then $v_s < v_i$ and the feedback is positive.

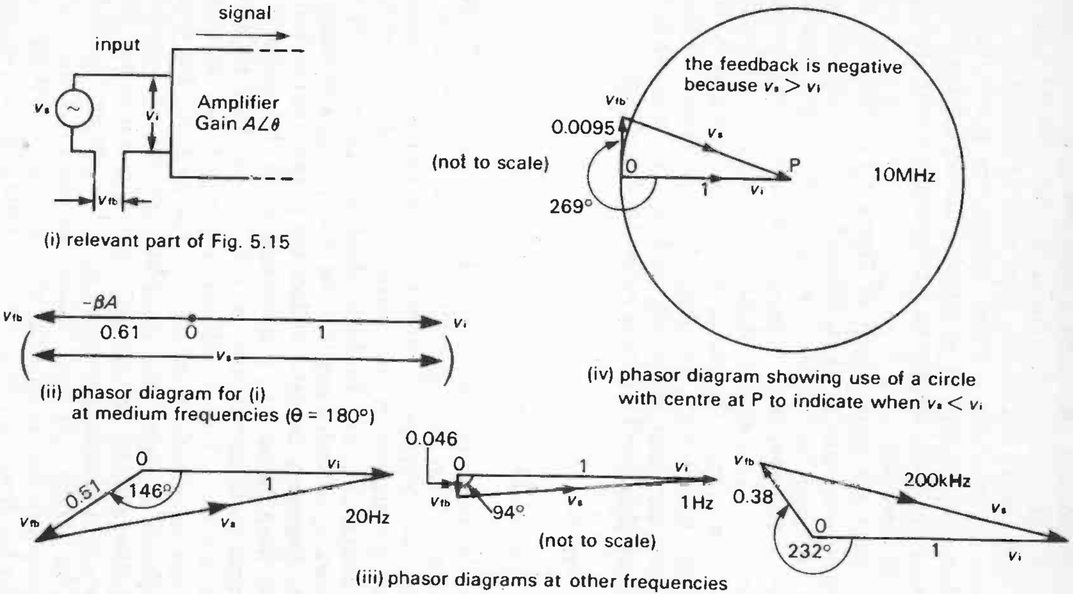


Fig. 5.17 Phasor diagrams for input circuit of Fig. 5.15

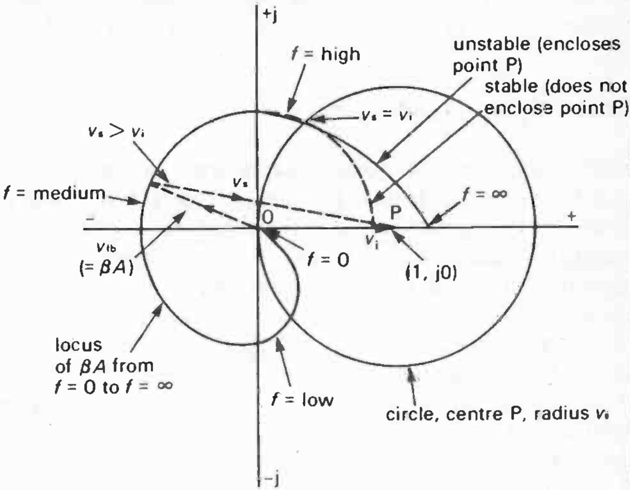


Fig. 5.18 Basic features of Nyquist principle

However, positive feedback does not necessarily mean that the amplifier is unstable, other factors must be taken into consideration and this is the main part of Nyquist's work. His conclusions on stability in practical terms show that only when the locus of βA *encloses* the point P (see Fig.5.18) is the system unstable. In the Figure an amplifier locus of βA is shown resulting in instability (the dotted curve shows a stable condition). The curve of βA is not that of the amplifier of Figure 5.12 because it is evident from Figure 5.17 that at 0 and ∞ Hz, θ is approaching 90° and 270° respectively so that βA does not enter the circle. This is true for all single-stage transistor amplifiers, the maximum shift of θ from 180° due to capacitances within the amplifier does not exceed 90°. However, when several stages are in cascade, the phase shifts add so the Nyquist criterion comes into use. Modern design, especially in I.C.'s avoids the use of coupling capacitors but it cannot entirely eliminate transistor input capacitance, thus stability problems are usually experienced only at high frequencies and on Figure 5.18 for example, only the high frequency end of the locus of βA would need to be plotted.

5.9.4 Distortion and Noise

Both distortion and noise voltages arise within an amplifier and in a negative feedback system a fraction of these is fed back to the input to be transmitted in antiphase to the originals. Some cancellation of the unwanted voltages is therefore to be expected but what we need to know is the extent of cancellation when the original gain is restored or alternatively when the input voltage, v_s is raised to compensate. By representing the interfering voltages as arising from a voltage generator v_{dn} in the output circuit and using the relationships already developed, it is straightforward mathematically to show that, using the general formula for $\theta = 180°$,

$$v_0 = A_v \times v_s + \frac{v_{dn}}{1 + \beta A_v}$$

(where A_v is the voltage gain), showing that the distortion and noise components have been reduced by

$$\frac{1}{1 + \beta A_v},$$

that is:

$$\text{distortion and/or noise with feedback}$$

$$= \frac{\text{distortion and/or noise without feedback}}{1 + \beta A_v} \qquad (5.53)$$

If therefore an amplifier has an initial signal output voltage v_0 with distortion and noise output voltages v_d and v_n, then with negative feedback applied and the gain or input signal increased to restore v_0 :

total output voltage with feedback

$$= v_0 + \frac{v_d + v_n}{1 + \beta A_v} \tag{5.54}$$

EXAMPLE 5.24:
An amplifier of voltage gain 1000 is found to have a harmonic distortion of 5%. The distortion must be reduced to 1% but the gain maintained. What value of β should be used and what should the gain of an additional stage be?

From Equation (5.53):

$$0.01 = \frac{0.05}{1 + 1000\,\beta},$$

from which $\beta = 0.004$. Then

$$A_{vf} = \frac{A_v}{1 + \beta A_v} = \frac{1000}{1 + (0.004 \times 1000)}$$

$$= \frac{1000}{5} = 200$$

i.e. gain of additional stage $= \dfrac{1000}{200} = 5$.

(or more simply: $\dfrac{A_v}{A_{vf}} = 1 + \beta A_v = 5$).

5.10 SUMMARY OF KEY FORMULAE

A	= amplifier gain
e	= emitter
b	= base
c	= collector
c-e	= common emitter
c-b	= common base
c-c	= common collector
f_1	= amplifier lower 3 dB frequency
f_2	= amplifier upper 3 dB frequency
F	= fraction of current fed back
i_i	= input current
i_o	= output current
I	= current
R_L	= amplifier load resistance
R_S	= source resistance
v_i	= input voltage
v_o	= output voltage
v_d	= distortion voltage
v_n	= noise voltage
V_q, I_q	= quiescent values
β	= feedback fraction
θ	= amplifier phase angle
ϕ	= feedback network phase angle

QUANTITY	FORMULA	UNIT	SECTION
Transistor current equation	$I_e = I_b + I_c$	A	5
Transistor voltage equation	$V_{ce} = V_{cb} + V_{be}$	V	5
Collector leakage current (c-e)	$I_c = h_{FE} I_b + I_{CEO}$	A	5.4.1
Collector leakage current (c-b)	$I_c = h_{FB} I_e + I_{CBO}$	A	5.4.1
Total leakage current (c-e)	$I_{CEO} = (h_{FE} + 1) I_{CBO}$	A	5.4.1
Hybrid Parameters [see Fig.5.2(iii)]			
Basic network equations	$v_i = h_i i_i + h_r v_o$	V	
	$i_o = h_f i_i + h_o v_o$	A	
Input resistance	$h_i = \dfrac{v_i}{i_i}$	Ω	5.1.2
Current gain	$h_f = \dfrac{i_o}{i_i}$		5.1.2

QUANTITY	FORMULA	UNIT	SECTION
Output conductance	$h_o = \dfrac{i_o}{v_o}$	S	5.1.2
Voltage feedback ratio	$h_r = \dfrac{v_i}{v_o}$		5.1.2
Amplifier current gain (c-e)	$K_i = \dfrac{h_{fe}}{1 + h_{oe}R_L} = h_{fe}$ (approx)		5.1.3.1
Amplifier voltage gain (c-e)	$K_v = -\dfrac{h_{fe}R_L}{h_{ie}(1+h_{oe}R_L) - h_{fe}h_{re}R_L}$		5.1.3.4
	$= -\dfrac{h_{fe}R_L}{h_{ie}}$ (approx)		
Input resistance (c-e)	$R_{in} = h_{ie} - \dfrac{h_{fe}h_{re}R_L}{1 + h_{oe}R_L} \approx h_{ie}$	Ω	5.1.3.2

QUANTITY	FORMULA	UNIT	SECTION
Output resistance (c-e)	$R_{out} = \dfrac{h_{ie} + R_s}{h_{oe}(h_{ie} + R_s) - h_{fe}h_{re}}$ $\approx \dfrac{1}{h_{oe}}$	Ω	5.1.3.3
Power gain (c-e)	$K_p = \dfrac{K_i{}^2 R_L}{R_{in}} \approx \dfrac{h_{fe}{}^2 R_L}{h_{ie}}$		5.1.3.4
Conversion c-e to c-b	$h_{ib} = \dfrac{h_{ie}}{1 + h_{fe}}$ $h_{fb} = -\dfrac{h_{fe}}{1 + h_{fe}}$		5.1.4
	$h_{ob} = \dfrac{hoe}{1 + h_{fe}}$ $h_{rb} = \dfrac{h_{ie}h_{oe}}{1 + h_{fe}}$		5.1.4

QUANTITY	FORMULA	UNIT	SECTION
Conversion c-e to c-c	$h_{ic} = h_{ie} \quad h_{fc} = -(1 + h_{fe})$		5.1.4
	$h_{oc} = h_{oe} \quad h_{rc} = \dfrac{1}{1 + h_{re}}$		5.1.4
Formulae for approximate design – see Table 5.1 (Sect.5.1.4)			
y-Parameters [see Fig.5.5(ii)]			
Voltage gain	$K_v = -\dfrac{y_{fs}}{y_{os} + Y_L} \approx \dfrac{-y_{fs}}{Y_L}$		5.2
,, ,,	$K_v = -g_m \dfrac{r_{os}R_L}{r_{os} + R_L}$ $\approx -g_m R_L$		5.2

QUANTITY	FORMULA	UNIT	SECTION
Amplifier Response (c-e) [see Fig.5.11(iii)			
Voltage gain at low frequencies	$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{(R_L + R_2 - j/\omega C_c)}$		5.5.1.1
Voltage gain at medium frequencies	$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{(R_L + R_2)}$		5.5.1.2
Voltage gain at high frequencies	$A_v = -\frac{h_{fe}}{h_{ie}} \times \frac{R_L R_2}{R_L + R_2 + j\omega C_2 R_L R_2}$		5.5.1.3
Lower 3 dB frequency	$f_1 = \frac{1}{2\pi C_c (R_L + R_2)}$	Hz	5.5.1.4
Upper 3 dB frequency	$f_2 = \frac{R_L + R_2}{2\pi C_2 R_L R_2}$	Hz	5.5.1.4

QUANTITY	FORMULA	UNIT	SECTION
Power Amplifiers			
Max. power dissipation of transistor	$P_{tot} = V_{CE}I_C + V_{BE}I_B$	W	5.6.1
Power in load	$P = \frac{(V_{max} - V_{min})(I_{max} - I_{min})}{8}$	W	5.6.1
Conversion efficiency	$\eta = \frac{(V_{max} - V_{min})(I_{max} - I_{min})}{8V_qI_q}$		5.6.2
Second harmonic distortion (see Fig.5.1.4)	$d = \frac{i_1 - i_2}{2(i_1 + i_2)} \times 100$	%	5.6.3
Amplifier Pulse Transmission			
Pulse sag	$\approx \frac{\pi f_1 \times 100}{f}$	%	5.7.1
Pulse rise-time	$t_r = \frac{0.35}{f_2}$	s	5.7.2

QUANTITY	FORMULA	UNIT	SECTION
Amplifier Stability (d.c.)			
Stability factor	$S = \frac{1}{1 + h_{FE}F}$		5.4.2
Negative Feedback			
Gain with feedback	$A_f = \frac{A\angle\theta}{1 - \beta A\underline{/\theta+\phi}} \approx -\frac{A}{1+\beta A}$		5.9.1
Gain when $\beta A \gg 1$	$A_f = -\frac{1}{\beta}$		5.9.2
Total output voltage with feedback	$V_f = v_0 + \frac{v_d + v_n}{1 + \beta A_v}$	V	5.9.4

6. SIGNAL GENERATION AND PROCESSING

The meaning of the word "signal" is precise, an "intelligible sign conveying information" and in electronics the sign is given mainly by changes in voltage or current. In this Chapter however, the word may be used rather loosely, not through any disregard of the purity of language, but because it is a convenient one for labelling any waveform whether it actually conveys information or not.

In the first part of the Chapter oscillators and multivibrators are considered, in the second part is an analysis of some of the processes by which signals are combined or modified.

6.1 SINE WAVE OSCILLATORS

Feedback over an amplifier is considered in Section 5.9, usually negative for if instead positive, the danger arises of instability or oscillation. An oscillator deliberately employs positive feedback, as shown in Figure 6.1, using the feedback as its amplifier input. The *frequency control network* maintains the output at that required or may be variable so that a range of frequencies is available. Then, assuming no loss in the network:

$$v_2 = A_v \times v_1 \quad \text{and} \quad v_1 = \beta v_2$$

$$\therefore \quad v_1 = \beta A_v \times v_1$$

$$\therefore \quad v_1(1 - \beta A_v) = 0$$

if the oscillator is running, v_1 cannot be 0, hence $1 - \beta A_v = 0$

$$\therefore \quad \beta A_v = 1 \tag{6.1}$$

and because of the likelihood of both A_v and β being complex, more precisely:

$$\beta \angle\phi \times A_v \angle\theta = 1 \angle 0°$$

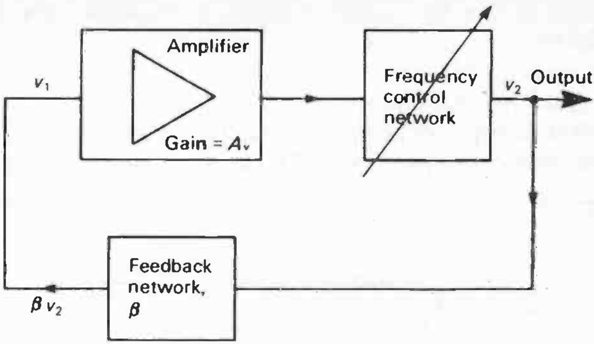


Fig. 6.1 Principle of oscillator

(where β and A_V are the moduli), i.e.

$$\beta A_V = 1\,, \quad (\phi + \theta) = 0° \qquad (6.2)$$

There are many different types of circuit which achieve this condition. In this Section (6.1) all oscillators generate sine waves.

6.1.1 LC Oscillators

This particular type has as its frequency control element a resonant circuit consisting of an inductor with a capacitor of appropriate value for resonance at the frequency required:

$$f_{osc} = \frac{1}{2\pi\sqrt{LC}} \text{ Hz} \qquad (6.3)$$

where L is the inductance in henries and C is the capacitance in farads.

With miniaturization and IC's designers tend to avoid the use of inductors hence the use of this type is diminishing in favour of the RC oscillator which is discussed later (Sect. 6.1.2).

6.1.1.1 Tuned Collector

This is perhaps one of the easier circuits to understand because each of the components of Figure 6.1 is relatively clear. A

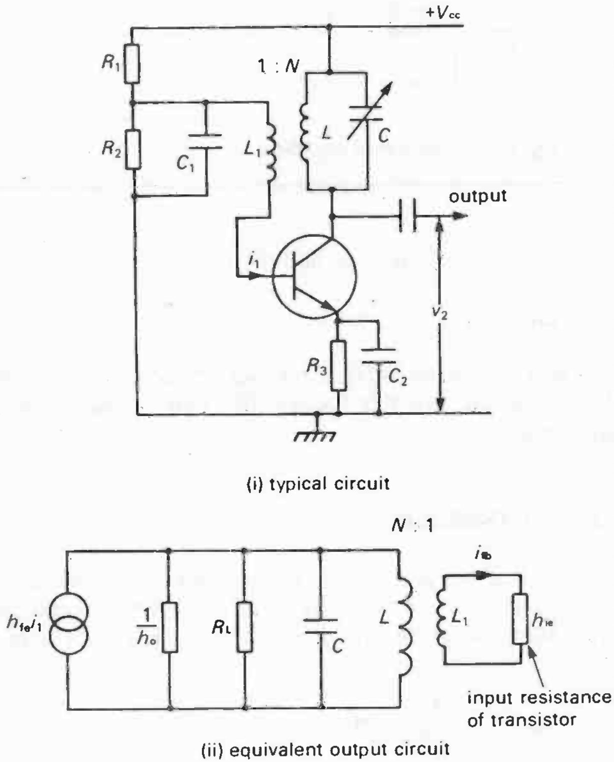


Fig. 6.2 An LC oscillator

typical circuit is shown in Figure 6.2(i). The frequency control network is LC resonating according to Equation (6.3), C may be variable or fixed. This and the output circuit form the transistor load. The transistor is the central component of the amplifier, deriving its bias from R_1, R_2 and R_3 (Sect. 5.4). The feedback network consists of the winding L_1 mutually coupled to L and connected to the transistor base and via C_1 to the common line. The connexions to L_1 are such as to provide positive feedback. C_2 decouples R_3 to avoid a.c. feedback.

Switching on gives rise to a small pulse of current into the capacitances, LC selects its own resonant frequency and injects a small voltage into the base which is amplified so that the oscillatory current in LC grows until it reaches its steady state. This is accomplished within a very few cycles.

If the base current is i_1 then the collector circuit can be considered as containing a constant current generator $h_{fe}i_1$ (see Fig.5.2 or for revision, Sect.5.1). As shown in Figure 6.2(ii) the generator is connected to several components in parallel. The analysis is more manageable if h_0 is neglected which is permissible because generally it is of low value. Also h_{re} has little effect on the conclusions so we neglect that too. However we cannot ignore the impedance reflected into L from the input resistance of the transistor, h_{ie}, this is N^2h_{ie}. Then total resistance across generator

$$= \frac{R_L \times N^2 h_{ie}}{R_L + N^2 h_{ie}}$$

and voltage across L

$$= h_{fe}i_1 \left(\frac{R_L N^2 h_{ie}}{R_L + N^2 h_{ie}}\right)$$

therefore voltage across L_1

$$= \frac{h_{fe}i_1}{N}\left(\frac{R_L N^2 h_{ie}}{R_L + N^2 h_{ie}}\right)$$

and current, i_{fb} flowing into base circuit, i.e. in h_{ie}

$$= \frac{h_{fe}i_1}{Nh_{ie}} \left(\frac{R_L N^2 h_{ie}}{R_L + N^2 h_{ie}} \right)$$

$$= h_{fe}i_1 \left(\frac{NR_L}{R_L + N^2 h_{ie}} \right)$$

for oscillation this must be equal to i_1 (for $\beta A = 1$)

$$\therefore \quad h_{fe}i_1 \left(\frac{NR_L}{R_L + N^2 h_{ie}} \right) = i_1$$

$$\therefore \quad h_{fe} \times NR_L = R_L + N^2 h_{ie}$$

$$\therefore \quad h_{fe} = \frac{R_L + N^2 h_{ie}}{NR_L} \quad \left(\text{or } \frac{N \times h_{ie}}{R_L} + \frac{1}{N} \right) \tag{6.4}$$

so if h_{fe} is less than this value, oscillation is not maintained. The formula is slightly approximate because no account is taken of h_o and h_{re}, nevertheless it is sufficient to indicate the requirements.

EXAMPLE 6.1:
A tuned-collector oscillator at 250 kHz has an air-cored transformer with a main winding [L on Fig.6.2(i)] of 1980 turns. How many turns are required on winding L_1 if the transistor parameters are $h_{ie} = 2.5\ \text{k}\Omega$, h_{fe} not less than 30? The load R_L is 3.0 kΩ.

From Equation (6.4):

$$h_{ie} \times N^2 - h_{fe}R_L \,.\, N + R_L = 0$$

i.e. $$2500\,N^2 - 90{,}000N + 3000 = 0$$

$$\therefore \quad 5N^2 - 180N + 6 = 0\,.$$

Solving by use of Equation (A2.3):

$$N = \frac{180 \pm \sqrt{180^2 - (4 \times 5 \times 6)}}{10} = \frac{180 \pm 179.67}{10}$$

$$= 36 \text{ or } 0.033 ,$$

clearly the second result is inadmissible hence:

$$\text{Number of turns required} = \frac{1980}{36} = 55 .$$

6.1.1.2 Colpitts and Hartley

These are two resonant-circuit oscillators named after their American engineer inventors. In both types the frequency control network is basically LC but the method of tapping the circuit for an emitter connexion differs as shown in Figure 6.3. The biassing arrangements are omitted in the Figure so that the principle of operation is clear. C_3 and C_4 are both of low reactance at the oscillation frequency, From the point of view of the emitter, if at resonance one side of the LC circuit is positive, the other side is negative and so the signal returned from collector to base changes phase by 180°. Together with the 180° change through the transistor, the complete loop has a phase change of 360° (or 0°), hence with sufficient amplifier gain, oscillation is maintained. Ignoring the slight shunting effect of the amplifier circuit on the tuned circuit, allows the general frequency of oscillation formula to apply modified by the split of the capacitance (Colpitts) or inductance (Hartley):

Colpitts:

$$f_{osc} = \frac{1}{2\pi \sqrt{L\left(\frac{C_1 C_2}{C_1 + C_2}\right)}} \qquad (6.5)$$

$$= \frac{1}{2\pi}\sqrt{\frac{C_1 + C_2}{LC_1C_2}} \text{ Hz}$$

Hartley:

$$f_{osc} = \frac{1}{2\pi\sqrt{(L_1 + L_2)\,C}} \text{ Hz} \qquad (6.6)$$

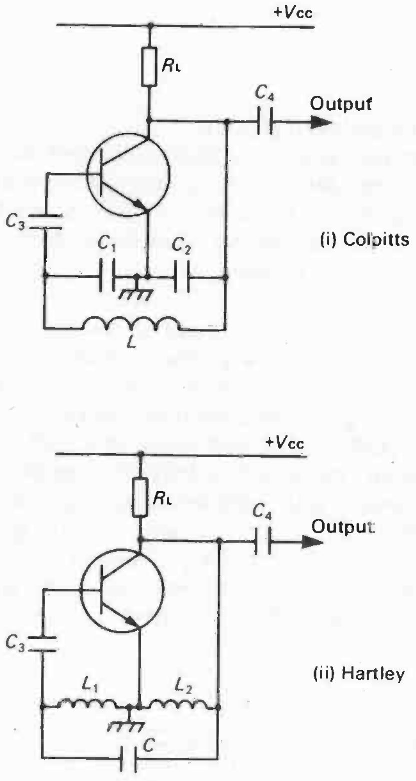


Fig. 6.3 Colpitts and Hartley oscillators

for zero mutual coupling between L_1 and L_2. Should coupling exist, say of M henries, then the combined inductance of L_1 and L_2 is $(L_1 + L_2 \pm 2M)$ and

Hartley:

$$f_{osc} = \frac{1}{2\pi\sqrt{(L_1 + L_2 \pm 2M)C}} \text{ Hz} \qquad (6.7)$$

[inductors in series opposition are unlikely to be used so the minus sign in the denominator can be omitted].

6.1.2 RC Oscillators

The resistance-capacitance oscillator makes its feedback positive by interposing an RC network between output and input of the amplifier. The phase-shift through the network is arranged to complement that through the amplifier for an overall of 360° or 0°. Such an oscillator is sinusoidal and as previously mentioned has the advantage that no inductor is required.

6.1.2.1 Phase-Shift

A single transistor amplifier has a phase-shift of 180° hence the feedback network of Figure 6.1 must add or subtract 180° to satisfy the angle requirement of Equation (6.2). A phase-shift of up to 90° arises in a simple RC network as shown in Figure 6.4(i) but as the angle approaches 90° the ratio between X_C and R becomes unmanageable and in fact at exactly 90° it is infinite, hence not realizable in practice. Accordingly two simple RC networks in series are incapable of shifting the phase by 180° so three must be used of about 60° shift each. We cannot however design one single network for 60° phase-shift, and then connect three similar ones in tandem because each network has an effect on the other two. The analysis therefore needs Kirchhoff's Laws and as is so often found, some assumptions are helpful to avoid the mathematics getting out of hand. Hence in Figure 6.4(ii) it is assumed that the input resistance of the amplifier is low enough and the output resistance high enough to terminate the network suitably. In practice this is hardly so, accordingly our conclusions can only be taken as approximate. Labelling the network currents i_1 to i_4 as shown:

Mesh I: $i_2(2R + jX_c) - i_1R - i_3R = 0$ (i)

Mesh II: $i_3(2R + jX_c) - i_2R - i_4R = 0$ (ii)

Mesh III: $i_4(R + jX_c) - i_3R = 0$ (iii)

By eliminating i_2 and i_3 the magnitude of i_4 in terms of i_1 can be obtained thus:

multiply (ii) by $\dfrac{(2R + jX_c)}{R}$ and add to (i) to give:

$$-i_1R - i_3R + \frac{i_3(2R + jX_c)^2}{R} \qquad \text{(iv)}$$

$$-i_4(2R + jX_c) = 0$$

then from (iii):

$$i_3 = \frac{i_4(R + jX_c)}{R}$$

and substituting for i_3 in (iv) gives:

$$-i_1R - i_4(R + jX_c) + \frac{i_4(R + jX_c)(2R + jX_c)^2}{R^2}$$

$$-i_4(2R + jX_c) = 0$$

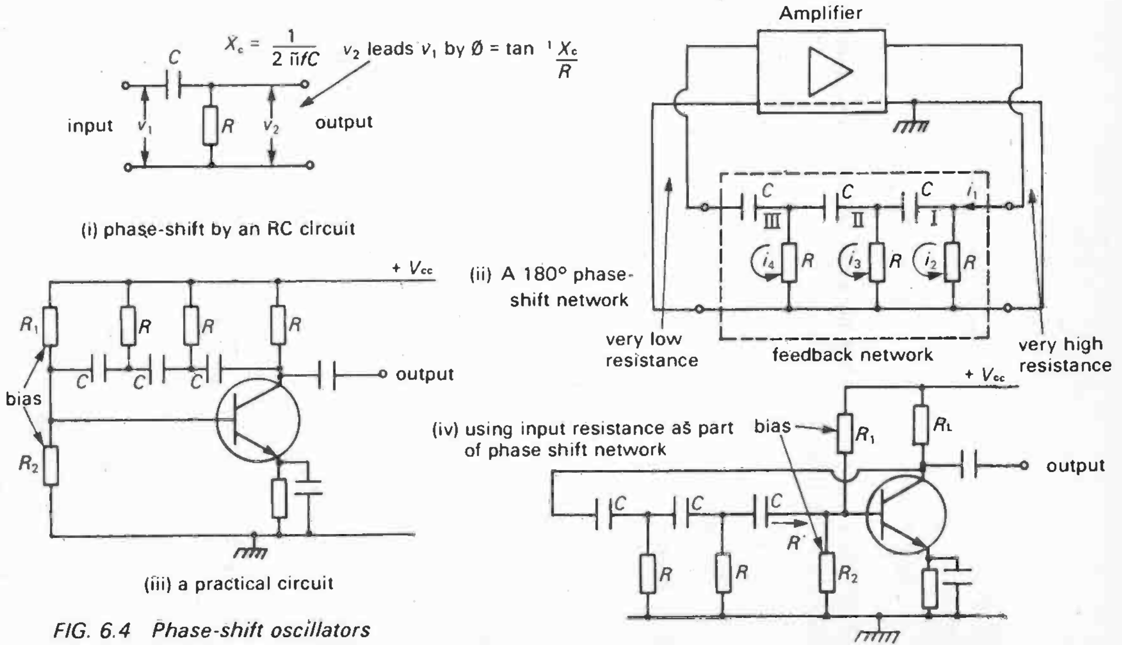


FIG. 6.4 Phase-shift oscillators

reducing to:

$$i_4 = \frac{i_1 R^3}{R^3 - 5RX_c^2 + j(6R^2 X_c - X_c^3)} \qquad (6.8)$$

Now the j term disappears when the phase shift is 180°, hence:

$$6R^2 X_c - X_c^3 = 0 \quad \therefore X_c = \sqrt{6} \times R \qquad (6.9)$$

which gives the relationship between the two components at the oscillation frequency. Also:

$$\frac{1}{2\pi f_{osc} C} = \sqrt{6} \times R \,,$$

from which

$$C = \frac{0.065}{f_{osc} \times R} \text{ Farads} \qquad (6.10)$$

and

$$f_{osc} = \frac{1}{2\pi \times \sqrt{6} \times RC} \text{ or } \frac{0.065}{RC} \text{ Hz} \qquad (6.11)$$

Also with 180° phase-shift the amplifier current gain, A_i follows from Equation (6.8):

$$A_i = \frac{i_1}{i_4} = \frac{R^3 - 5RX_c^2}{R^3}$$

and since

$$X_c^2 = 6R^2$$

$$A_i = \frac{R^3 - 30R^3}{R^3} = -29 \qquad (6.12)$$

which is therefore the minimum value to sustain oscillation.

An example of a practical amplifier is given in Figure 6.4(iii). Some circuits reverse the feedback network, using the input resistance of the transistor as the final resistor. In this case the transistor dictates the value of R to be used and C is calculated accordingly as in the next example. Figure 6.4(iv) shows a typical circuit.

EXAMPLE 6.2:
An RC oscillator as in Figure 6.4(iv) is to be designed for a frequency output of 433 Hz. The input resistance of the transistor is 1500 Ω. What value capacitors should be used?

From Equation (6.9):

$$\text{reactance of capacitors, } X_c = \sqrt{6} \times R = \sqrt{6} \times 1500$$

$$= 3674\ \Omega$$

$$\therefore \quad C = \frac{1}{2\pi f_{osc} \times X_c}\ \text{F} = \frac{10^6}{2\pi \times 433 \times 3674}\ \mu\text{F} = 0.1\ \mu\text{F}$$

[or directly from Equation (6.10)

$$C = \frac{0.065 \times 10^6}{433 \times 1500} = 0.1\ \mu\text{F}\ ].$$

6.1.2.2 Wien Bridge

So named after its inventor[(A3)] the bridge which is shown in Figure 6.5(i) is particularly useful for the measurement of frequency because it balances at one frequency only. Note that the input and detector may be interchanged as in (ii). The bridge principle is used in an oscillator in which the series and parallel RC networks constitute the frequency control as shown in (iii). The network is capable of zero phase-shift at a single frequency at which therefore the associated amplifier must also have zero or 360° phase-shift. For this a single stage is insufficient, two or more are required.

Consider the equivalent network shown in Figure 6.5(iv) where Z_s and Z_p represent the impedances of the series and parallel combinations. Then:

$$Z_s = R + jX_c$$

$$Z_p = \frac{jRX_c}{R + jX_c}$$

and

$$X_c = -\frac{1}{\omega C}$$

Also

$$\beta = \frac{Z_p}{Z_s + Z_p} = \frac{\dfrac{jRX_c}{R + jX_c}}{R + jX_c + \dfrac{jRX_c}{R + jX_c}}$$

$$= \frac{jRX_c}{(R + jX_c)^2 + jRX_c} = \frac{jRX_c}{(R^2 - X_c^2) + 3jRX_c}$$

rationalizing (Sect.4.2.2.1):

$$\beta = \frac{jRX_c(R^2 - X^2) + 3R^2X_c^2}{(R^2 - X_c^2)^2 + 9R^2X_c^2}$$

and for zero phase shift the j terms equate to zero, hence:

$$RX_c(R^2 - X_c^2) = 0$$

$$\therefore \quad R^2 - X_c^2 = 0, \quad \text{i.e. } R = X_c$$

$$\therefore \quad R = \frac{1}{2\pi f_{osc} C}$$

and

$$f_{osc} = \frac{1}{2\pi RC} \tag{6.13}$$

and also when the j terms are zero:

$$\beta = \frac{3R^2 X_c^{\,2}}{(R^2 - X_c^{\,2})^2 + 9R^2 X_c^{\,2}}$$

and since

$$R = X_c, \quad (R^2 - X_c^{\,2})^2 = 0$$

$$\therefore \quad \beta = \frac{3R^2 X_c^{\,2}}{9R^2 X_c^{\,2}} = \frac{1}{3} \tag{6.14}$$

hence from Equation (6.1):

$$\beta A_v = 1 \quad \therefore A_v = 3 \tag{6.15}$$

this is the minimum value, something slightly greater is usually required. Circuits based on the alternative bridge arrangement of Figure 6.5(ii) are also used.

To keep the mathematical solution reasonably manageable, we have considered the network to be one in which the resistors in the two sections have equal values, similarly with the capacitors. By further analysis it can be shown that with different values for the components the equation for the frequency of oscillation becomes:

$$f_{osc} = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}} \text{ Hz} \tag{6.16}$$

where, for example, $R_1 C_1$ form the series network and

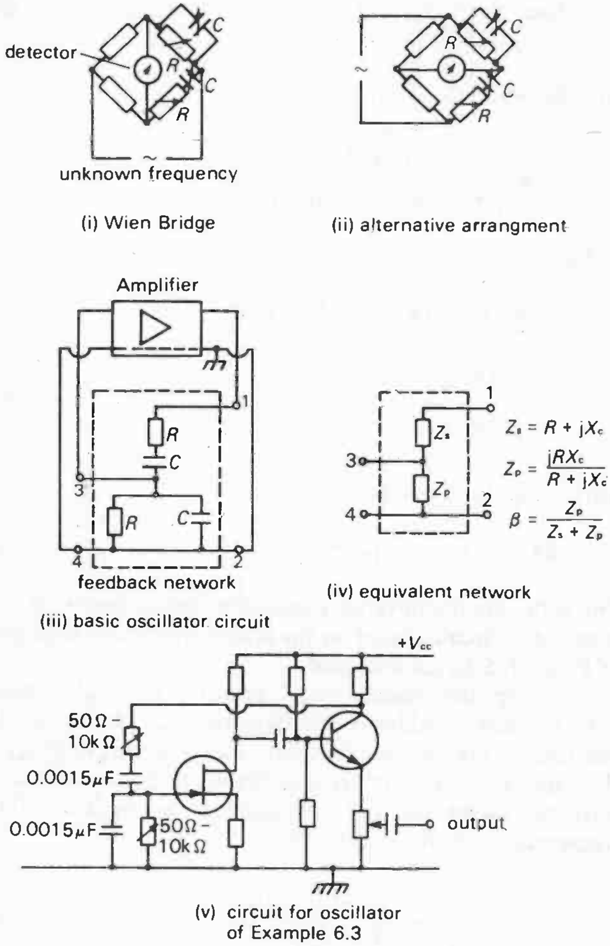


Fig. 6.5 Wien bridge oscillator

$R_2 C_2$ form the parallel network. The amplifier gain required under these conditions remains at 3.

EXAMPLE 6.3:
A Wien bridge oscillator is to be designed for a frequency range of 11 kHz – 2 MHz. Preferred value capacitors of 0.0015 μF are to be used with variable resistors. What is the range for the resistors?

From Equation (6.13):

$$R = \frac{1}{2\pi f_{osc} C}$$

At 11.0 kHz

$$R = \frac{10^6}{2\pi \times 11 \times 10^3 \times 0.0015} = 9646\ \Omega$$

At 2.0 MHz

$$R = \frac{10^6}{2\pi \times 2 \times 10^6 \times 0.0015} = 53.05\ \Omega$$

A suitable range for R is therefore 50 Ω – 10 kΩ, this covers the frequency range with a little to spare. The elements of a typical practical circuit are added to Figure 6.5 at (v). The first stage employs a common source FET (Sect.5.2) so that there is little shunting of the feedback network. The emitter resistors are not decoupled so that the overall gain is kept low (but not lower than 3).

6.1.3 Operational Amplifier Oscillators

The *operational amplifier* which abounds in plenty in integrated form is ideal as the basis of any of the types of oscillator mentioned above. Its versatility arises mainly from the fact that it has two inputs which result in outputs which are either in-phase (input marked +) or 180° out-of-phase (input marked −). Such an amplifier has high gain which is reduced to the required value by negative feedback and it is especially suited to the Wien oscillator, a typical diagram of which is given in Figure 6.6.

Equation (5.52) indicates that with a very high gain amplifier with negative feedback, the gain is approximately $1/\beta$ so because in the Figure the negative feedback is supplied via the resistor chain R_1R_2, then the fraction of the output fed back,

$$\beta = \frac{R_2}{R_1 + R_2} .$$

Accordingly the gain

$$A = \frac{R_1 + R_2}{R_2}$$

so in this particular case to reduce it to 3:

$$\frac{R_1 + R_2}{R_2} = 3 \qquad \therefore R_1 = 2R_2 .$$

The feedback from the frequency control network is connected to the + input terminal. The phase-shift from input to output of the amplifier is zero and at the oscillation frequency the phase-shift from output to input is also zero.

6.2 SAWTOOTH WAVE GENERATION

The *sawtooth* waveform, aptly named because of its shape, is particularly known for its use in building up television and oscilloscope *rasters* (the patterns of scanning lines traced on the tube). In these the spot travels linearly across the tube but flies back more quickly. Figure 6.7(i) shows the waveform

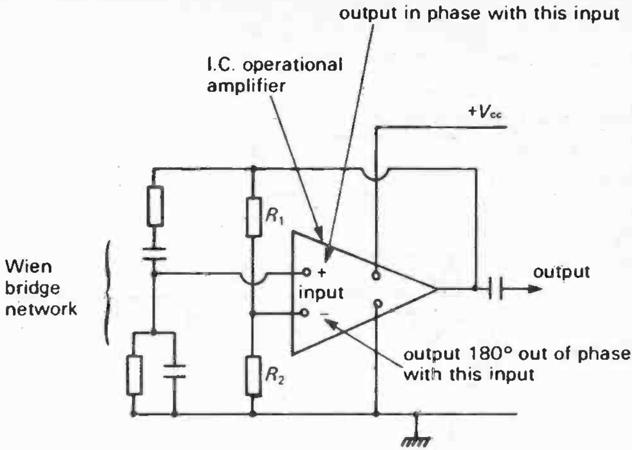


Fig. 6.6 Wien bridge oscillator embodying I.C. operational amplifier

with its linear rise known as the *ramp* from t_0 to t_1 and the fast fall to the starting potential from t_1 to t_2. The cycle is completed at t_2. In Book 1 (Sect.5.2) it is shown that the wave comprises all harmonics, both odd and even.

The ramp is usually obtained from a charging capacitor. Because the charging is exponential the ramp is not linear but by working over a small section only of the curve, sufficient linearity for many purposes can be obtained. The basic principle of operation is illustrated in Figure 6.7(ii) in which C charges through R until a negative-going pulse is applied to the base of the transistor (in this case, p-n-p) whereupon the transistor conducts and takes its collector current mainly from C. There is a partial discharge of C therefore throughout the duration of the input pulse. This circuit is *driven* by the supply of input pulses but one which needs no drive and runs continuously is illustrated in Figure 6.7(iii). A voltage applied to a neon lamp has little effect

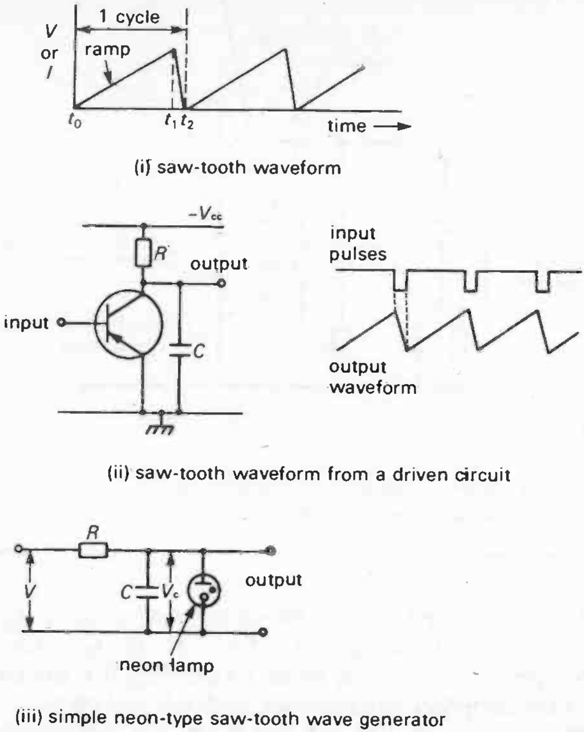
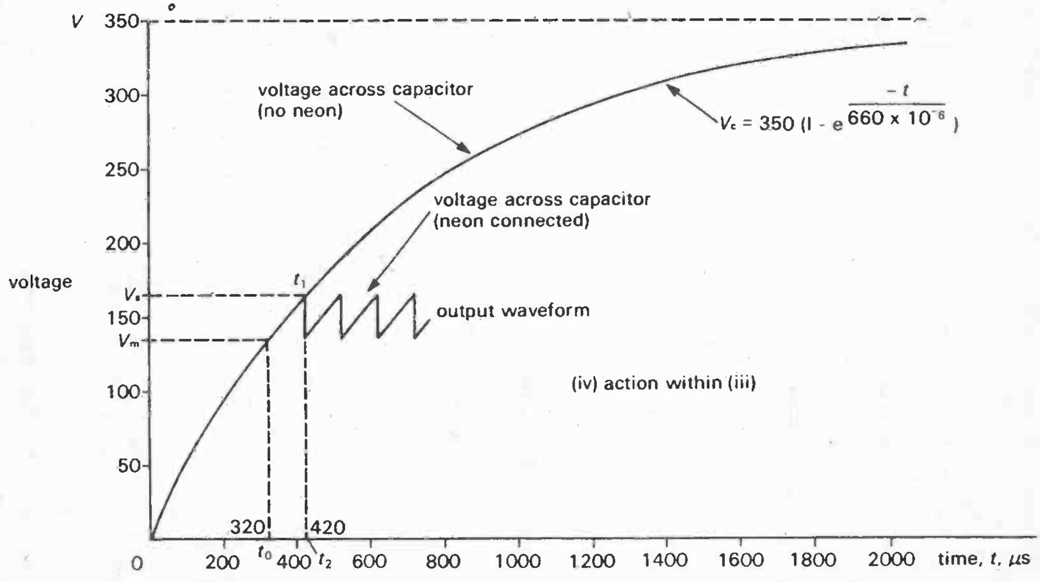


Fig. 6.7 Saw-tooth wave generation

unless it is at or above the value required for ionizing the gas whereupon the gas glows and its resistance falls. This is the *striking* voltage, V_s. The ionization continues if V_s is next reduced but not below a certain *maintaining* value, V_m for below this the gas reverts to its unionized, high resistance state. Generally V_s will be between 80 and 200 V and V_m up to around 180 V but not lower than 60 V. The action within the circuit of (iii) of the Figure is illustrated by (iv). Ignore the figures on the drawing at present.



V
350
300
250
200
150
100
50
0
voltage
V_e
V_m
voltage across capacitor
(no neon)
$V_c = 350\,(1 - e^{\frac{-t}{660 \times 10^{-6}}})$
voltage across capacitor
(neon connected)
t_1
output waveform
(iv) action within (iii)
320
420
t_0
t_2
200
400
600
800
1000
1200
1400
1600
1800
2000
time, t, μs

(Fig. 6.7 contd)

With no neon connected, the application of V to the series combination of R and C results in an exponential curve of voltage across the capacitor as time progresses. With the neon connected however the voltage is prevented from exceeding V_s because at this the low resistance of the neon quickly discharges C, the voltage of which falls until reaching V_m when the neon "cuts off". The charging then recommences so producing the sawtooth waveform. Let V_c represent the voltage across the capacitor, i.e. the output voltage of Figure 6.7(iii). Then:

$$V_c = V(1 - e^{(-t/CR)})$$

and if V_m is reached at time t_0,

$$V_m = V(1 - e^{(-t_0/CR)})$$

$$\therefore \quad V - V_m = V \times e^{(-t_0/CR)}$$

and V_s is reached at time t_1,

$$V - V_s = V \times e^{(-t_1/CR)}.$$

Then

$$\frac{V - V_m}{V - V_s} = \frac{e^{(-t_0/CR)}}{e^{(-t_1/CR)}} = \frac{e^{(t_1/CR)}}{e^{(t_0/CR)}} = e^{(t_1 - t_0)/CR)}$$

The only way to handle this is by taking natural logarithms of both sides, i.e.

$$\log_e \left(\frac{V - V_m}{V - V_s}\right) = \frac{t_1 - t_0}{CR}$$

and if we simplify the reasoning by considering that the discharge time is zero, then $t_1 = t_2$ [Fig.6.7(iv)] and $t_1 - t_0$ is the period of the waveform which is also the reciprocal of the frequency, hence:

$$f_{osc} = \frac{1}{CR \log_e \left(\dfrac{V - V_m}{V - V_s} \right)} \quad \text{Hz} \qquad (6.17)$$

EXAMPLE 6.4:
A neon lamp has striking and maintaining voltages of 165 and 135 respectively. It is to be used in a 10 kHz sawtooth generator circuit as in Figure 6.7(iii). If the supply voltage is 350 and the capacitor is chosen as 33 nF, what value of resistance is required?

$$f_{osc} = 10^4 \text{ Hz},$$

$$V = 350 \text{ V},$$

$$V_s = 165 \text{ V},$$

$$V_m = 135 \text{ V},$$

$$C = 33 \times 10^{-9} \text{F},$$

$$R = ?$$

From Equation (6.17):

$$R = \frac{1}{f_{osc} \times C \times \log_e \left(\dfrac{V - V_m}{V - V_s} \right)}$$

$$= \frac{1}{10^4 \times 33 \times 10^{-9} \log_e(215/185)}$$

$$= \frac{10^5}{33 \times 0.1503} \approx 20 \text{ k}\Omega .$$

The graph of Figure 6.7(iv) is in fact drawn to these values (CR = 660 μs). t_0 occurs 320 μs after V is applied, t_1 after 420 μs. Hence the time of one cycle (the period, T) $= t_1 - t_0 = 100\ \mu$s and

$$f_{osc} = \frac{1}{T} = \frac{1}{100 \times 10^{-6}} = 10\ \text{kHz}$$

showing that the circuit can also be analysed graphically. The graph also illustrates the technique of using a relatively high value for V so that a small section only of the exponential curve is employed to obtain a fairly linear ramp.

More sophisticated methods of discharging the capacitor can be envisaged but the basic principle developed above is unchanged.

6.3 PULSE GENERATION

A *pulse* is formed when a voltage or current changes in value and after a period of time reverts to its original value. Ideally a *rectangular pulse* has a shape as shown typically in Figure 6.8(i). Several pulses in succession are called a "waveform" although there is little resemblance to our usual conception of a wave. Unless existing in an environment of infinite bandwidth a pulse cannot have the ideal shape so we label the deviations as follows:

Pulse duration (or *width* or *length*) is perhaps most frequently stated by the time interval between its two half-amplitude points as in (ii) of the Figure, i.e. $t_2 - t_1$.

Pulse rise-time is that taken for the voltage or current to rise from 10% to 90% of its maximum value as shown in (iii). The *fall-time* is similarly illustrated.

Pulse sag is illustrated in (iv) and is usually measured as a percentage, i.e.

$$\%\ \text{sag} = \frac{V_{max} - V}{V_{max}} \times 100\ .$$

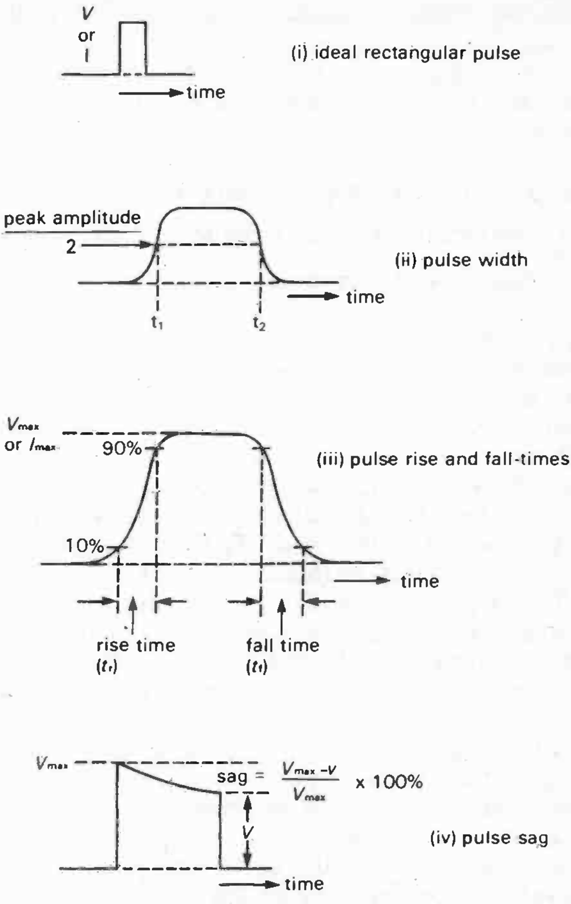


Fig. 6.8 Pulse terminology

6.3.1 Multivibrators

A *multivibrator* is basically a two-transistor circuit capable of generating pulses. Its general forms are shown in Figure 6.9(i) to (iii), the only differences being in the feedback networks from the collector of each transistor to the base of its partner. The form of feedback determines the multivibrator type as follows:

(i) both networks resistive – *bistable*

(ii) one network resistive, the other capacitive – *monostable*

(iii) both networks capacitive – *astable*.

6.3.1.1 Bistable

When either transistor is "on" (i.e. the base potential is such that saturation collector current flows), its low collector potential ensures that the other transistor is "off" (minimum collector current). This is a reciprocal arrangement because the OFF transistor high collector potential also ensures that its partner is held ON. The circuit *locks* or *latches* with T_1 ON , T_2 OFF or alternatively with T_2 ON , T_1 OFF . It is said to have two stable states (*bistable*, bi = 2).

To change over between the two states an external signal or *trigger* is applied, for example to one of the bases. The applied potential must be such as to swing the transistor into its opposite state so causing the same to happen to its partner. The change-over is very rapid, each transistor driving the other, thus the trigger need only be a pulse of short duration. To change the circuit back again a second pulse is applied to the other transistor. Thus when driven by a series of pulses, square waveforms appear at the two collectors. Because an input pulse can "flip" the circuit over while a second one causes it to "flop" back, the bistable is known in the computer world as a *flip-flop*. It has considerable application for by locking on to the polarity last received it "remembers" it and so is the basic cell of many memory systems.

6.3.1.2 Monostable

If one of the feedback networks is a series capacitor instead of a resistor as in (ii) of the Figure, one d.c. connexion is lost, hence T_1 cannot be kept OFF, the resistor R ensuring that it always reverts to ON with T_2 therefore OFF. This is the single stable state (*monostable*, mono = 1) to which the circuit always returns. With the n-p-n transistors shown, a negative triggering pulse applied to T_1 cuts it OFF, hence T_2 ON with a rapid collector potential fall which via C reinforces the negative potential on T_1 base. There now follows an exponential discharge of C through R and the base potential of T_1 becomes progressively less negative until eventually it is able to switch back to ON and the circuit reverts to its stable state. The time taken for this to happen which is also that of the output pulse duration naturally depends on the type of transistors in use, on the supply voltage and on R and C. To obtain some sort of relationship it is generally assumed that T_1 conducts when the voltage across C has fallen to half. Thus if V is the initial voltage across C on triggering, the end of the pulse duration occurs at $V/2$, hence if t = pulse duration:

$$\frac{V}{2} = V \times e^{(-t/CR)} \quad \therefore \; e^{(-t/CR)} = 0.5$$

$$\therefore \quad -\frac{t}{CR} = \log_e 0.5 = -0.693$$

$$\therefore \quad t = 0.693\,CR \tag{6.18}$$

EXAMPLE 6.5:

A monostable multivibrator circuit as in Figure 6.9(ii) operates with a supply voltage of 10 and the transistors have a d.c. current gain of 50 with a saturation collector current of 4 mA. Calculate suitable values for C and R if the output pulse duration is to be 0.25 ms.

$$I_b = \frac{4 \text{ mA}}{50} = 80\,\mu\text{A}$$

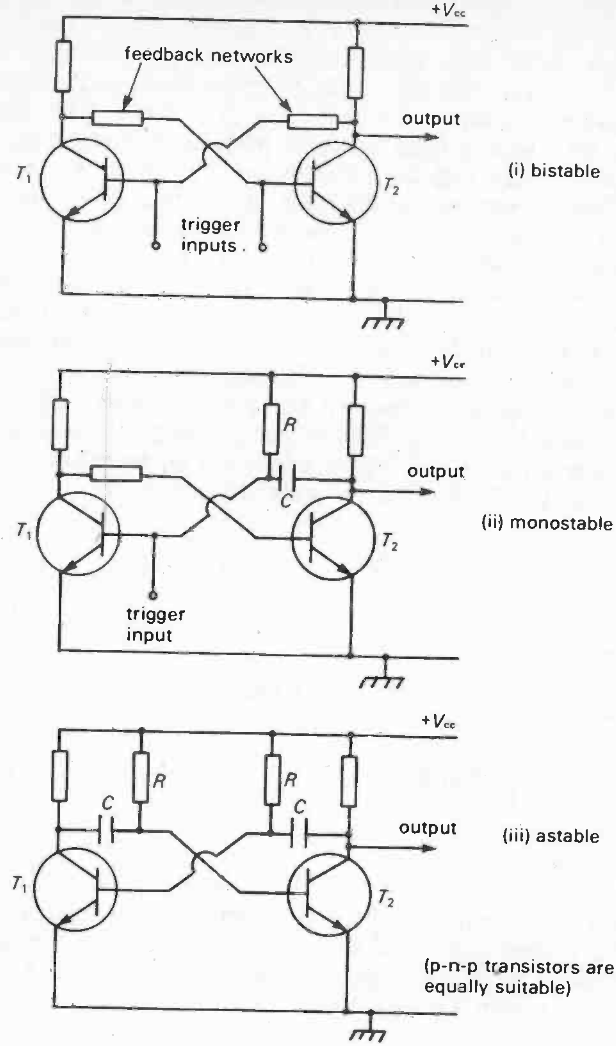
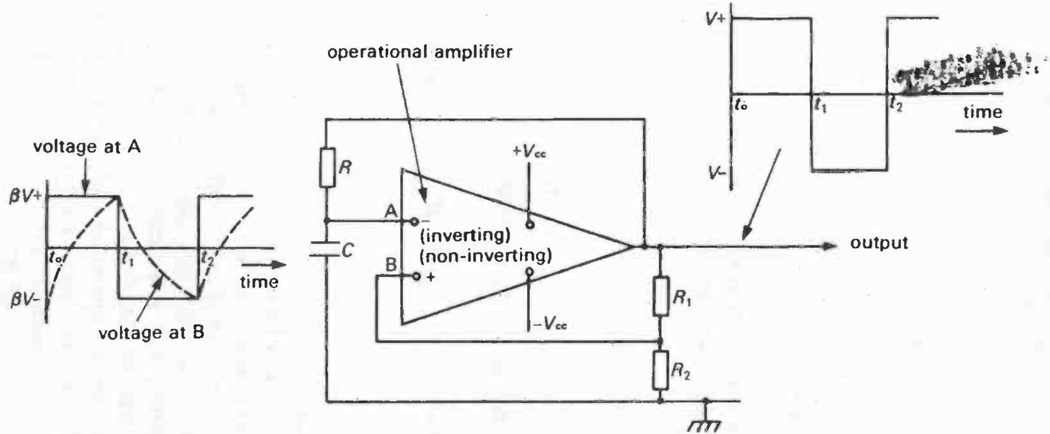


Fig. 6.9 Basic forms of multivibrators



(iv) astable embodying operational amplifier

Fig. 6.9 (contd)

and to provide $I_b = 80\ \mu A$,

$$R = \frac{(V_{cc} - V_{be})}{I_b}$$

and allowing 0.7 V for V_{be},

$$R = \frac{10 - 0.7}{80 \times 10^{-6}} = 116\ k\Omega .$$

This is the maximum value of R for saturation, to be on the safe side and use a preferred value, $R = 100\ k\Omega$ might be preferable.

From Equation (6.18):

$$C = \frac{t}{0.693R} = \frac{0.25 \times 10^{-3}}{0.693 \times 10^5} = 3.6\ nF$$

Suitable values for C and R are therefore 3.6 nF and 100 kΩ.

6.3.1.3 Astable

This is a particularly interesting circuit because it can function as a square-wave generator with no requirement of triggering. The feedback networks are both capacitors as in Figure 6.9(iii). The circuit therefore cannot latch in either state but switches regularly and symmetrically between the two. It is therefore classed as *astable* (a = not) and is said to be *free-running*. The circuit operation follows from the explanation above for the monostable. Then from Equation (6.18), if each transistor conducts for a time $t = 0.693\ CR$ the period, T of the output waveform is:

$$T = 2t = 2 \times 0.693CR = 1.386CR \qquad (6.19)$$

and the square-wave frequency:

$$f_{osc} = \frac{1}{T} = \frac{1}{1.386CR} \quad \text{or} \quad \frac{0.722}{CR} \qquad (6.20)$$

EXAMPLE 6.6:
An astable multivibrator as in Figure 6.9(iii) has a symmetrical square-wave output at 2400 Hz. The capacitors are 0.01 μF. What resistors are required?

From Equation (6.20):

$$R = \frac{1}{f_{osc} \times 1.386 \times C} = \frac{10^6}{2400 \times 1.386 \times 0.01}$$

$$= 30 \text{ k}\Omega .$$

6.3.1.4 Astable (Op. Amp.)
An IC approach to the generation of a square waveform embodies an operational amplifier. The elements of such a multivibrator are given in Figure 6.9(iv). Both positive and negative feedback are applied to the amplifier. With sufficient positive feedback alone the amplifier is driven into saturation with its output terminal at, say, V^+ , with negative feedback only the output swings to V^- . With both feedback paths connected the amplifier saturates according to which of the two types is greater.

Positive feedback is applied from the resistance chain, R_1R_2 , hence with

$$\beta = \frac{R_2}{R_1 + R_2}$$

and this on its own drives the amplifier to V^+ as shown at t_0 in the Figure. The same output voltage is applied via resistor R to terminal A but it cannot rise immediately to full value because of current drain into capacitor C so actually rises according to the time-constant CR. Meanwhile the output voltage remains fairly constant. When however the voltage at A rises to a value greater than that at B (i.e. > βV^+), the op. amp. immediately swings to opposite saturation, (at t_1) and as with the 2-transistor astable circuit the voltage changes are such as to reinforce the swing to make the switch to V^- very rapid. C next begins to charge in the opposite direction and the output reverts to V^+ as soon as A becomes more negative than B (at t_2).

From this it is evident that switching over takes place when the voltage across C is either βV^+ or βV^- and from Equation (6.17) it follows that in this particular case the times between the two conditions are:

$$t_1 - t_0 = CR \log_e \frac{V^+ - \beta V^-}{V^+ - \beta V^+}$$

and

$$t_2 - t_1 = CR \log_e \frac{V^- - \beta V^+}{V^- - \beta V^-}$$

For a symmetrical waveform $t_1 - t_0 = t_2 - t_1$ and V^+ has the same magnitude as V^- hence βV^+ has the same magnitude as βV^- so that $V^+ - \beta V^-$ can be rewritten as $V^+ + \beta V^+ = V^+(1 + \beta)$ and similarly $V^- - \beta V^+$ becomes $V^- + \beta V^- = V^-(1 + \beta)$. Hence the waveform period:

$$T = (t_1 - t_0) + (t_2 - t_1)$$

$$= CR \left\{ \log_e \frac{V^+(1 + \beta)}{V^+(1 - \beta)} + \log_e \frac{V^-(1 + \beta)}{V^-(1 - \beta)} \right\} \quad \text{secs}$$

so from Equation (A2.12):

$$T = CR \log_e \left\{ \frac{(1 + \beta)}{(1 - \beta)} \right\}^2$$

and next from Equation (A2.14):

$$T = 2CR \log_e \frac{(1+\beta)}{(1-\beta)} \text{ s} \qquad (6.21)$$

and substituting

$$\beta = \frac{R_2}{R_1 + R_2}$$

$$T = 2CR \log_e \left(\frac{R_1 + 2R_2}{R_1} \right)$$

$$= 2CR \log_e \left(1 + \frac{2R_2}{R_1} \right) \text{ s} \qquad (6.22)$$

and f_{osc} is the reciprocal of this. When $R_1 = R_2$ then:

$$T = 2CR \log_e 3 = 2.2CR \text{ s} \qquad (6.23)$$

EXAMPLE 6.7:
A square-wave generator circuit as in Figure 6.9(iv) is to cover the range 100 – 4500 Hz. $R_1 = R_2 = 4.7$ kΩ . R is a variable 1 kΩ – 47 kΩ. What value of C is suitable?

Since $R_1 = R_2$, Equation (6.23) applies.

$$f_{osc} = \frac{1}{2.2CR} \text{ Hz} \qquad \therefore C = \frac{10^6}{2.2R \times f_{osc}} \ \mu\text{F}$$

When $R = 1$ kΩ ,

$$C = \frac{10^6}{2.2 \times 1000 \times 4500} = 0.1 \ \mu\text{F} .$$

When $R = 47\ \text{k}\Omega$,

$$C = \frac{10^6}{2.2 \times 47000 \times 100} = 0.097\ \mu\text{F} .$$

A 0.1 μF capacitor therefore appears to be suitable, it is correct at the lower end of R but produces a frequency of

$$\frac{10^6}{2.2 \times 0.1 \times 47000} = 97\ \text{Hz}$$

at the upper end. It has the advantage of being a preferred value and the range required is only slightly exceeded.

6.4 ADDITION OF TWO SINE WAVES

Before becoming involved in the more extensive subject of modulation a look at the formulae involved when two sine waves are added together under linear conditions may be profitable. This is because it is essential to appreciate that different frequencies can be mixed but it is only under non-linear conditions that modulation takes place. There is no difficulty with analysis when two waves acting together are of the same frequency even though they are not in phase for a phasor diagram provides all the answers. When the frequencies differ however, say of the two waves

$$v_1 = V \sin \omega_1 t \quad \text{and} \quad v_2 = V \sin \omega_2 t$$

(choosing waves of the same amplitude simplifies the mathematics) then:

$$v_1 + v_2 = V \sin \omega_1 t + V \sin \omega_2 t$$

and by expanding via Equation (A2.48)

$$v_1 + v_2 = 2V \sin \left(\frac{\omega_1 - \omega_2}{2}\right) t \times \cos \left(\frac{\omega_1 + \omega_2}{2}\right) t \qquad (6.24)$$

so because $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$ it is evident that two *new* frequencies are developed,

$$\frac{f_1 + f_2}{2} \quad \text{and} \quad \frac{f_1 - f_2}{2},$$

the former varying in amplitude at the frequency of the latter. As an example, mixing 1000 Hz with 800 Hz produces a new frequency of 900 Hz varying in amplitude at 100 Hz, the *beat note.*

6.5 MODULATION

When one wave is not added to another (as in the Section above) but is *impressed* on it by multiplication, then modulation takes place. It is a process through which some characteristic (amplitude, frequency or phase) of a *carrier wave*, f_c is varied in accordance with a *modulating signal*, f_m. It is also effected through variations in a pulse train. The most widely used systems are amplitude, frequency and pulse-code modulation.

6.5.1 Amplitude Modulation

In this the amplitude of the carrier wave is varied in accordance with the characteristics of the modulating signal. Thus if the carrier wave has an original amplitude of V_c at a frequency f_c, a picture of its modulated form shows an amplitude varying above and below V_c, but the wave still maintains its frequency f_c as shown in Figure 6.10(i). However, because the amplitude of the sine wave has been made to vary we would expect additional frequencies to be generated.

The amplitude of the carrier wave has that of the modulating wave added to it so it is equal to $(V_c + V_m \sin \omega_m t)$ and this varies at the frequency of the carrier wave so the instantaneous value of the modulated wave is:

$$v = (V_c + V_m \sin \omega_m t) \sin \omega_c t \text{ volts}$$

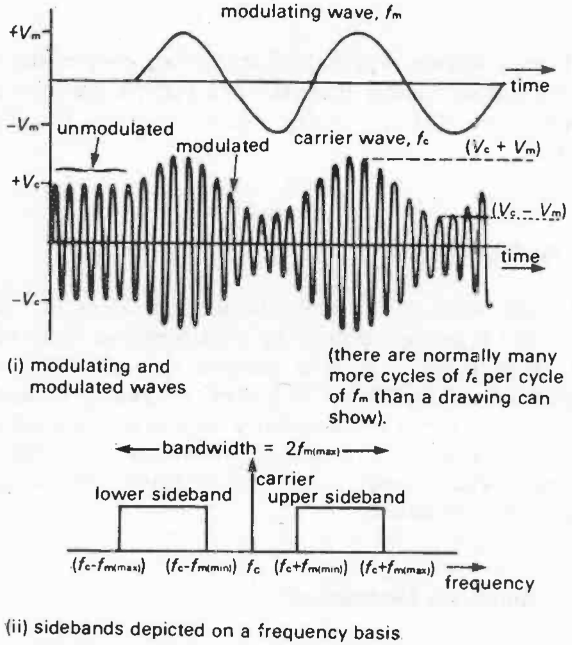


Fig. 6.10 Amplitude modulation

hence

$$v = V_c \sin \omega_c t + V_m \sin \omega_m t \times \sin \omega_c t \qquad (6.25)$$

and we already see that direct multiplication of the two frequencies is involved.

Expanding via Equation (A2.46) gives:

$$v = V_c \sin \omega_c t + \frac{V_m}{2} \times \cos(\omega_c - \omega_m)t - \frac{V_m}{2} \times \cos(\omega_c + \omega_m)t \qquad (6.26)$$

or

$$v = V_c \sin 2\pi f_c t + \frac{V_m}{2} \times \cos 2\pi(f_c - f_m)t - \frac{V_m}{2} \times \cos 2\pi(f_c + f_m)t \qquad (6.27)$$

the familiar mixture of carrier plus lower and upper side frequencies, i.e.

$$f_c \text{ plus } (f_c - f_m) \text{ and } (f_c + f_m).$$

6.5.1.1 Modulation Factor

This is defined as the ratio of the difference to the sum of the largest and smallest amplitudes of the carrier reached during one cycle of modulation, that is:

$$\textit{modulation factor, } m = \frac{(V_c + V_m) - (V_c - V_m)}{(V_c + V_m) + (V_c - V_m)}$$

which reduces very simply to

$$m = \frac{V_m}{V_c} \tag{6.28}$$

and accordingly Equation (6.25) can be rewritten as

$$v = V_c \left(1 + \frac{V_m}{V_c} \times \sin \omega_m t\right) \sin \omega_c t$$

i.e. $$v = V_c(1 + m \sin \omega_m t) \sin \omega_c t \tag{6.29}$$

6.5.1.2 *Bandwidth*

Generally f_m is not a single frequency but a band of frequencies, say from lowest to highest, $f_{m(min)}$ to $f_{m(max)}$. An a.m. (amplitude-modulated) wave may therefore be depicted as in Figure 6.10(ii) in which the term side-frequency gives way to *sideband.* From the sketch it is clear that the bandwidth is given by $(f_c + f_{m(max)}) - (f_c - f_{m(max)})$ i.e.

$$\text{bandwidth} = 2f_{m(max)} \tag{6.30}$$

or simply twice the highest modulating frequency.

EXAMPLE 6.8:
A modulated wave has an amplitude of 6 V, a bandwidth of 30 kHz and modulation factor 0.2. If the modulation is a single sine wave what is its frequency and the amplitude in dB relative to that of the carrier?

$V_c = 6$ V $\qquad m = 0.2$.

For a bandwidth of 30 kHz, from Equation (6.30),

$$f_m = \frac{30 \text{ kHz}}{2} = 15 \text{ kHz}$$

From Equation (6.28)

$$\frac{V_m}{V_c} = m = 0.2 \quad \therefore V_m = 0.2 \times 6 = 1.2 \text{ V}$$

in decibels, $20 \log_{10} 0.2 = 20 \times -0.699 = -14$ dB ,

i.e. the modulating wave has a frequency of 15 kHz and an amplitude 14 dB below that of the carrier.

6.5.1.3 The Modulation Process

What is required is a practical means of creating the results expressed by Equations (6.26) or (6.27) and for this the non-linear device (for example, the input characteristic of a transistor) is used. A non-linear characteristic can be represented by the power series

$$i = a + bv + cv^2 + dv^3 + \ldots$$

as explained in detail in Appendix 2.1.4.1. The Appendix also shows that because the constant d and those which follow are likely to be small, generally only the first three terms need to be considered. This is fortunate because otherwise the mathematics get truly out of hand. The current i, flowing through a non-linear device when the voltage v is expressed by $(v_c + v_m)$ is therefore

$$i = a + b(v_c + v_m) + c(v_c + v_m)^2$$

$$\therefore \quad i = a + bv_c + bv_m + cv_c^2 + 2cv_cv_m + cv_m^2$$

so because

$$v_c = V_c \sin \omega_c t \quad \text{and} \quad v_m = V_m \sin \omega_m t ,$$

$$i = a + bV_c \sin \omega_c t + bV_m \sin \omega_m t$$
$$+ cV_c^2 \sin^2 \omega_c t + 2cV_cV_m \sin \omega_c t \times \sin \omega_m t$$
$$+ cV_m^2 \sin^2 \omega_m t \qquad (6.31)$$

Of these six terms the second and fifth are of greatest interest for it is evident that they produce the frequencies required, that is, for these two terms only, from Equation (A2.46):

$$i_{mod} = bV_c \sin \omega_c t + cV_c V_m \cos(\omega_c - \omega_m)t$$
$$- cV_c V_m \cos(\omega_c + \omega_m)t$$

thus containing the carrier with its lower and upper side frequencies.

The remaining four terms are unwanted extras: a is a d.c. component; $bV_m \sin \omega_m t$ is a small amplitude version of the modulating frequency (V_m reduced by the factor b); and these are easily filtered out because they are of much lower frequency than that of the modulated wave.

We are left with

$$cV_m^2 \sin^2 \omega_m t \quad \text{and} \quad cV_c^2 \sin^2 \omega_c t$$

which reveal their contents on application of Equation (A2.42), for example:

$$cV_m^2 \sin^2 \omega_m t = \frac{cV_m^2}{2}(1 - \cos 2\omega_m t)$$

$$= \frac{cV_m^2}{2} - \frac{cV_m^2}{2} \cos 2\omega_m t,$$

a d.c. component plus second harmonic of the modulating frequency. The last term, $cV_c^2 \sin^2 \omega_c t$ also produces a d.c. component but with a second harmonic of the carrier frequency. Both the second harmonics are sufficiently remote from the carrier that they too are easily removed by filtering. This is a further reminder that such non-linearity gives rise to second harmonic distortion.

EXAMPLE 6.9:

A carrier of 1 MHz at 0.2 V peak and a signal at 800 Hz and 20 mV peak are applied in series to the base of a transistor which has a collector current represented by $i_c = (1.2 + 1.5v_{be} + 0.5v_{be}^2)$ mA. What are the frequency components of the collector current?

$$V_c = 0.2 \text{ V}$$

$$V_m = 0.02 \text{ V}$$

$$f_c = 10^6 \text{ Hz}$$

$$f_m = 800 \text{ Hz}$$

$$a = 1.2\,, \quad b = 1.5\,, \quad c = 0.5\,.$$

By looking back through this Section and compiling a complete list of the mathematical expressions for the components, the answer can be set out as in the Table 6.1 (overleaf).

6.5.1.4 Demodulation

To show that *demodulation* or *detection* is possible we must be able to uncover a component at modulation frequency from within the expression for the modulated wave and also it must arise from a component which is not filtered out before transmission. In the process of demodulation non-linearity is again involved so using the general form $i = a + bv + cv^2$, (Sect.A2.1.4.1) from Equation (6.29):

$$v = V_c(1 + m \sin \omega_m t) \sin \omega_c t$$

$$\therefore \quad i = a + bV_c(1 + m \sin \omega_m t) \sin \omega_c t$$

$$+ cV^2 (1 + m \sin \omega_m t)^2 \sin^2 \omega_c t\,.$$

Expanding the second term does not lead to a component relating to ω_m but expanding the third term does:

$$cV_c^2 (1 + m \sin \omega_m t)^2 \sin^2 \omega_c t$$

$$= cV_c^2 \sin^2 \omega_c t(1 + 2m \sin \omega_m t$$

$$+ m^2 \sin^2 \omega_m t)$$

Table 6.1: WAVEFORM COMPONENTS (EXAMPLE 6.9)

COMPONENT(S)	DESCRIPTION	PRACTICAL EXAMPLE	
		Magnitude	Frequency
$a + \frac{cV_c^2}{2} + \frac{cV_m^2}{2}$	total dc component	1.2 + 0.01 + 0.0001 = 1.21 mA	–
$bV_c \sin \omega_c t$	carrier	0.3 mA	1.0 MHz
$bV_m \sin \omega_m t$	signal	0.03 mA	800 Hz
$cV_cV_m \cos(\omega_c - \omega_m)t$	lower side-frequency	2 μA	999,200 Hz
$-cV_cV_m \cos(\omega_c + \omega_m)t$	upper side-frequency	2 μA	1,000,800 Hz
$-\frac{cV_c^2}{2} \cos 2\omega_c t$	2nd harmonic of carrier	10 μA	2.0 MHz
$-\frac{cV_m^2}{2} \cos 2\omega_m t$	2nd harmonic of signal	0.1 μA	1,600 Hz

the middle term within the brackets looks promising for:

$$cV_c^2 \sin^2\omega_c t \times 2m \sin \omega_m t$$

$$= 2mcV_c^2 \sin^2\omega_c t \times \sin \omega_m t$$

and from Equation (A2.42)

$$= 2mcV_c^2 \sin \omega_m t \times \tfrac{1}{2}(1 - \cos 2\omega_c t)$$

$$= mcV_c^2 \sin \omega_m t - mcV_c^2 \sin \omega_m t \times \cos 2\omega_c t$$

The first term shows that the modulating frequency

$$f_m \left(= \frac{\omega_m}{2\pi}\right)$$

is in fact regained. It must, of course, be separated out by filtering.

Summing up: Modulating frequency component of a demodulated wave

$$= mcV_c^2 \sin \omega_m t \qquad (6.32)$$

Two practical forms of demodulation are:

(i) the *diode detector* which uses a resistance-capacitance network as a filter for separating out the modulating frequency;

(ii) the *square-law detector* which employs that part of the input characteristic of a transistor which conforms mainly with a power series.

6.5.2 Frequency Modulation

Compared with the amplitude form, *frequency modulation* (f.m.) is a system in which signal-noise ratio is improved at the expense of a greater bandwidth requirement (see Sect.7.3).

The modulated wave has constant amplitude but its frequency deviates from the nominal value according to the amplitude of the modulating wave. Noise which changes the amplitude of a wave and is therefore detrimental to a.m. can be removed by *limiting* the level in f.m. and no information is lost because this resides in the wave frequency, not its amplitude. The process is summed up pictorially in Figure 6.11. The amount by which the carrier frequency, f_c is varied is proportional to the amplitude of the modulating wave so, calling the frequency deviation Δf, then since $\Delta f \propto v_m$, the deviation at any instant is $\Delta f \sin \omega_m t$. The instantaneous frequency, f of the f.m. wave is therefore (nominal frequency + deviation), i.e.

$$f = f_c + \Delta f \sin \omega_m t \tag{6.33}$$

The maximum frequency deviation which can occur in a system is fixed at the design stage so to put some figures in Figure 6.11, a 20 kHz frequency is shown modulating a 250 kHz carrier (carrier frequencies are normally in MHz but this cannot be shown satisfactorily on such a graph). Δf is given a value of 100 kHz and some of the frequencies encountered after modulation are also shown.

EXAMPLE 6.10:
On Figure 6.11 what is the f.m. wave frequency at $t = 40\ \mu s$?

At $t = 40\ \mu s$ the cycle has actually been in existence for a time of 30 μs, so from Equation (6.33):

$$f = f_c + \Delta f \sin \omega_m t = 250 \times 10^3$$

$$+ 100 \times 10^3 \sin(2\pi \times 20 \times 10^3 \times 30 \times 10^{-6})$$

i.e. $f = 250 \times 10^3 + 10^5 \sin 3.77$ radians

(or $\sin 3.77 \times \dfrac{180}{\pi} = \sin 216^\circ$)

$\therefore \quad f = 250 \times 10^3 - 0.588 \times 10^5\ \text{Hz} = 191.2\ \text{kHz}$.

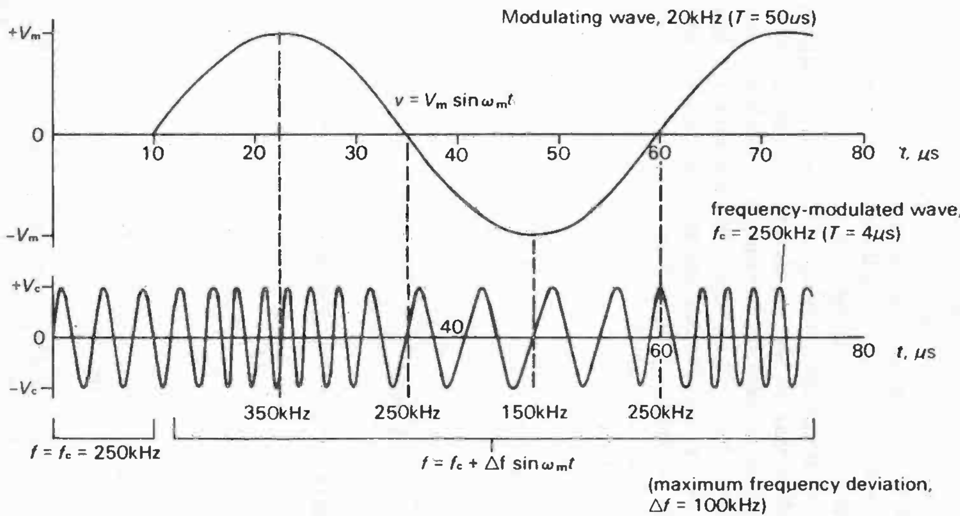


Fig. 6.11 Frequency modulation

Equation (6.33) gives the modulated wave frequency at any time, t but what we also need is a mathematical expression for the *components* of the waveform. A wave of constant frequency is easily handled mathematically because it always moves through ωt radians in t secs. The f.m. wave does not have a constant angular velocity so an expression is first needed to show how many radians the carrier moves through in a time of t seconds when it is modulated and therefore *changing* in frequency. Trying to do this graphically is a most unrewarding task so we are obliged to have recourse to the mathematical process of *integration.* Readers proficient in this art will have no difficulty in following the integration step but others will perhaps accept it on the basis that we do not wish to become involved in the more complex mathematics. Thus the angle ϕ through which the carrier moves in t secs is given by:

$$\phi = 2\pi \int_0^t f\mathrm{d}t = 2\pi \int_0^t (f_c + \Delta f \sin \omega_m t)$$

$$= 2\pi \left[f_c t - \frac{\Delta f}{\omega_m} \cos \omega_m t \right]$$

(these are the steps referred to above)

$$\therefore \quad \phi = 2\pi f_c t - \frac{\Delta f}{f_m} \times \cos \omega_m t$$

$$= \omega_c t - \frac{\Delta f}{f_m} \times \cos \omega_m t$$

The modulation index, m of an f.m. wave is defined as the ratio of the maximum deviation to the modulating frequency, i.e.

$$m = \frac{\Delta f}{f_m} \qquad (6.34)$$

and when Δf and f_m both have maximum values, m becomes the *deviation ratio.*

$$\therefore \quad \phi = \omega_c t - m \times \cos \omega_m t \tag{6.35}$$

and substituting this in the general formula for the f.m. wave, i.e.

$$v = V_c \sin \phi$$

then the instantaneous value:

$$v = V_c \times \sin(\omega_c t - m \cos \omega_m t) \tag{6.36}$$

Hence from Equation (A2.38)

$$v = V_c[\sin \omega_c t \times \cos(m \times \cos \omega_m t)$$
$$- \cos \omega_c t \times \sin(m \times \cos \omega_m t)]$$

The formula is becoming rather involved but its complexity can be eased by expanding sin and cos terms into series and then approximating as follows.

From Equations (A2.26) and (A2.27)

$$v = V_c \left[\sin \omega_c t \left(1 - \frac{m^2 \cos^2 \omega_m t}{2}\right.\right.$$
$$\left. + \frac{m^4 \cos^4 \omega_m t}{24} - \ldots\right)$$
$$\left. - \cos \omega_c t \left(m \times \cos \omega_m t - \frac{m^3 \cos^3 \omega_m t}{6} + \ldots\right)\right]$$

As the order of the term increases so its effect decreases, especially with small values of m so to avoid mathematical confusion we are justified in neglecting all terms above the second order so that approximately:

$$v = V_c \left[\sin \omega_c t \left(1 - \frac{m^2 \cos^2 \omega_m t}{2}\right) - m \times \cos \omega_c t \times \cos \omega_m t \right]$$

and from Equation (A2.42)

$$v = V_c \left[\sin \omega_c t \left\{1 - \frac{m^2}{4} (\cos 2\omega_m t + 1)\right\} - m \times \cos \omega_c t \times \cos \omega_m t \right]$$

$$= V_c \left[\sin \omega_c t \left(1 - \frac{m^2}{4}\right) - m \times \cos \omega_c t \times \cos \omega_m t - \frac{m^2}{4} \sin \omega_c t \times \cos 2\omega_m t \right]$$

and finally expanding the second and third terms by using Equations (A2.47) and (A2.44) and also by changing from ω to f:

$$v = V_c \left(1 - \frac{m^2}{4}\right) \sin \omega_c t - \frac{m V_c}{2} \left[\cos 2\pi(f_c + f_m)t + \cos 2\pi(f_c - f_m)t \right] - \frac{m^2 V_c}{8} \left[\sin 2\pi(f_c + 2f_m)t + \sin 2\pi(f_c - 2f_m)t \right] \tag{6.37}$$

and it is now evident that the f.m. waye comprises:

(first term) the carrier at an amplitude of $V_c\left(1-\dfrac{m^2}{4}\right)$

(second term) a pair of side frequencies (f_c+f_m) and (f_c-f_m), amplitude $\dfrac{mV_c}{2}$

(third term) a second pair of side frequencies (f_c+2f_m) and (f_c-2f_m), amplitude $\dfrac{m^2V_c}{8}$.

The higher order terms which are neglected in the above proof result in side frequencies theoretically to infinity, i.e. $(f_c \pm 3f_m)$, $(f_c \pm 4f_m)$ etc., but generally of lesser account because of low amplitude. However this does demonstrate that f.m. requires a large bandwidth which might be looked upon as the result of continuously distorting the carrier wave as its frequency is varied.

6.5.2.1 Bandwidth

The actual bandwidth required for any particular system cannot be determined from Equation (6.37) because of the approximations we have been forced to use and because it is not known just how many side-frequencies should be catered for. It might at first be considered that the bandwidth = $2\Delta f$ as for example in Figure 6.11, 200 kHz since the lowest and highest wave frequencies are 150 and 350 kHz. But the point already mentioned that changing the wave frequency produces harmonics indicates that something greater is required. If Equation (6.37) is extended, it will be found that the side-frequency amplitude falls to around 1% of that of the carrier at the eighth order term indicating that a band 8 times the modulating frequency is required on each side of the nominal carrier frequency. Working systems frequently manage with less and more complicated mathematical work which brings in the relationship between bandwidth and modulation index has established that:

$$\text{bandwidth required} = 2(\text{peak frequency deviation} + \text{modulating frequency}) = 2(\Delta f_{max} + f_m)\ \text{Hz} \quad (6.38)$$

which Equation (6.34) shows to be equivalent to

$$2f_m(m + 1)\ \text{Hz} \quad (6.39)$$

EXAMPLE 6.11:
A v.h.f. f.m. radio broadcast channel has a system deviation of 75 kHz and caters for modulating frequencies up to 15 kHz. What is the approximate bandwidth required?

From Equation (6.38)

$$\text{bandwidth required} = 2(75 \times 10^3 + 15 \times 10^3)$$

$$= 180\ \text{kHz}$$

or using Equation (6.39),

$$m = \frac{\Delta f}{f_m} = \frac{75}{15} = 5$$

$$\therefore \quad \text{bandwidth required} = 2 \times 15 \times 6 = 180\ \text{kHz}$$

which is considerably more than that required for a.m. which in this case is $2 \times 15 = 30$ kHz.

EXAMPLE 6.12:
A 20 MHz f.m. carrier wave has an amplitude of 0.5 V and its frequency deviates by its maximum of 7.5 kHz when a 15 kHz modulating signal is applied. What are the amplitudes and frequencies of the carrier and first and second order side-frequencies?

$f_c = 20$ MHz, $\quad \Delta f = 7.5$ kHz,

$f_m = 15$ kHz, $\quad V_c = 0.5$ V.

From Equation (6.34),

$$m = \frac{\Delta f}{f_m} = \frac{7.5 \text{ kHz}}{15 \text{ kHz}} = 0.5 \,,$$

then from Equation (6.37)

amplitude of carrier component

$$= V_c\left(1 - \frac{m^2}{4}\right) = 0.5\left(1 - \frac{0.25}{4}\right)$$

$$= 0.47 \text{ V}, \quad \text{frequency} = 20 \text{ MHz}$$

amplitudes of 1st side-frequencies

$$= \frac{mV_c}{2} = \frac{0.5 \times 0.5}{2} = 0.125 \text{ V}$$

1st side-frequencies

$$= (f_c \pm f_m) = 20.015 \text{ and } 19.985 \text{ MHz}$$

amplitudes of 2nd side-frequencies

$$= \frac{m^2 V_c}{8} = \frac{0.5^2 \times 0.5}{8} = 0.016 \text{ V}$$

2nd side-frequencies

$$= (f_c \pm 2f_m) = 20.03 \text{ and } 19.97 \text{ MHz}$$

which supports the earlier suggestion that with low values of m the amplitudes of the side frequencies fall rapidly as the order increases.

6.5.2.2 *Phase Deviation*

When the frequency of a carrier wave is increased then, as Figure 6.11 indicates, the wave passes a given point *before* its nominal time, i.e. the phase leads that of the unmodulated carrier. Equally as the frequency decreases the phase lags. The *peak phase deviation* is calculated from Equation (6.35), i.e. $\phi = \omega_c t - m \cos \omega_m t$ and ϕ is minimum when $\cos \omega_m t = +1$ for then $\phi = \omega_c t - m$ and also ϕ is maximum when $\cos \omega_m t = -1$ for then $\phi = \omega_c t + m$. Thus the peak excursions of ϕ are equal to the modulation index, m.

EXAMPLE 6.13:
A carrier is frequency modulated by a 0.4 V, 12 kHz sinusoidal modulating wave. If the maximum system deviation is 60 kHz and maximum amplitude of the carrier is 0.5 V, by how much does the carrier phase shift relative to when it is unmodulated?

$$\Delta f = \frac{0.4}{0.5} \times 60 = 48 \text{ kHz}$$

From Equation (6.34)

$$m = \frac{\Delta f}{f_m} = \frac{48}{12} = 4$$

then phase deviation $= m = 4$ radians, or

$$4 \times \frac{180}{\pi} \text{ degrees} = 229^\circ .$$

6.5.3 Pulse Modulation

With a train of pulses as the carrier, information can be transmitted by variations in pulse amplitude, duration or by the time position of each pulse in the train. There is also a more complex system of pulse-code modulation (p.c.m.). This has the advantage of superior noise immunity which is one of the factors hastening the move out of an analogue world into the digital. It does not represent perfection however, there are disadvantages which are inherent in the basic principle so a compromise between conflicting factors has to be made. Because of the complexity of some of the circuitry there is little doubt that the advent of IC's has speeded development. But although the sphere of "digital" is steadily increasing, the fact remains that most of the human world is analogue so conversion from one form to the other is frequently required. Although the ultimate aim of this Section is to develop some of the ideas associated with p.c.m. it is necessary to consider pulse-amplitude modulation (p.a.m.) first because in fact p.c.m. is p.a.m. with something added.

6.5.3.1 Pulse-Amplitude Modulation

Pulse-amplitude modulation is developed diagrammatically in Figure 6.12 in which a pulse train is running at 7000 per second, that is, the *pulse repetition frequency*, $f_p = 7000$. This is modulated by a single sine wave at 1000 Hz. The contacts of the switch S close for a short period of time at 1/7000 s (= 0.143 ms) intervals, hence allowing the voltage v_m to be applied to R. It is said to be *sampling* the modulating signal. Sampling is the process of describing any waveform in terms of its value at frequent time intervals as illustrated in (ii) of the Figure.

The use of switch S is purely for illustration, pulse trains are invariably generated by semiconductors (Sect.6.3.1) for no mechanical switch could operate at the frequencies required.

The technique of pulse modulation rests on the *Sampling Theorem* first proposed by the American scientist, H. Nyquist and later developed further by an American mathematician, C. E. Shannon. Expressed in our own simplified terms the theorem states that given a band of modulating frequencies extending up to $f_{m(max)}$ Hz, then this can be completely recovered after sampling provided that:

$$\text{sampling rate} = \textit{at least}\ 2 \times f_{m(max)}\ \text{Hz} \qquad (6.40)$$

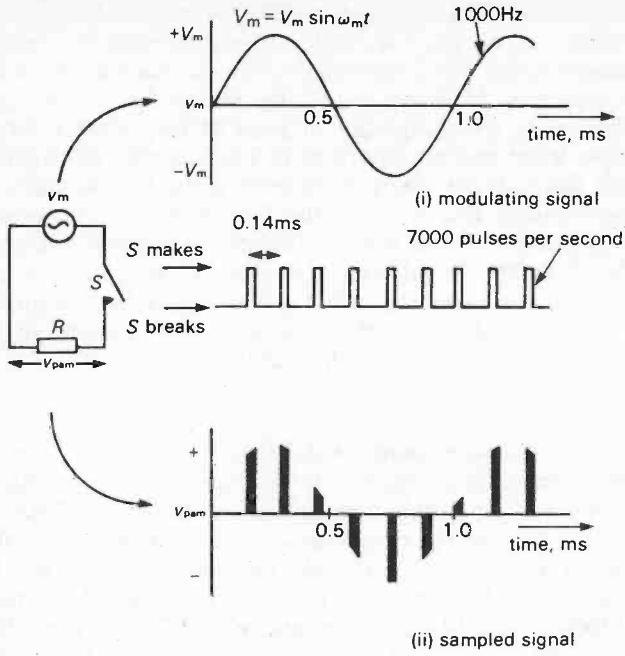


Fig. 6.12 Pulse amplitude modulation

and that recovery may be effected simply by means of a low-pass filter having a cut-off frequency at $f_{m(max)}$ Hz (filters are considered in Book 1).

Recalling that Fourier's theorem enables us to analyse a periodic function into its sinusoidal components, the result of the switch action can be seen as multiplying the modulating frequency, f_m by each of the components of the pulse train. We will use a simplified form of the Fourier series for a rectangular pulse train but this in no way invalidates the conclusions. For a pulse repetition frequency f_p (ω_p =

$2\pi f_p$), the series for the pulse train (or carrier) voltage is:

$$v_p = a_0 + a_1 \sin \omega_p t + a_2 \sin 2\omega_p t + a_3 \sin 3\omega_p t + \ldots$$

where a_0 is the d.c. component and the subsequent terms are the harmonic sinusoidal components. The series runs to infinity but it will become evident that only the first few terms are significant. Multiplying this series for the pulse train by the modulating frequency $v_m = V_m \sin \omega_m t$ gives an expression for the p.a.m. voltage:

$$v_{pam} = V_m \sin \omega_m t(a_0 + a_1 \sin \omega_p t + a_2 \sin 2\omega_p t + \ldots)$$

$$= a_0 V_m \sin \omega_m t + a_1 V_m \sin \omega_p t \times \sin \omega_m t + a_2 V_m \sin 2\omega_p t \times \sin \omega_m t + \ldots$$

Then from Equation (A2.46) and since $\omega_m = 2\pi f_m$ and $\omega_p = 2\pi f_p$:

$$v_{pam} = (a_0 V_m) \sin 2\pi f_m t + \frac{a_1 V_m}{2} [\cos 2\pi(f_p - f_m)t - \cos 2\pi(f_p + f_m)t] + \frac{a_2 V_m}{2} [\cos 2\pi(2f_p - f_m)t - \cos 2\pi(2f_p + f_m)t] + \ldots \quad (6.41)$$

showing that the components of a p.am. wave are:

(i) the modulating frequency;

(ii) lower and upper side-frequencies (or sidebands if f_m is a range of frequencies) of each harmonic contained within the carrier.

It is the fact that the modulating frequency is one of the components which makes demodulation simple.

The validity of the Sampling Theorem can be seen by expressing graphically on a frequency basis each of the components of Equation (6.41) as in Figure 6.13. Practical figures are used to demonstrate pulse modulation of a 50 Hz – 10 kHz music circuit, i.e. f_m of Equation (6.41) now becomes a range $f_{m(min)} = 50$ Hz to $f_{m(max)} = 10$ kHz. To comply with the theorem, sampling must be as expressed by Equation (6.40), i.e. at a rate 2 x 10 kHz = 20 kHz, but note that this is the theoretical minimum value. Accordingly Figure 6.13 shows at (i) the frequency spectrum for a sampling rate or pulse repetition frequency, $f_p = 20$ kHz and clearly only a perfect low-pass filter cutting off at 10 kHz could separate the modulating frequency from the remainder. The cure is of course to increase f_p slightly so that there is a gap between the modulating frequency band and the lower sideband of f_p. The significance of the proviso "at least" in Equation (6.40) is now apparent and the more practical arrangement is shown in Figure 6.13(ii) in which a low-pass filter easily separates the modulating frequency band from all other components of modulation, i.e. satisfactory demodulation is achieved. On the other hand if $f_p < 2f_{m(max)}$ there is overlapping of the modulating frequency band with the lower sideband of f_p, which is clearly inadmissible.

6.5.3.2 *Quantising Error*

In the move from p.a.m. to the more complex p.c.m. comes the need to understand *quantising error* for it is this which results in an important compromise which has to be made. Figure 6.12 shows that for the duration of each pulse its amplitude follows exactly that of the modulating signal. The more advanced systems do not transmit these pulses but code signals instead representing their amplitudes. There is an infinite number of amplitudes possible between the maximum + and – values of v_{pam} [Fig.6.12(ii)] but in a

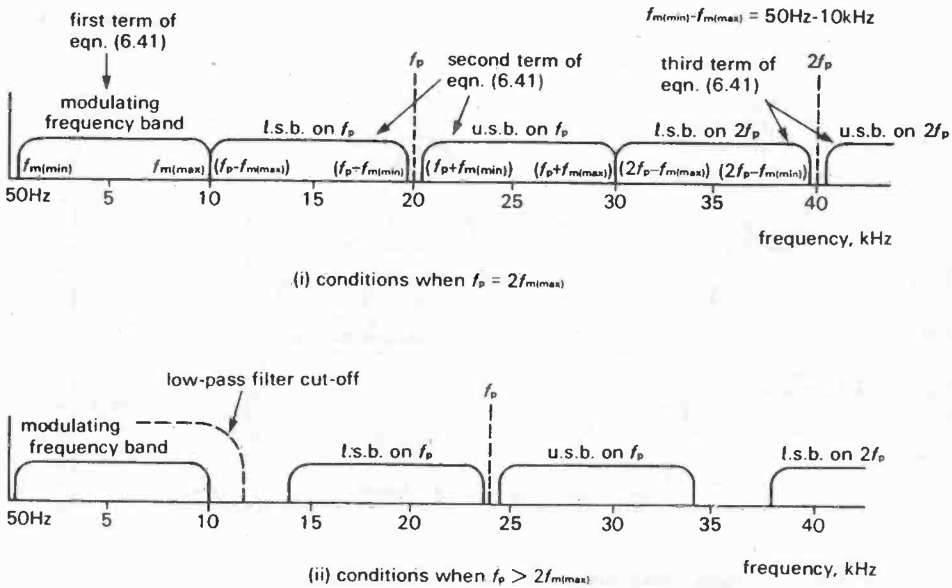


Fig. 6.13 Frequency spectra of P.A.M. pulse trains

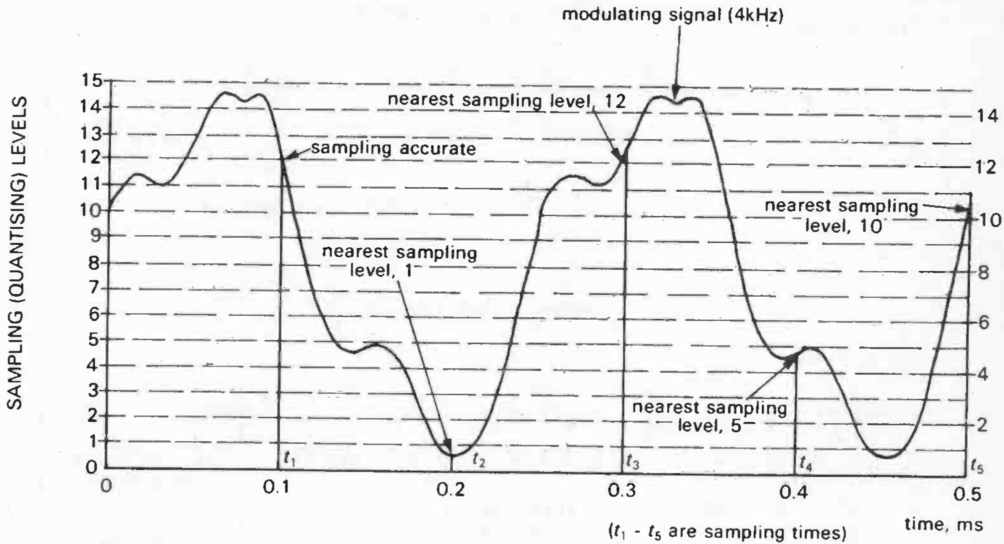


Fig. 6.14 Sampling a 4kHz wave at 10kHz

practical system a discrete number of levels must be chosen. Figure 6.14 illustrates a sampling technique. The maximum peak-peak voltage which a modulating signal might have is divided in this case into 16 evenly spaced *sampling levels* (marked 0–15). The signal is sampled (i.e. measured) regularly either to the nearest level or to the next below, in the case shown at 0.1 ms intervals. At t_1 the signal voltage is exactly that represented by sampling level 12 so the sampling is accurate. At $t_2 - t_5$ however the signal voltage is not in agreement with that for any sampling level, hence as shown, the nearest one is taken. The modulating signal in the Figure is now described by the five sampling levels 12, 1, 12, 5, 10 which may be transmitted in the form of pulses of relative amplitudes.

Demodulation however cannot reproduce the original waveform exactly because of the inaccuracies in sampling, the reconstructed signal dithers about its true value. This in an audio system is evident as an additional background noise, known as *quantisation noise* and due to *quantisation error. Quantisation* means measuring in discrete amounts and in this application is defined as the process of sampling a signal waveform for its conversion into digital form.

Figure 6.14 illustrates the case where the maximum error is equal to half a quantising step (the difference between two sampling levels). If, for example each step were equivalent to 0.1 V then the quantisation could be considered to be producing an additional *error signal* of maximum amplitude 0.05 V. To reduce the error signal it follows that the number of quantising steps must be increased. As an example, suppose the 16 steps of Figure 6.14 cover a maximum incoming signal range of 1 V (peak-peak):

then each quantising step represents

$$\frac{1000}{16} = 62.5 \text{ mV}$$

and maximum error signal amplitude

$$= \frac{62.5}{2} = 31.25 \text{ mV}$$

Now if the number of steps is increased to 32, the maximum signal error amplitude is

$$\frac{1000}{32 \times 2} = 15.625 \text{ mV}$$

$\therefore$ Change in quantising noise =

$$20 \log_{10} \frac{15.625}{31.25} = -6.02 \text{ dB},$$

showing that for every doubling of the number of quantising steps the quantisation noise is reduced by 6 dB.

Generally audio systems use 128 or 256 quantising steps and with these the error signal becomes manageable, the actual signal/quantisation noise ratios being of the order of 42 and 48 dB respectively.

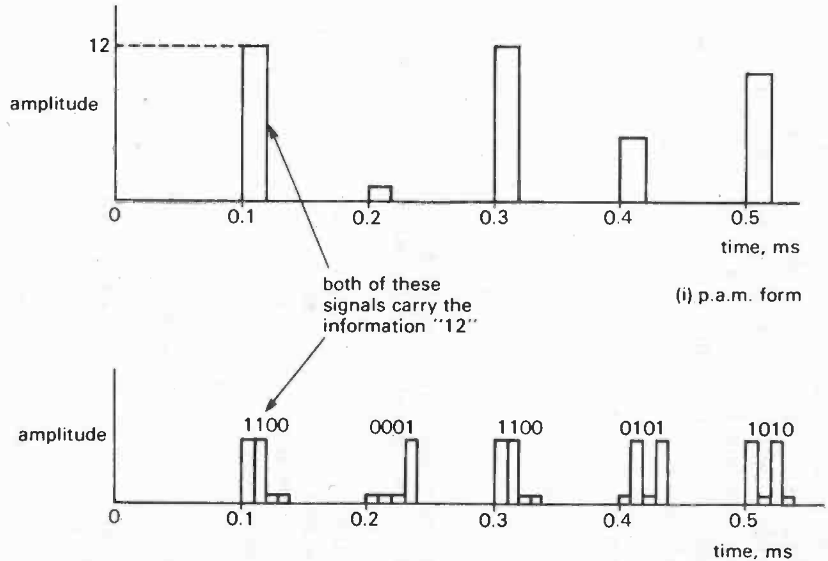
6.5.3.3 Pulse-Code Modulation

If pulses having amplitudes representing sampling levels are transmitted over a line, the inevitable distortions lead to further errors in the reconstituted signal. PCM avoids this by using instead groups of *constant amplitude* pulses with spaces (the 1's and 0's of the computer world) corresponding to codes assigned to the sampling levels. The *binary code* (Sect.13.1) is in widespread use for this and the change from p.a.m. to p.c.m. is demonstrated in Figure 6.15 which at (i) repeats the p.a.m. signals for Figure 6.14. As an example, the level of the modulating wave at $t = 0.3$ ms (which is rounded down to level 12) is coded 1100 (Sect.13.1.2). Two pulses followed by two spaces therefore express level 12.

From (ii) of the Figure the *bit rate* (a 1 or a 0 is a *b*inary dig*it* or *bit*) can be calculated for there are four bits per 0.1 ms, i.e. 40,000 bits per second (b/s) or generally:

$$\text{transmission rate} = n \times f_s \text{ b/s} \tag{6.42}$$

and $$\text{number of sampling levels} = 2^n \tag{6.43}$$



Decimal	Binary
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
10	1010
11	1011
12	1100

Fig. 6.15 PAM and PCM line signals of Fig. 6.14

where n is the number of digits in the code and f_s is the sampling frequency.

To increase the number of sampling levels to 256 as earlier suggested therefore requires $n = 8$ ($2^8 = 256$, see Table 13.1) hence giving a transmission rate of $n \times f_s = 8 \times 10^4 =$ 80,000 b/s, double that for $n = 4$.

It is more easily said than done to determine an exact relationship between bit rate and bandwidth for all practical cases. As a rough guide however we might assume that the bandwidth required for successful transmission of p.c.m. signals lies between 0.5 and 1.0 times the bit rate, say about 0.6. Accordingly to transmit a 4 kHz wave at a sampling frequency of 10 kHz (as in Fig.6.14) using, say, 128 sampling levels ($2^7 = 128$, i.e. $n = 7$):

$$\text{transmission rate} = n \times f_s = 7 \times 10^4 = 70{,}000 \text{ b/s}$$

and approximate bandwidth required

$$= 0.6 \times 70{,}000 = 42 \text{ kHz},$$

a figure many times that required for analogue transmission but less if n is reduced although then quantisation noise increases.

So far we have brought to light some of the disadvantages of the system, specifically quantisation noise and requirement of wide bandwidth. To more than balance these there are major advantages which are inherent in the p.c.m. signal itself for it comprises two signals of utmost simplicity only:

(i) in computer systems the p.c.m. signals can be stored in conventional computer memory, especially convenient when $n = 7$ or 8 for the standard memory unit caters for 8 bits;

(ii) in transmission over distances distortion of the pulses is of little consequence as long as the presence or absence of a pulse can be detected. When the p.c.m. signals are weak, amplifiers are not used, each detected pulse is *regenerated*, for example by triggering a monostable multivibrator as in Figure 6.9(ii). The signals therefore start out again as new with all incoming noise blocked. In contrast analogue systems gather noise as the signal progresses along a line. In fact the use of p.c.m. in telephone connexions is now such that we frequently talk over connexions with no inkling whatsoever that we have "gone digital".

Reference again to Figure 6.15(ii) shows that there is unused time between the separate pulse groups so groups from other separate circuits are inserted, in fact modern telephony systems carry up to 32 separate *channels* by this method, known as *time division multiplex*. In this case:

$$\text{system transmission rate} = N \times n \times f_s \qquad (6.44)$$

where N is the number of channels.

EXAMPLE 6.14:
A p.c.m. telephony system has 32 channels with 256 sampling levels, sampled at 8 kHz. What is the system transmission rate?

$$N = 32 \qquad f_s = 8000$$

From Equation (6.43):

$$256 = 2^n \qquad \therefore \log 256 = n \log 2$$

$$\therefore \quad n = \frac{\log 256}{\log 2} = 8$$

(or directly from Table 13.1). Then from Equation (6.44):

$$\text{system transmission rate} = 32 \times 8 \times 8000 = 2.048 \text{ Mb/s}$$

In the computer world most of the p.c.m. processing is undertaken by an *analogue to digital converter* for modulation and a *digital to analogue converter* for demodulation. Very briefly the a/d converter compares the incoming analogue voltage with internally generated ones and then codes them accordingly. For demodulation a d/a converter accepts the groups of pulses and provides a current or voltage at its output in agreement with the numerical value of each code. Most of the conversion is accomplished by the use of logic circuits (Chapter 13).

6.6 SUMMARY OF KEY FORMULAE

A_i = current gain

A_v = voltage gain

C = capacitance

f_c = carrier frequency

f_m = modulating frequency

f_s = sampling frequency

Δf = frequency deviation

h_{fe} = transistor current gain (c-e)

h_{ie} = transistor input resistance (c-e)

L = inductance

M = mutual inductance

N = turns ratio

n = number of digits in binary code

p = $2\pi \times$ modulating frequency ($2\pi f_m$) – rads/s

R = resistance

R_L = load resistance

t = time

V = voltage — V_m = maintaining voltage; V_s = striking voltage

X = reactance

β = feedback fraction

θ = amplifier phase angle

ϕ = feedback network phase angle

ω = $2\pi \times$ carrier frequency ($2\pi f_c$) – rads/s.

QUANTITY	FORMULA	UNIT	SECTION
Oscillators (Figs.6.2 – 6.5)			
Conditions for oscillation:			
General	$\beta A_{\mathrm{v}} = 1, \ (\phi + \theta) = 0$		6.1
LC oscillator	$f_{\mathrm{osc}} = \dfrac{1}{2\pi\sqrt{LC}}$	Hz	6.1.1
Tuned collector oscillator	$h_{\mathrm{fe}} = \dfrac{R_{\mathrm{L}} + N^2 h_{\mathrm{ie}}}{NR_{\mathrm{L}}}$		6.1.1.1
Colpitts oscillator	$f_{\mathrm{osc}} = \dfrac{1}{2\pi}\sqrt{\dfrac{C_1 + C_2}{LC_1C_2}}$	Hz	6.1.1.2
Hartley oscillator	$f_{\mathrm{osc}} = \dfrac{1}{2\pi\sqrt{(L_1 + L_2 \pm 2M)C}}$	Hz	6.1.1.2

QUANTITY	FORMULA	UNIT	SECTION
Phase-shift oscillator	$f_{osc} = \frac{1}{2\pi\sqrt{6} \times RC} = \frac{0.065}{RC}$	Hz	6.1.2.1
	$X_c = \sqrt{6} \times R \quad A_i = -29$		
Wien Bridge oscillator	$f_{osc} = \frac{1}{2\pi RC}$	Hz	6.1.2.2
	$\beta = \frac{1}{3} \quad , \quad A_v = 3$		
Sawtooth Wave generator (Fig.6.7)	$f_{osc} = \frac{1}{CR \log_e \left(\frac{V - V_m}{V - V_s} \right)}$	Hz	6.2

7. COMMUNICATION

Communication in the electronics sense is the process of conveying information from one point to another. In so doing a *channel* is required as the communication link between those two points. The *rate* of transfer of information over any channel has an upper limit depending not only on the channel itself but often also on the conditions surrounding it. This Chapter looks at the formulae and calculations involved in assessment of channel performance.

7.1 NOISE

Noise is present in all electronic systems and later in the Chapter it is shown to be one of the limiting factors to the usefulness of a communication channel. It is generally defined as an unwanted disturbance in a communication channel, "unwanted" because its effects are almost without exception, detrimental. What usually matters most is the level of the noise relative to that of the signal in which case the degree of noise interference is expressed by the *signal-to-noise ratio* which in decibels is simply:

Signal-to-noise ratio (S/N)

$$= 10 \log \frac{\text{signal power}}{\text{noise power}} \text{ dB}$$

$$= 20 \log \frac{\text{signal voltage}}{\text{noise voltage}} \text{ db} \qquad (7.1)$$

Noise may arise within a communication channel both through generation within the channel and also by injection from outside.

7.1.1 Thermal Noise

This is generated within a channel and is due to the random nature of electron movement in any resistance. From Section 1.4 it is easy to visualize that at any point within a conductor the number of electrons is continually varying. Between any two points therefore there is a minute random noise voltage which increases with temperature because heat provides the electrons with energy, the effect is known as *thermal agitation.* The voltage constantly changes in value as electrons come and go and because the changes occur at random the generated noise frequency extends over the whole spectrum (*white noise*). The channel itself acts as a band-pass filter, hence the actual noise voltage measurable is proportional to the channel bandwidth. Thus the noise power $P_n = kT$ watts per Hz bandwidth where k is a factor introduced to relate energy with temperature and is known as *Boltzmann's*(A3) *Constant* which has a value of 1.38×10^{-23} joules per degree Kelvin(A3) (0 K = –273.15°C). T is the temperature in °K.

The complete formula is expressed in terms of the noise voltage, V_n and includes the value of the resistance, R in which the noise power is dissipated.

$$V_n = \sqrt{4kTRB} \text{ volts (r.m.s.)} \qquad (7.2)$$

B can generally be taken as between the upper and lower frequencies at which the response has fallen by 3 dB (Sect. 5.5).

The formula can be simplified for a room temperature of 290 K to:

$$V_{n(\text{room})} = 1.265 \times 10^{-10} \sqrt{RB} \text{ V} \qquad (7.3)$$

EXAMPLE 7.1:
A microphone has a frequency range from 100 Hz to 12 kHz and an impedance of 200 ohms. Its output with a certain input test signal is –70 dB/1V. What is the signal-to-noise ratio?

Assuming room temperature, $T = 290$ K, $R = 200\ \Omega$, $B = 11{,}900$ Hz, then from Equation (7.3):

$$V_n = 1.265 \times 10^{-10} \sqrt{200 \times 11{,}900}$$

$$= 1.95 \times 10^{-7}\ \text{V}.$$

In dB/1V,

$$V_n = 20 \log 1.95 \times 10^{-7} = -134.$$

Working in dB/1V is legitimate because both noise and microphone voltages are generated in the same impedance, hence:

$$\text{Signal-to-noise ratio} = -70 - (-134) = 64\ \text{dB}.$$

If the noise voltage feeds power into a matched load resistance, R as in Figure 7.1(i) the resultant voltage across the load is equal to $V_n/2$ and the power into the load

$$= \frac{\left(\dfrac{V_n}{2}\right)^2}{R}$$

hence,

noise power P_n delivered into matched load

$$= kTB \text{ watts} \tag{7.4}$$

which, perhaps rather surprisingly, is independent of the resistances. This is not the complete story however for the load itself generates thermal noise, this is catered for by (ii) of the Figure which shows that for two different resistances in parallel at the same temperature, the net noise voltage:

$$V_n = \sqrt{4kTB\,R_T}\ \text{V}$$

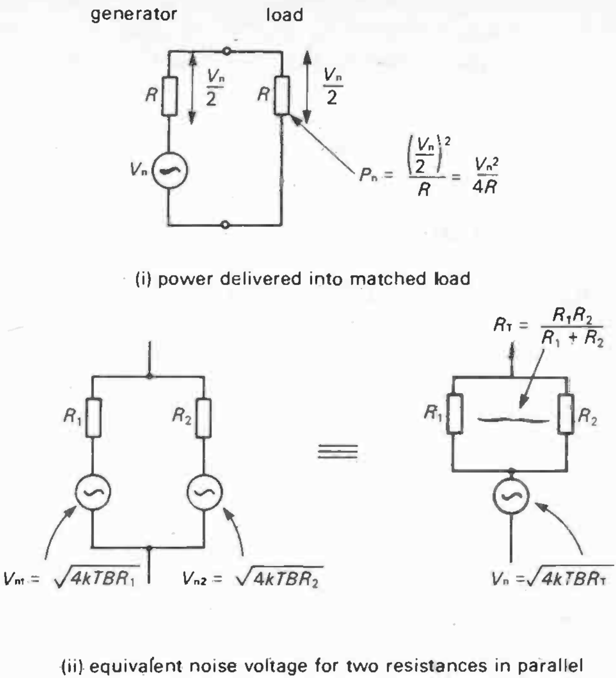


Fig. 7.1 Thermal noise generation

where

$$R_T = \frac{R_1 R_2}{R_1 + R_2} \tag{7.5}$$

and in the particular case of a matched termination, R_T is half the termination value, R, therefore:

$$V_n = \sqrt{4kTB\frac{R}{2}} = \sqrt{2kTB\,R} \text{ volts} \tag{7.6}$$

EXAMPLE 7.2:
What noise voltage will be measured across a 1000 ohm resistor at room temperature (17°C = 290 K) if the measuring equipment has a bandwidth of (i) 10 kHz, (ii) 1 MHz?

(i) $T = 290$ K $\quad B = 10^4$ Hz $\quad R = 10^3\ \Omega$

From Equation (7.2):

$$V_n = \sqrt{4 \times 1.38 \times 10^{-23} \times 290 \times 10^3 \times 10^4}$$

$$= 4 \times 10^{-7}\ \text{V} = 0.4\ \mu\text{V}.$$

(ii) at a bandwidth of 10^6 Hz, V_n increases by

$$\sqrt{\frac{10^6}{10^4}},$$

i.e. 10 times to 4.0 μV [the same answer is of course given using Eqn.(7.3)].

EXAMPLE 7.3:
A 75 ohm coaxial cable is matched into a television receiver of bandwidth 5.5 MHz. What noise power does this inject into the receiver at room temperature?

$B = 5.5 \times 10^6 \quad T = 290$ K

From Equation (7.4):

$$P_n = kTB \text{ watts} = 1.38 \times 10^{-23} \times 290 \times 5.5 \times 10^6\text{W}$$

$$= 0.022\ \text{pW}.$$

Alternatively P_n can be calculated directly in decibels by inserting the values of T and B into a basic formula with $T = 0$ K and $B = 1$ Hz whereupon $P_n = k$ watts =

−228.6 dBW, then

$$P_n = -228.6 + 10 \log T + 10 \log B \text{ dBW} \qquad (7.7)$$

$$\therefore \quad P_n = -228.6 + 10 \log 290 + 10 \log 5.5 \times 10^6$$

$$= -228.6 + 24.62 + 67.40 = -136.58 \text{ dBW}$$

(i.e. 0.022 pW).

EXAMPLE 7.4:
A 600 ohm transmission line is matched to its termination. What is the minimum signal voltage level at the termination which will maintain a signal-to-noise ratio of 80 dB at 300 K over a 1 MHz bandwidth?

$$T = 300 \text{ K} \qquad B = 10^6 \text{ Hz} \qquad R = 600\ \Omega$$

From Equation (7.6):

$$V_n = \sqrt{2 \times 1.38 \times 10^{-23} \times 300 \times 10^6 \times 600}$$

$$= 2.229 \times 10^{-6} \text{ V}.$$

Let V_s be the signal voltage, then from Equation (7.1)

$$80 \text{ dB} = 20 \log \frac{V_s}{V_n} \qquad \therefore 4 = \log \frac{V_s}{V_n}$$

$$\therefore \quad \text{antilog } 4 = \frac{V_s}{V_n} \qquad \therefore V_s = V_n \times 10^4$$

i.e. signal voltage for 80 dB S/N ratio

$$= 2.229 \times 10^{-6} \times 10^4 \text{ V} = 22.29 \text{ mV}.$$

7.1.2 Crosstalk

This is interference between two separate channels through unintentional electrical couplings between them. In telephony circuits it may actually be intelligible but generally crosstalk injected into one circuit from several others results in a *babble* of noise. Crosstalk is usually measured as the loss between the disturbing and disturbed circuits hence

$$\text{Crosstalk attenuation} = 10 \log \frac{P_s}{P_c} \text{ dB} \tag{7.8}$$

where P_s = signal power in disturbing circuit and P_c = signal power in disturbed circuit (crosstalk). Thus if a signal level at −3 dBm in one circuit results in a crosstalk signal at −60 dBm in a second circuit, then the crosstalk attenuation is simply 57 dB (not dBm).

Crosstalk is most likely to vary with frequency so not only must the method of measurement be stated but also the frequency or band of frequencies.

7.1.3 Intermodulation Noise

Non-linearity in any system carrying more than one frequency leads to *intermodulation*. For example, two frequencies f_1 and f_2 in a non-linear channel produce $(f_1 + f_2)$, $(f_1 + 2f_2)$, $(2f_1 + f_2)$, $(2f_1 + 2f_2)$ and higher order products. Although likely to be at relatively low levels, the net effect when many separate frequencies are transmitted can exceed the level of thermal noise.

7.1.4 Other Miscellaneous Sources

Several other sources of noise exist:

(i) *transistor noise* – see Section 5.8;

(ii) *man-made noise* arising mainly from impulsive currents and effective over a wide frequency band

(iii) *radio-path noise* which includes *atmospherics* due to lightning and *galactic* noise which arrives from the stars.

7.1.5 Noise Figure

Various definitions exist for both *noise figure* and *noise factor*. We can simply express them both as the ratio of the input and output signal-to-noise ratios of any system thus:

noise figure (or factor) F

$$= 10 \log \frac{\text{input } S/N \text{ power ratio}}{\text{output } S/N \text{ power ratio}} \text{ dB} \qquad (7.9)$$

so conveniently, if the two S/N ratios are expressed in decibels, the noise figure (or factor) is the difference between them. A positive figure therefore indicates that in its passage through the system the signal-to-noise ratio has worsened.

EXAMPLE 7.5:
A radio receiver has a noise figure of 8 dB. The input signal-to-noise ratio is 62 dB. What is the expected output signal-to-noise ratio?

From Equation (7.9):

$$F \text{ dB} = \text{input } S/N \text{ (dB)} - \text{output } S/N \text{ (dB)}$$

$$\therefore \quad 8 = 62 - \text{output } S/N \text{ (dB)}$$

$$\therefore \quad \text{Output } S/N \text{ ratio} = 54 \text{ dB}.$$

7.2 INFORMATION THEORY

In the late 1940's C. E. Shannon (see also Sect.6.5.3.1) published his now universally accepted work on *Information*

Theory which can broadly be described as an attempt to analyze and quantify *information*, a word which in everyday use has a rather imprecise meaning. To the less mathematically minded Shannon's paper is truly frightening so here we merely dabble in one or two facets which demonstrate the practical advantages. It is possible to calculate *information flow* for any waveform such as of speech, music or television but to keep things manageable we consider digital information only in the normal binary form of 1's and 0's represented electrically by pulses and spaces.

7.2.1 Information Content

Consider that at the receiving end of a channel the equipment is waiting for the next binary signal to arrive. The *information content* of this signal is a function of its *probability* (also see Sect.8.3.1) which is on a scale running from 0 to 1. With two different signals only there is one chance in two that a 0 will arrive next and the same chance that it will be a 1. This is expressed as:

$$P(0) = 1/2 \text{ or } 0.5 \quad \text{and} \quad P(1) = 0.5 .$$

With a denary system there is one chance in 10 that a certain signal will arrive, hence that signal brings more informaion with it and the probability is 1/10 or 0.1, expressed as $P(0) = 0.1$, $P(1) = 0.1$, $P(2) = 0.1$ etc. Put in everyday language, with no knowledge whatsoever of odds, jockeys or horses, the probability of picking the winner of a two-horse race is 0.5 which greatly exceeds that of picking the winner of a 10-horse race at only 0.1. But when the result is declared more information is given for the 10-horse race simply because 10 horses are concerned instead of 2.

A measure of the *amount* of information in any particular case is given by:

Information content of signal

$$I = \log_2 P^{-1} \text{ bits} \qquad (7.10)$$

so for binary, since $P = 0.5$

$$I = \log_2 \frac{1}{0.5} = \log_2 2 = 1 \text{ (bit)}$$

meaning that one *bit* of information is required to distinguish between two equiprobable signals.

EXAMPLE 7.6:
In a monitoring system any one of 8 different voltage levels is transmitted to indicate the level of liquid in a tank. How much information is required for each indication?

$$P = 1/8 = 0.125 .$$

From Equation (7.10):

Information content,

$$I = \log_2 P^{-1} = \log_2 \frac{1}{0.125} = \log_2 8$$

which from Equation (A2.16)

$$= \frac{\log_{10} 8}{\log_{10} 2} = \frac{0.9031}{0.3010} = 3 \text{ bits}$$

which demonstrates that lower probabilities require more information to be transmitted for error-free reception.

7.2.2 Information Flow

The section above shows how the information carried by a single digital signal is calculated in bits, the most elementary signal (binary) containing 1 bit, other signals more. The various generators produce information at different rates, measured in bits per second (b/s), for example high-quality speech at 40,000 b/s, music at 90,000 b/s and television at 50–100 million b/s (these figures are very approximate and continuously vary with the signal itself). As far as a practical channel is concerned therefore what matters is just how many bits of information can be transmitted over it per second. It is not simply a question of bandwidth of the channel for as shown earlier all channels contain some sort of noise.

7.3 CHANNEL CAPACITY

In Section 6.5.3 it is shown that in a pulse-code modulation system the input waveform is first sampled with the result coded into binary and then transmitted to line as pulses. Assuming thermal noise (and things get rather complicated if we do not), then the information capacity, C of the channel can be expressed by:

$$C = \frac{1}{2} \log_2 \left(1 + \frac{S}{N}\right) \text{ bits per sample} \quad (7.11)$$

where S/N is the average signal-to-noise ratio.

The Section also shows that the sampling rate must be at least twice the highest frequency of the input waveform so if B Hz represents the bandwidth required and assuming that the band extends down to 0 Hz, there are $2B$ samples per second, therefore:

$$C = B \log_2 \left(1 + \frac{S}{N}\right) \text{ b/s} \quad (7.12)$$

Now it must be emphasized that we are doing no more than scratching the surface of the theory, nevertheless from Equation (7.12) much of practical value can be deduced.

EXAMPLE 7.7:
If a channel has a bandwidth of 200 kHz and a signal-to-noise ratio of 20 dB, what is its capacity? If the signal-to-noise ratio is worsened to 10 dB what channel bandwidth will be required for the same information rate?

$$B = 2 \times 10^5 \text{ Hz} \qquad S/N = 20 \text{ dB} = 100$$

Then from Equation (7.12)

$$C = B \log_2 (1 + 100) = 2 \times 10^5 \log_2 101$$

$$= 2 \times 10^5 \times \frac{\log_{10} 101}{\log_{10} 2} = 1.33 \text{ Mb/s} .$$

If S/N is reduced to 10 dB (= 10)

$$B = \frac{C}{\log_2 (1 + 10)} = \frac{1.33 \times 10^6}{3.459} = 385 \text{ kHz} .$$

The exchange of signal-to-noise ratio for bandwidth is evident. Worsening of this particular S/N ratio by 10 dB requires an increase in bandwidth from 200 to 385 kHz for an equally efficient transmission of information.

For convenience Table 7.1 lists the values of $\log_2(1 + S/N)$ for a range of values of S/N. Using this table and repeating the example, when

$$S/N = 20 \text{ dB} , \ \log_2 (1 + S/N) = 6.66$$

$$\therefore \quad C = 2 \times 10^5 \times 6.66 = 1.33 \times 10^6 \text{ b/s} .$$

When $S/N = 10$ dB, $\log_2(1 + S/N) = 3.46$ as already calculated.

The formula only relates channel capacity with bandwidth and noise. This may not be the whole story, for example, *intersymbol interference* which arises when the tail of one pulse overlaps the rising edge of the subsequent one, also

Table 7.1: CALCULATION OF CHANNEL CAPACITY

S/N, dB	S/N ratio	$\log_2(1 + S/N)$	S/N, dB	S/N ratio	$\log_2(1 + S/N)$
0	1.00	1	18	63.10	6.00
1	1.26	1.18	20	100.00	6.66
2	1.58	1.37	25	316.23	8.31
3	2.00	1.58	30	1000.00	9.97
4	2.51	1.81	35	3162.30	11.63
5	3.16	2.06	40	10,000	13.29
6	3.98	2.32	45	31,623	14.95
7	5.01	2.59	50	100,000	16.61
8	6.31	2.87	55	316,228	18.27
9	7.94	3.16	60	10^6	19.93
10	10.00	3.46	65	3,162,278	21.59
12	15.85	4.07	70	10^7	23.25
14	25.12	4.71	75	3.16×10^7	24.91
16	39.81	5.35	80	10^8	26.58

degrades a channel. In fact Nyquist(A3) has shown that when this is the overriding factor, C can be as low as 2 B b/s.

7.4 SUMMARY OF KEY FORMULAE

B = bandwidth (Hz)

P = probability

P_n = noise power

P_s = signal power

P_c = crosstalk power

R = resistance

T = temperature (°K)

V_n = noise voltage

V_s = signal voltage

QUANTITY	FORMULA	UNIT	SECTION
Boltzmann's Constant	$k = 1.38 \times 10^{-23}$	J/K	7.1.1
Signal-to-noise Ratio	$= 10 \log \frac{P_s}{P_n} = 20 \log \frac{V_s}{V_n}$	dB	7.1
Thermal noise	$V_n = \sqrt{4kTRB}$	V	7.1.1
Thermal noise at room temperature	$V_{n(room)} = 1.265 \times 10^{-10}\sqrt{RB}$	V	7.1.1
Thermal noise in matched load	$V_n = \sqrt{2kTBR}$	V	7.1.1
	$P_n = -228.6 + 10 \log T + 10 \log B$	dBW	7.1.1
Crosstalk attenuation	$= 10 \log \frac{P_s}{P_c}$	dB	7.1.2
Noise figure	$F = 10 \log \frac{\text{input } S/N \text{ power ratio}}{\text{output } S/N \text{ power ratio}}$	dB	7.1.5
Information content of signal	$I = \log_2 P^{-1}$	bits	7.2.1
Channel capacity	$C = B \log_2(1 + S/N)$	b/s	7.3

8. STATISTICS

Statistics is the science of collection and analysis of numerical data. By use of the techniques sensible conclusions can be drawn and through probability calculations predictions are possible but never without an element of uncertainty. This Chapter is not intended as a full survey of statistical methods but rather as an introduction to some techniques of interest in electronic component production and measurements.

To provide a range of values to work on, consider the measurements made on a batch of 1000 Ω ± 2% resistors as in Table 8.1. All figures are rounded *down* to the nearest ohm, that this is a reasonable approximation becomes apparent later. The figures are said to refer to a *sample* of the *population* (i.e. total number) of a particular batch of resistors and are taken so that calculations on the sample may enable reasonable predictions to be made about the whole batch.

Table 8.1: MEASUREMENTS ON A SAMPLE OF 60 1000 Ω ± 2% RESISTORS (OHMS)

1009	987	1014	986	995	1005	998	1006	1001	1011
996	1008	990	1000	1023	1000	1003	999	979	995
1028	1012	1011	987	991	1006	1007	993	994	1001
1017	984	1003	991	987	997	1015	985	1003	998
983	1013	997	990	1004	1002	974	1008	1015	995
993	1001	1002	991	975	985	981	981	997	989

As in mathematics generally, the Greek capital sigma, Σ, is used as shorthand for "the sum of" and if x is used as a general symbol for the value of any given item in the sample, then Σx represents the sum of all the individual values. Let there be n of these.

8.1 MEAN AND STANDARD DEVIATION

The *mean* or more colloquially, the *average* of a set of values is simply the sum of all those values divided by their number, i.e. $(\Sigma x)/n$. The sample mean is denoted by $\bar{x}$ (x bar or barred x) and it is unlikely to be the same as the population mean, only an estimate of it. Hence:

$$\bar{x} = \frac{\Sigma x}{n} \tag{8.1}$$

and for Table 8.1, $\Sigma x = 59891$, $n = 60$, $\bar{x} = 998.18$.

The mean is one way of representing a whole set by a single figure. Although frequently used, it can lead to wrong conclusions for considering, as a simple example, the two sets of figures 8, 9, 10 and 3, 7, 11, 15, both have a mean of 9, yet clearly the sets differ widely. In the second there is more *dispersion*, i.e. the extent to which the individual values differ from the mean. In assessing dispersion first thoughts might suggest taking the average deviation from the mean but this gives an answer of zero because the very position of the mean results in positive deviations cancelling out the negative. Averaging the squares of the deviations makes everything positive and so removes this complication but because a measure of the deviations is required, not of the squares, the square root is finally taken. The result of all this is the *standard deviation*, denoted by the Greek lower-case sigma, σ .

The standard deviation is therefore defined as the square root of the average value of the squares of the deviations from the mean of the original values. It is an r.m.s. quantity and mathematically:

$$\sigma = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n}} \tag{8.2}$$

EXAMPLE 8.1:
Calculate the mean and standard deviation of the two sets of figures 8, 9, 10 and 3, 7, 11, 15 .

First set

Numbers (x)	Deviation from mean, $(x - \bar{x})$	$(\text{Deviations})^2$ $(x - \bar{x})^2$
8	−1	1
9	0	0
10	+1	1
$\Sigma x = 27$	$\Sigma(x - \bar{x})^2 = 2$	

where $n = 3$, mean, $\bar{x} = 27/3 = 9$, standard deviation, $\sigma = \sqrt{2/3} = 0.82$.

Second set

3	−6	36
7	−2	4
11	+2	4
15	+6	36
$\Sigma x = 36$	$\Sigma(x - \bar{x})^2 = 80$	

where $n = 4$, mean, $\bar{x} = 36/4 = 9$, standard deviation, $\sigma = \sqrt{80/4} = 4.47$.

8.1.1 Degrees of Freedom

Taking the extreme, a sample consisting of one item only is useless because no estimate of the standard deviation can be made. It is not until a second item is selected that a meaningful calculation is possible. With a sample of n items therefore, only $(n - 1)$ actually contribute to the result and statisticians refer to this as the number of *degrees of freedom*. When n is large, the difference between n and $(n - 1)$ is small so the point may seem trivial, but it is made because some scientific calculators which take over the tedious work involved in calculating means and standard deviations may be found to use $(n - 1)$ degrees of freedom or *weighting*, i.e.

$$\text{standard deviation, } s = \sqrt{\frac{\Sigma(x - x)^2}{n - 1}} \qquad (8.3)$$

and when this formula is used s represents the s.d. instead of σ to show the difference. When reasonable we ourselves avoid this extra complication.

EXAMPLE 8.2:
The set of figures in Table 8.1 is a sample taken from a production line for resistors. What is the mean resistance value and what are the standard deviations, σ and s?

Calculation directly as shown in Example 8.1 is tedious and is usually carried out with scientific calculator help (many calculators provide a mean and s.d. facility). There are short-cut methods (see Sect.8.2.1) but even with these much arithmetic is involved. Thus, via a calculator into which all 60 resistance values have been entered,

$$\text{mean } (\bar{x}) = \frac{59891}{60} = 998.18 \text{ ohms,}$$

standard deviation, $\sigma = 11.55$ ohms and $s = 11.65$ ohms.

8.2 HISTOGRAMS

The example above already gives some form to the mass of figures in Table 8.1, at this stage the s.d. may not seem to convey much information but to the expert who handles

Table 8.2: DISTRIBUTION OF RESISTANCE VALUES

	1	2	3	4	5	6	7	8	9	10	11	12
Class Intervals (Ω)	970–974	975–979	980–984	985–989	990–994	995–999	1000–1004	1005–1009	1010–1014	1015–1019	1020–1024	1025–1029
Frequency (f)	1	2	4	7	8	10	11	7	5	3	1	1

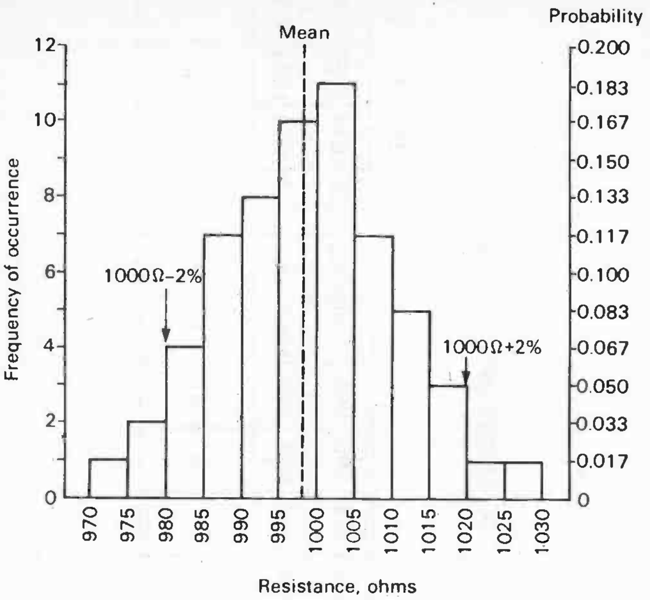


Fig. 8.1 Histogram showing frequency distribution

production line sampling it does. We ourselves will find it of use later.

A pictorial way of seeing Table 8.1 is by the *histogram*. Firstly a *frequency of occurrence* table is drawn up by grouping the data into a number of equal *class intervals* (in this case resistance ranges) and the frequency or number of times that values occur within the interval is recorded as shown in Table 8.2.

From these results a histogram (from Greek, histos = mast) is constructed as in Figure 8.1. This has the class intervals extending along the x-axis with the frequency per class interval indicated by the height of the pillar. There is little doubt that this pictorial representation tells more than a batch of figures. Note that gaps do not occur between the class intervals, for example between 974 and 975 in the first two because in fact 974 includes all values up to 974.999 . . . . Alternatively each class may be labelled with the centre value of the resistance range, e.g. 972.5, 977.5 etc.

8.2.1 Simplified Calculations

When a scientific calculator or computer is not available, any means of reducing the volume of arithmetical work is welcome. Even then an ordinary calculator is invaluable although as we will see below, not essential. Several methods of reducing the work involved in calculating a standard deviation are available, usually by manipulation of the formula

$$\sigma = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n}} \quad \text{or} \quad s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$

(Sect.8.1.1). For example it can be shown that the latter can be expanded to

$$s = \sqrt{\frac{\Sigma x^2}{n-1} - \frac{(\Sigma x)^2}{n(n-1)}} \ ,$$

a formula suitable for the ordinary calculator because it requires mainly totals of the individual values and their squares. There are other formulae equally useful which work with no approximation, however more effort is saved if some small inaccuracy is accepted. The work is reduced by calculating in classes as used in the section above (Table 8.2). The inaccuracy arises from the assumption that the values in each class are distributed evenly so that their mean is equal to the centre value. Although this is unlikely to be the case, to a certain extent the class inaccuracies cancel out leaving a relatively insignificant error.

The number of class intervals must firstly be chosen, in this example 12 are used although generally a smaller number is sufficient, say 6 or more for values of n up to about 30 rising to 9–12 when n reaches some 200 or more. The important point is to choose class intervals which are conveniently arithmetically. Secondly it is possible to work from a *fictitious mean* to reduce the arithmetic in each calculation, readjusting at the last stage. Let u represent any class position relative to the mean, what this actually implies in practice will become apparent as we progress through a typical calculation.

EXAMPLE 8.3:
Find the mean and standard deviation for the values in Table 8.1 using a fictitious mean as a short-cut method.

Table 8.2 already assembles the data of Table 8.1 in classes with the frequency of occurrence of each. It is repeated here in Table 8.3 with the additional calculations added. Recalling that in this particular case each of the resistance figures is rounded down to the nearest ohm, then the class interval, $C = 5\ \Omega$. First a class is chosen which appears to contain the mean. In this case the mean *ought* to be around 1000 Ω, hence class 6 seems most appropriate, accordingly its own mean value is noted, i.e. 997.5 Ω. This is the class fictitious mean $\bar{x}_f$. Choosing wrongly does not affect the issue, but if we do, some of the figures become larger.

From the totals column in the Table:

$$\Sigma f = 60\,, \quad \Sigma fu = 16\,, \quad \Sigma fu^2 = 332\,.$$

It can be shown that the mean,

$$\bar{x} = \bar{x}_f + C\left(\frac{\Sigma fu}{\Sigma f}\right) \tag{8.4}$$

i.e. the fictitious mean is adjusted according to the degree by which it exceeds or falls short of the true mean and is then multiplied by C to restore class intervals to resistance values.

Table 8.3: SIMPLIFIED CALCULATION OF MEAN AND STANDARD DEVIATION

Fictition mean class chosen = 6 ↓Mid-point = 997.5

Class No.	1	2	3	4	5	6	7	8	9	10	11	12	TOTALS
Class interval (Ω)	970–974	975–979	980–984	985–989	990–994	995–999	1000–1004	1005–1009	1010–1014	1015–1019	1020–1024	1025–1029	
Frequency, f	1	2	4	7	8	10	11	7	5	3	1	1	60
Deviation from fictitious mean, u	–5	–4	–3	–2	–1	0	1	2	3	4	5	6	
fu	–5	–8	–12	–14	–8	0	11	14	15	12	5	6	–47 + 63 = 16
u^2	25	16	9	4	1	0	1	4	9	16	25	36	
fu^2	25	32	36	28	8	0	11	28	45	48	25	36	322

Also the s.d.

$$\sigma = C \sqrt{\frac{\Sigma f u^2}{\Sigma f} - \left(\frac{\Sigma f u}{\Sigma f}\right)^2} \tag{8.5}$$

again in class intervals, adjusted and multiplied by C. Substituting in Equation (8.4):

$$\bar{x} = 997.5 + 5\left(\frac{16}{60}\right) = 997.5 + 1.33 = 998.83$$

(Sect.8.1.1 gives the accurate answer as 998.18). From Equation (8.5):

$$\sigma = 5\sqrt{\frac{322}{60} - \left(\frac{16}{60}\right)^2} = 5\sqrt{5.367 - 0.071}$$

$$= 11.51$$

(Sect.8.1.1 gives 11.55). The error is less than 0.5% in both calculations.

8.3 PROBABILITY DISTRIBUTIONS

Probability is the essence of prediction. In everyday life we might predict that an event is *likely* to happen but in statistics the possibility has to be expressed precisely. As mentioned earlier such *prediction* assumes some *uncertainty* and the methods not only reduce the uncertainty but also quantify it.

8.3.1 Probability

The scale of measurement is simple. The impossibility of an event occurring is rated as zero and certainty as 1.0. Few events are in these categories but fall somewhere in between, hence a probability figure of between 0 and 1 is assigned. For example, a coin tossed properly (i.e. with no *bias*) is as likely to produce a head as a tail, the probability of a head is expressed as $P(\text{H}) = 0.5$, equally the probability of a tail is $P(\text{T}) = 0.5$. $P(\text{H}) + P(\text{T}) = 0.5 + 0.5 = 1$ showing the certainty of either a head or tail. Taking the roulette wheel as another example, with any of its numbers 0–36 as a possible result, the probability of any one is 1 in 37, i.e. 0.027.

8.3.2 Distributions

Looking again at Figure 8.1 and taking as an example the range of resistance values 1000–1005, of the 60 measurements, 11 are contained in this range. The probability of a resistor having a value in the range is therefore 11 in 60 or 0.183. The calculation is similar for each of the other ranges and the probabilities are marked on a scale to the right. With this scale instead of the frequency scale, Figure 8.1 is correctly termed a *probability distribution*. The probability scale is evidently the frequency scale divided by the number of items, n.

EXAMPLE 8.4:
From Figure 8.1, what is the probability of a resistor being outside of the intended range 1000 Ω ± 2%?

The probability of a resistor being in the range

970– 975	is 0.017
975– 980	is 0.033
1020–1025	is 0.017
1025–1030	is 0.017

The probability of being outside the 2% limits is the sum of these, i.e. 0.084, seemingly low, but in fact predicting that 84 resistors in every batch of 1000 will be outside of tolerance limits. Note that the probability of *not* being outside of the limits is 1 – 0.084 = 0.916.

8.3.2.1 The Normal Distribution Curve

Table 8.1 represents a small sample, a larger one would have better predictive powers. The true answer is only given when *all* items are tested but this is exactly what is being avoided because of the amount of work involved. What is possible instead is to calculate a theoretical curve from the results of the sample. Gauss(A3) was one of the first to study how curves can be fitted to histograms with minimum error. They are appropriately known as *Gaussian curves* and by far the one most frequently utilized is the *Normal Distribution Curve* which is symmetrical about the mean value and seems to suit many investigations. Referring to the histogram of Figure 8.1 for which the original aim was to produce resistors of 1000 Ω it is evident that the probability of error decreases as the error itself increases. In the normal Gaussian curve which expresses this the presumption is also made that the elementary errors which make up the total error at each point

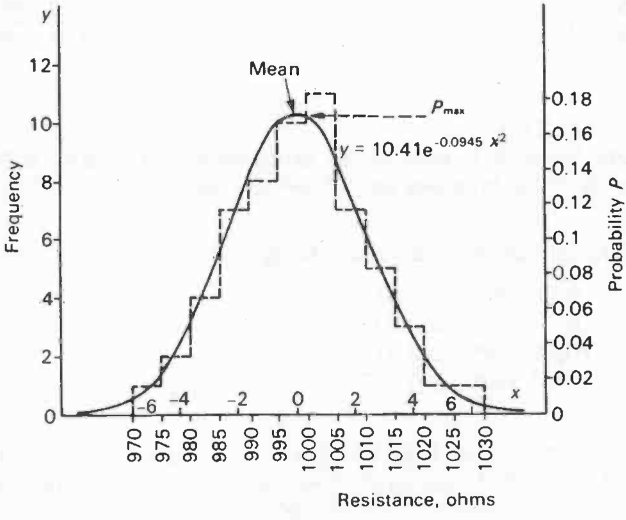


Fig. 8.2 Normal distribution curve

have occurred by pure chance. That this is reasonably so in many manufacturing processes accounts for the favourable reception of the curve.

Figure 8.2 shows the outline of the histogram of Figure 8.1 with a best fit normal distribution curve added. The latter is therefore a smooth curve which, derived only from the measurements on the sample, predicts the probability of the various values arising in the total population. The mathematical analysis published by Gauss is somewhat complicated but it results in a general equation:

$$y = \frac{1}{\sigma\sqrt{2\pi}} \times e^{(-x^2)/(2\sigma^2)} \tag{8.6}$$

where in this case y represents probability and x is in terms of class intervals. The curve of Figure 8.2 is calculated from this formula.

From this equation a simplified method of producing the curve can be developed based on the fact that it is defined by two parameters only, the mean and the standard deviation. As seen from Figure 8.2 the maximum probability P_{max} occurs at the mean and from the formula it has a value

$$P_{max} = \frac{1}{\sigma\sqrt{2\pi}}, \text{ i.e. } \frac{0.3989}{\sigma} \tag{8.7}$$

(for when $x = 0$, $e^{(-x^2)/(2\sigma^2)} = 1$).

Next let us introduce the symbol z to represent a number of standard deviations above or below the mean, then $x = z\sigma$ and probability, $P = P_{max} \times e^{(-z^2)/2}$,

so at 1 s.d. from the mean:

$$(z = 1), \quad P_1 = P_{max} \times e^{-0.5} = P_{max} \times 0.6065$$

and at 2 s.d. from the mean

$$(z = 2), \quad P_2 = P_{max} \times e^{-2} = P_{max} \times 0.135$$

and at 3 s.d. from the mean

$$(z = 3), \quad P_3 = P_{max} \times e^{-4.5} = P_{max} \times 0.011$$

already sufficient points to sketch in the curve. More points can be added as required.

In Figure 8.1 the area of each pillar relative to the total area of the histogram represents the proportion of items in the sample having values between those stated at the base of the pillar. As an example, for 1005 – 1010 Ω (strictly to 1009.999 . . .) the area is 7/60 of the total simply because the frequency is 7 out of a total of 60. For the normal distribution curve of Figure 8.2 (which is in fact the histogram idealized) the same therefore applies, i.e. the area under the curve over a certain resistance range represents the proportion of resistors within that range. The most interesting and useful feature is that irrespective of the values of the mean and s.d., that proportion of the total population which occurs between the mean and a specified number of s.d.'s away from it is constant. What this means is shown pictorially in Figure 8.3(i) and (ii) which show two completely different distributions [there is more dispersion in (ii)], each with a shaded area extending over, for example, one s.d. from the mean. In both cases the shaded areas represent the same proportion of the total area under the curve (here 0.341 or 34.1%). Accordingly it can be said that in (i) 34.1% of the items are likely to fall within the range 18–20 and because of the symmetry of the curve, also from 20–22. In (ii) the range is 20–24, also implying 16–20.

Provided that the equation to the curve is known then mathematically the area under it between any two points on the x-axis can be calculated. This has been done for a range of value of z from published standard *error function* tables which relate the probability of error (i.e. digression from the mean) with the various ranges, expressed as a curve in (iii) of Figure 8.3. The value of the technique is shown by the examples which follow and a further time-saving feature is that the curve itself need not be drawn.

EXAMPLE 8.5:

Measurements on a sample of 1000 Ω resistors from a production line show a mean of 1000 Ω with a s.d. of 10 Ω. What is the probability that a resistor selected from the production at random will have a value between 1000 and 1010 Ω?

This is simple because 1000–1010 extends from the mean over 1 s.d. and Figure 8.3(i) and (ii) show that the probability is 0.341.

EXAMPLE 8.6:
With mean and s.d. as in the example above, what is the proportion of resistors likely to lie in the range 995–1005 Ω.

Although 995–1005 = 10, z is not equal to 1 because it is always in terms of s.d.'s *from the mean.* Take first 1005, it is 1005–1000 = 5 Ω from the mean, hence $z = 5/10 = 0.5$. From Figure 8.3(iii), probability = 0.1915, i.e. approximately 19% lie in the range 1000–1005. Equally because the normal distribution curve is symmetrical about the mean, 19% lie in the range 995–1000.

Thus proportion within range 995–1005 = 38%.

EXAMPLE 8.7:
With mean and s.d. as in the above two examples, how many resistors per 1000 may be expected to have values outside of the tolerances ± 1%, ± 2%, ± 3%?

The calculations are as shown in Table 8.4 (overleaf).

The last three examples demonstrate the use of the technique in production control. Although the normal distribution is most common, others may be used when the histogram is *skew* (not symmetrical). To sum up, Table 8.5 expands on the last example to indicate the percentage of a production batch likely to fall outside of various ranges.

8.3.3 Sample Size

Throughout the foregoing sections a sample of 60 items has been used to predict the outcome of a production run many

Table 8.4: CALCULATIONS FOR EXAMPLE 8.7

Tolerance	Range	z	Probability [Fig.8.3(iii)]	Total probability	No. likely to have values within range	No. likely to have values outside range
± 1%	990–1010	+1 –1	0.3413 0.3413	0.6826	683	317
± 2%	980–1020	+2 –2	0.4772 0.4772	0.9544	954	46
± 3%	970–1030	+3 –3	0.4987 0.4987	0.9974	997	3

Table 8.5: **PERCENTAGES OF ITEMS FALLING OUTSIDE GIVEN RANGES FOR NORMAL DISTRIBUTION**

Range, in terms of $\pm \sigma$	Percentage of items falling outside range	Range, in terms of $\pm \sigma$	Percentage of items falling outside range
0	100	1.6	11.0
0.2	84.1	1.8	7.2
0.4	68.9	2.0	4.6
0.5	61.7	2.2	2.8
0.6	54.9	2.4	1.6
0.8	42.4	2.5	1.2
1.0	31.7	2.6	0.9
1.2	23.0	2.8	0.5
1.4	16.2	3.0	0.3
1.5	13.4	3.1	0.2

times greater. The question remains as to what size a sample should be, evidently as small as is practicable consistent with producing a realistic forecast. The subject is complicated, nevertheless we can look at the calculations involved in one important facet, the relationship between *percentage confidence* and the *forecast tolerance*, a concept which may be difficult to appreciate at first but which will be clarified by examples.

Confidence is a human impression which in statistics has to be expressed numerically. 100% confidence means that we know something *will* happen, less means that there is an element of doubt. From Figure 8.3(iii) the probability that an item will have a value within ± 2 s.d.'s of the mean is 2 × 0.4772, i.e. approximately 0.955. Equally it can be said that there is 95.5% confidence that a value will lie within ± 2 s.d.'s. But there is still a 4.5% chance that it will not. Because this type of investigation involves prediction from a sample and there cannot be certainty, answers can only tell the whole story if the particular level of confidence is also quoted.

We will see later that the higher the level of confidence, the less precise is the result hence a compromise is necessary

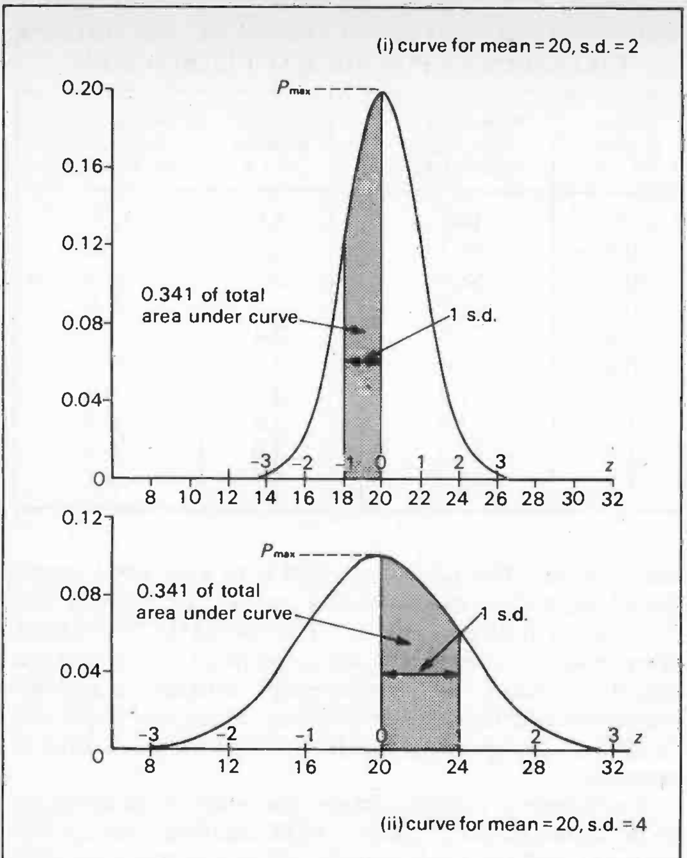
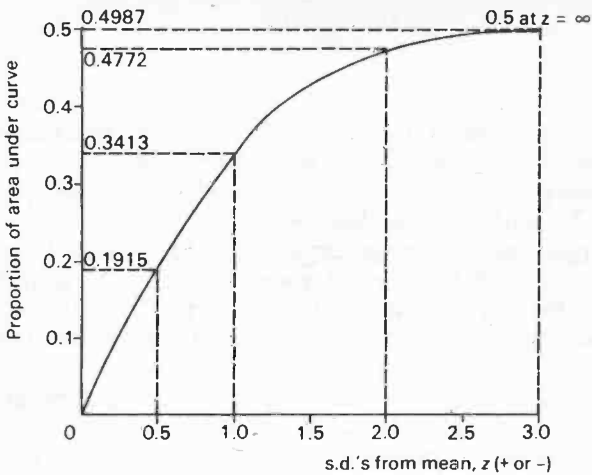


Fig. 8.3 Normal distribution curves

and a commonly used one is the 95.5% level mentioned above.

A formula has been developed from which are derived the *confidence limits* (*CL*) on a population percentage, P , it is:

$$CL = \pm \sigma \sqrt{\frac{P(100 - P)}{n}} \qquad (8.8)$$



(iii) Proportion of area under curve between mean and z s.d.'s above or below mean

Fig. 8.3 (contd)

P however is the unknown, all we have is the sample estimate of it, S. However, by substituting S for P in the formula the loss of accuracy is generally tolerable so we use:

$$CL = \pm\sigma\sqrt{\frac{S(100-S)}{n}} \tag{8.9}$$

What this means in practice can be seen by considering the resistor measurements of Table 8.1 and the fact discovered in Section 8.3.2 that 8.4% of the sample is outside the ± 2% tolerance. It is the accuracy of the figure of 8.4% which interests us. So far this one figure is all we have, now with the formula the range within which the true answer lies can also be estimated. The range is, from Equation (8.9)

$$S \pm \sigma\sqrt{\frac{S(100-S)}{n}}$$

so for 95.5% confidence it is

$$8.4 \pm 2\sqrt{\frac{8.4(100-8.4)}{60}}\ \%$$

i.e. (8.4 ± 7.2)% or 1.2% – 15.6%, the figures being known as the *95.5% confidence limits.* Summing up therefore, the percentage of resistors failing to be within ± 2% tolerance is 8.4 ± 7.2, with 95.5% confidence.

Estimating that the percentage may be between 1.2 and 15.6 may seem to be a pretty useless result but Figure 8.1 shows that it is based on the fact that 5 resistors only out of 60 are responsible. Further samples of 60 resistors might easily produce several more or less than 5 outside of tolerance, hence the large range within which the true percentage is expected to lie. This one sample constitutes the only information available so there is little to be gained by juggling with confidence levels and ranges for one can only be traded for the other. At 80% confidence, $\sigma = 1.28$ and the range is 8.4 ± 4.6% so although the answer is more precise, there is less confidence in it. Alternatively with increased confidence the limits widen, e.g. at 98% confidence the range is 8.4 ± 8.3%.

The only way of decreasing the range at a given level of confidence is to provide more information by increasing n, i.e. taking a larger sample.

EXAMPLE 8.8:

Following an advertising campaign for a new type of computer, 600 people were interviewed to ascertain whether they had seen the advertisements. 252 answered "yes". This is 42%. For 95.5% confidence in what range does this percentage actually lie?

$$S = 42, \quad n = 600, \quad \text{for 95.5\% confidence, } \sigma = 2$$

From Equation (8.9)

$$\text{Percentage range} = S \pm \sigma\sqrt{\frac{S(100-S)}{n}}$$

$$= 42 \pm 2 \sqrt{\frac{42 \times 58}{600}} = 42 \pm 4\%$$

therefore the true percentage is most likely to be lie within the range 38–46%.

EXAMPLE 8.9:
In the above example what sample size would be required to reduce the confidence limits to ± 2.5%?

From Equation (8.9)

$$2 \sqrt{\frac{S(100 - S)}{n}} = 2.5 \qquad \therefore n = \frac{42 \times 58}{1.5625} \approx 1560$$

thus 1560 people must be interviewed, i.e. about 2.6 times as many. Note that (i) before this last estimate of n can be made, some idea of S is required. This may therefore entail conducting a small investigation first although this is not as onerous as it may seem because the results may be used as part of the final work.

(ii) With the larger sample, S will not necessarily be 42% but some value near. At the stated confidence level however it will have limits on it of only ± 2.5%.

8.4 SUMMARY OF KEY FORMULAE

n = number of items

S = sample estimate of a population percentage

x = value of one item

Σ = the sum of

σ = standard deviation.

QUANTITY	FORMULA	UNIT	SECTION
Sample mean	$\bar{x} = \dfrac{\Sigma x}{n}$		8.1
Standard deviation	$\sigma = \sqrt{\dfrac{\Sigma(x - \bar{x})^2}{n}}$		8.1
Normal distribution curve	$y = \dfrac{1}{\sigma\sqrt{2\pi}} \times e^{(-x^2)/(2\sigma^2)}$		8.3.2.1
Confidence limits	$CL = \pm\sigma\sqrt{\dfrac{S(100 - S)}{n}}$	%	8.3.3

9. RELIABILITY

The importance of reliability hardly need be stressed nor need the general concept be explained. What this Chapter attempts to do is to highlight some of the basic formulae through which reliability can be expressed as a number and therefore with greater practical meaning. Through tests in advance and/or by observation of working equipment reasonable estimates of the reliability of a component or system can be made. Evidently statistics and probability calculations (Chapter 8) are needed for such predictions.

Reliability can be defined in this context as the probability that a component or system will function (or survive) within previously determined limits and operating conditions for a specified period of time. To express reliability numerically we make use of the term *probability of survival* (P_s) which has a range (as in Sect.8.3.1) of 0 to 1. Thus a reliability of 0 infers that an equipment will not function, of 1 that it will continue to function without failure (both are inconceivable). Practical systems obviously need reliabilities very much nearer to 1 than to 0.

9.1 ASSESSMENT OF FAILURE

Usually a fundamental to the concept of reliability is the factor of time for clearly the longer an item or system functions without failure, the greater is its reliability. However time is not used to the exclusion of all else for reliability could, for example, be assessed on the number of operations as in the case of a mechanical switch for which failure through wear may better be assessed with use rather than time. Two commonly used measurements of failure involving time (but not a specified period of time as is required in the definition of reliability) are as follows.

9.1.1 Mean Time to Failure

This is abbreviated to *m.t.t.f.* and is defined as the operating time of a single item of equipment divided by the total number of failures during that time. If several similar items are tested, each will have a different time to failure. The m.t.t.f. is then the sum of all the times for all the items divided by the number of items. Thus for n items, if t_0 is the time of commencement of operation (or of a trial) and t_1, t_2, t_3 . . . t_n are the times of the 1st, 2nd, 3rd . . . nth failures.

$$\text{m.t.t.f.} = \frac{(t_1 - t_0) + (t_2 - t_0) + (t_3 - t_0) + \ldots (t_n - t_0)}{n} \tag{9.1}$$

M.t.t.f. is especially useful with discrete components which are not repaired when they fail.

EXAMPLE 9.1:
Four similar items of equipment are put on test and failures occur after 37, 50, 43 and 30 hours. What is the m.t.t.f.?

From the above formula:

$$\text{m.t.t.f.} = \frac{37 + 50 + 43 + 30}{4} = \frac{160}{4} = 40 \text{ hours}.$$

9.1.2 Mean Time Between Failures

This is also known as mean time before failure and in both cases the abbreviation *m.t.b.f.* applies. It is a very useful indicator of reliability and is usable with repairable items or systems, the time in each case being simply that between two successive failures. For a single item of equipment suffering f failures, if t_0 is the time of commencement and t_1, t_2, t_3 . . . t_n are the times of the 1st, 2nd, 3rd . . . nth failures

$$\text{m.t.b.f.} = \frac{(t_1 - t_0) + (t_2 - t_1) + (t_3 - t_2) + \ldots (t_n - t_{n-1})}{f}$$

$$= \frac{t_n - t_0}{f} \tag{9.2}$$

Quite simply the time t_n is observed and this is divided by the number of failures and generally m.t.b.f. for more than one item is defined as

$$\text{m.t.b.f.} = \frac{\text{Number of items x operating time}}{\text{Total number of failures}}$$

$$= \frac{n \times t}{f} \tag{9.3}$$

The reciprocal of this gives the *failure rate*, λ , hence:

$$\text{m.t.b.f.} = \frac{1}{\lambda} \tag{9.4}$$

and

$$\lambda = \frac{f}{n \times t} \tag{9.5}$$

Multiplying λ by 100 gives the percentage failure rate:

$$100\,\lambda = \frac{100f}{n \times t} \tag{9.6}$$

EXAMPLE 9.2:
An item of equipment fails and is repaired on the 39th, 81st, 116th and finally on the 148th day after switching on. What is its m.t.b.f.?

Total number of failures, $f = 4$. From Equation (9.2)

$$\text{m.t.b.f.} = \frac{39 + (81-39)+(116-81)+(148-116)}{4}$$

$$= \frac{148}{4} = 37 \text{ days}$$

or simply from

$$\frac{t_n - t_0}{f} = \frac{148 - 0}{4} = 37 \text{ days} .$$

In reliability jargon, to cater for the very high reliabilities of some modern components, one *fit* (*f*ailure un*it*) represents one failure in 10^9 hours, its use possibly superseding that of the *bit* (nothing to do with the digital bit) which represents one failure in 10^8 hours. (Appendix 1 gives failure rate conversion factors.)

EXAMPLE 9.3:
If there are 2000 IC's in a small electronic telephone exchange and they have a failure rate of 57 fit, how frequently on average will one have to be changed?

Total expected failures in 10^9 hours = 57 x 2000. Therefore the average time between failures is equal to

$$\frac{10^9}{57 \times 2000} = 8772 \text{ hours} = 1 \text{ year} .$$

EXAMPLE 9.4:
It is estimated that a particular type of IC has an m.t.b.f. of 100 years. Would it be suitable for a range of calculators for which its failure rate should be better than 1000 fit?

$$\text{m.t.b.f.} = 100 \text{ years} = 100 \times 24 \times 365 \text{ hours}$$

$$= 876{,}000 \text{ hours} .$$

From Equation (9.4):

$$\text{failure rate} = \frac{10^9}{\text{m.t.b.f.}} \text{ fit} \quad \text{(where m.t.b.f. is in hours)}$$

$$= \frac{10^9}{876{,}000} = 1142 \text{ fit} .$$

The IC would therefore *not* be acceptable.

On first thoughts an m.t.b.f. of 100 years may seem remarkably good. It cannot be overemphasized however that the prediction is for one fault to occur every 100 years *on average.* By chance one or two faults could occur as early as in the first few years. An m.t.b.f. of 100 years does *not* mean that no fault will occur until 100 years have passed.

Nowadays IC's are proving to have reliabilities greatly in excess of what might be expected when one considers their complexities. Expectations are that IC's for certain applications will have failure rates as low as 10–20 fit, implying m.t.b.f.'s of 11,400 – 5700 years. Such figures tend to make reliability calculations appear non-sensical and bearing in mind that most equipment has a life span of a few decades only it would seem that such predictions have no practical parallel. This is certainly not the case as the single example of a large electronic telephone exchange employing say, 100,000 IC's each with a failure rate of 20 fit will show.

The m.t.b.f. is $10^9/20$ hours per IC, therefore $10^9/(20 \times 100{,}000)$ hours for the whole exchange, i.e. 3 weeks. Although therefore each IC seems to have an incredibly high m.t.b.f. of over 5,000 years, *on average* in this particular exchange, one is likely to fail every 3 weeks, a very practical consideration indeed.

9.1.3 Measurement

Such times as mentioned in the Section above highlight the difficulties encountered in measuring failure rates of modern components. Accelerated life tests may be used in which the failure rate is deliberately increased and the artificially high results are subsequently corrected by special techniques. Extreme care is necessary otherwise the abnormal conditions may activate failure mechanisms other than those experienced under normal working conditions.

Tests are conducted by operating a batch of components or units, possibly over several thousand hours and recording times at which failures occur.

9.2 ASSESSMENT OF RELIABILITY

Not unexpectedly certain assumptions have to be made in order to obtain a simple mathematical formula for reliability and one which is applicable and also applies reasonably in practical situations is that the failure rate, measured over a sufficiently long period, is constant. Given this we can arrive at a formula without too much complication.

9.2.1 The Exponential Law

A useful comparison is with the discharge of a capacitor (Book 1, Sect.3.3.6) which has an exponential form because the discharge current at any instant depends on the charge remaining. With a batch or population of similar components put into use or on test, the number of good components left falls with time as failures occur. Accordingly we make the capacitor charge analogous to the good components. Also, given a constant failure rate, the number of failures occurring during any given small time interval is controlled by the number of good components from which they arise. This process is therefore akin to the capacitor discharge current hence reliability can be expressed by a similar exponential function:

$$\text{Reliability (with time)}, R_t = e^{-t/m} \qquad (9.7)$$

where m.t.b.f. is further abbreviated to m, and in terms of the failure rate:

$$R_t = e^{-\lambda t} \qquad (9.8)$$

EXAMPLE 9.5:
Components put on test for 1 year are found to have an m.t.b.f. of 10 years. What is their reliability? To what must the m.t.b.f. be increased for a reliability of 0.99?

$$t = 1 \text{ year} \qquad m = 10 \text{ years} \qquad \frac{t}{m} = 0.1 .$$

From Equation (9.7):

$$\text{Reliability, } R_t = e^{-0.1} = 0.905 .$$

When $R_t = 0.99$, from Equation (9.7):

$$0.99 = e^{-1/m} \qquad \therefore \log_e 0.99 = -1/m$$

$$\therefore \qquad -0.01 = -1/m \qquad \therefore m = 100 \text{ years},$$

so demonstrating the order of m.t.b.f. required for high reliability.

EXAMPLE 9.6:
5000 rectifier diodes were put on test for 50 weeks during which time 105 failures occurred. What are the failure rate, m.t.b.f. and component reliability?

Number of items, $n = 5000$

Number of failures, $f = 105$

Operating time, $t = 50 \times 7 \times 24 = 8400$ hours

From Equation (9.5):

$$\text{failure rate, } \lambda = \frac{f}{nt} = \frac{105}{5000 \times 8400} = \frac{2.5}{10^6}$$

$$= \frac{2500}{10^9}, \text{ i.e. } \lambda = 2500 \text{ fit (or 250 bit)}.$$

$$\text{m.t.b.f.} = \frac{1}{\lambda} = \frac{10^6}{2.5} = 4 \times 10^5 \text{ hours}$$

$$\frac{t}{m} = \lambda t = 2.5 \times 10^{-6} \times 8400 = 0.021.$$

From Equation (9.8):

$$\text{Component reliability, } R_t = e^{-\lambda t} = e^{-0.021} = 0.979.$$

9.2.1.1 System Reliability

So far emphasis has been on batches of similar components. Seldom does a system use only one type hence for an estimate of *system* reliability the effects of the numbers and failure rates of all components must be included. Suppose that within a time interval, t, $n_1, n_2, n_3 \ldots$ components with failure rates $\lambda_1, \lambda_2, \lambda_3 \ldots$ give rise to $f_1, f_2, f_3 \ldots$ failures. Then from Equation (9.5):

$$\lambda = \frac{f}{n \times t} \qquad \therefore \frac{f}{t} = n\lambda$$

therefore

$$\frac{f_1 + f_2 + f_3 + \ldots}{t} = n_1\lambda_1 + n_2\lambda_2 + n_3\lambda_3 + \ldots$$

The left-hand expression represents the total failures in time t, i.e. the *system failure rate*, λ_s, hence

$$\lambda_s = n_1\lambda_1 + n_2\lambda_2 + n_3\lambda_3 + \ldots \qquad (9.9)$$

EXAMPLE 9.7:
An equipment, designed for high reliability comprises 40 transistors having an estimated failure rate of 32 fit, 80 diodes at 71 fit, 30 ceramic capacitors at 18 fit and 85 carbon resistors at 5 fit. Ignoring wiring connexions etc. what are the expected system failure rate, m.t.b.f. and reliability, the latter over a 10-year period?

From Equation (9.9), the system failure rate,

$$\lambda_s = (40 \times 32) + (80 \times 71) + (30 \times 18) + (85 \times 5)$$

$$= 1280 + 5680 + 540 + 425 = 7925 \text{ fit}$$

$$\text{m.t.b.f.} = \frac{1}{\lambda_s} = \frac{10^9}{7925} \text{ hours} = 1.26 \times 10^5 \text{ hours}$$

$$= 14.4 \text{ years} .$$

From Equation (9.7):

$$\text{Reliability } (t = 10 \text{ years}), R_t = e^{(-10)/14.4}$$

$$= e^{-0.694} = 0.5$$

(in plain English there is a fifty-fifty chance of survival without a fault during the 10-year period).

Note that this example is concerned with the components only. Reliability is likely to fall appreciably when wiring and connexions are also considered, for example, for 500 connexions with an estimated failure rate of 20 fit, the m.t.b.f. is reduced to 6.37 years and R_t to 0.21.

From the next example we can appreciate not only the value of reliability theory but also the need of high component reliability.

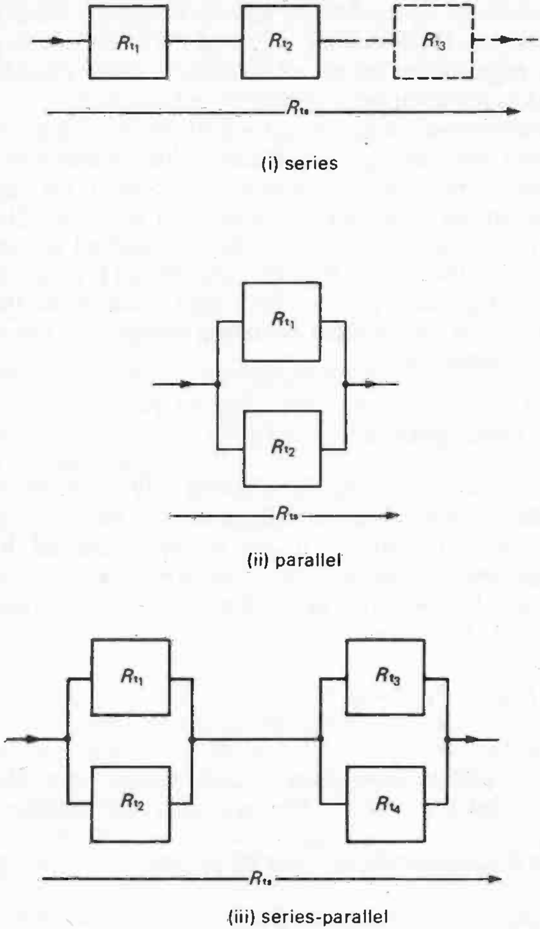


Fig. 9.1 Reliability of systems with redundancy

9.3.2 Parallel Systems

The simplest parallel arrangement is shown in Figure 9.1(ii) and consists of 2 units only. Generally,

$$R_{ts} = 1-(1-R_{t1})(1-R_{t2})(1-R_{t3})(\ldots \qquad (9.11)$$

so for two units only

$$R_{ts} = 1-(1-R_{t1}-R_{t2}+R_{t1}R_{t2})$$

$$= R_{t1}+R_{t2}-R_{t1}\times R_{t2} \qquad (9.12)$$

EXAMPLE 9.10:
Two similar units having reliabilities of 0.9 are connected in parallel. What is the overall reliability?

$$R_{t1} = 0.9\,, \quad R_{t2} = 0.9$$

$$R_{ts} = 0.9+0.9-0.81 = 0.99\,,$$

a considerable improvement.

EXAMPLE 9.11:
As above but for 4 units in parallel. This is simply 2 units in parallel, each of reliability 0.99.

$$\therefore \quad R_{ts} = 0.99+0.99-0.99^2 = 0.9999\,.$$

OR, from Equation (9.11):

$$R_{ts} = 1-(1-R_{t1})(1-R_{t2})(1-R_{t3})(1-R_{t4})$$

$$= 1-0.1^4 = 0.9999\,,$$

a greatly improved reliability but at some 4 times the cost.

9.3.3 Series-Parallel Systems

Although from the above it is evident that by running units in parallel (where this is electrically practicable) large gains in reliability are realized, there may be situations where only a series-parallel arrangement can be used. One employing 4 units is shown in Figure 9.1(iii). For this the series and parallel formulae combine to give:

$$R_{ts} = [1 - (1 - R_{t1})(1 - R_{t2})] \times [1 - (1 - R_{t3})(1 - R_{t4})] \qquad (9.13)$$

EXAMPLE 9.12:
Four units, each of reliability 0.9 are connected as in (iii) of the Figure. What is the overall reliability?

$$R_{ts} = (1 - 0.1^2)(1 - 0.1^2) = 0.99 \times 0.99 = 0.98 .$$

When compared with the previous example, the reduction in reliability on changing from full parallel to series-parallel is evident.

9.4 SUMMARY OF KEY FORMULAE

f = number of failures

m = m.t.b.f.

n = number of items

t = time

R_t = reliability (on the basis of time)

QUANTITY	FORMULA	UNIT	SECTION
Mean time to failure	$\text{m.t.t.f.} = \dfrac{(t_1-t_0)+(t_2-t_0)+\ldots}{n}$	as for t	9.1.1
Mean time between failures	$\text{m.t.b.f.} = \dfrac{t_n-t_0}{f} = \dfrac{n \times t}{f}$	as for t	9.1.2
Failure rate	$\lambda = \dfrac{f}{n \times t} = \dfrac{1}{\text{m.t.b.f.}}$	fit, bit, etc.	9.1.2
Percentage failure rate	$\% \lambda = \dfrac{100f}{n \times t}$	%	9.1.2
Reliability (with time and for constant λ)	$R_t = e^{-\lambda t} = e^{-(t/m)}$	a number between 0 and 1)	9.2.1
System failure rate	$\lambda_s = n_1 \lambda_1 + n_2 \lambda_2 + \ldots$	fit, bit, etc.	9.2.2.1
System reliability (two in series)	$R_{ts} = R_{t1} \times R_{t2}$	a number between 0 and 1)	9.3.1
System reliability (two in parallel)	$R_{ts} = R_{t1} + R_{t2} - R_{t1} \times R_{t2}$	ditto	9.3.2

10. AUDIO

Although "audio" embraces a wide range of topics, this Chapter is necessarily brief because the general principles of audio amplification are examined separately in Chapter 5 and what remains does not give much scope for development of formulae. Some aspects of the subject of *acoustics* are included however for fundamental to everything audio is the sound wave.

10.1 THE SOUND WAVE

Sound waves are set up whenever air (or any other substance) is in contact with a vibrating body. The air particles are alternately displaced forwards and backwards at the frequency of the vibration. In terms of the air itself it is said to undergo alternate *compressions* and *rarefactions.* Because of the elasticity of the air coupled with the inertia of its particles the disturbance is able to travel outwards as a *sound wave.* Such vibrational waves only become *sound* when intercepted by an ear and subsequently translated in the brain. For human beings the frequency range is from some 20 Hz to 20 kHz, tailing off as we get older.

10.1.1 Velocity

Compared with light, sound waves travel slowly. The speed is in fact determined by the rate at which each air particle can transfer energy to the next. Newton[(A3)] understood this and produced an early formula for the velocity of a sound wave in a gas from the latter's elasticity and density. He equated elasticity with pressure and so considered that:

$$\text{Velocity of sound wave} = \sqrt{\frac{\text{pressure}}{\text{density}}}\,.$$

The results however were not in accord with those from practical experiments and Laplace(A3) added a correction for the fact that the pressure changes within the gas were too rapid for the heat variations to disperse. The term *adiabatic* is used to describe such a condition and the necessary correction which is constant for any given gas is labelled γ and so the formula becomes:

velocity of sound wave in a gas,

$$c = \sqrt{\frac{\gamma P_0}{\rho_0}} \text{ m/s} \qquad (10.1)$$

where P_0 is the mean gas pressure in newtons/sq. metre (N/m^2), ρ_0 is the mean density in kilogrammes/cubic metre (kg/m^3) and γ is the correction factor for adiabatic conditions. γ can be calculated for any gas.

EXAMPLE 10.1:
What is the velocity of sound in air at atmospheric pressure (760 mm mercury) and at 0°C if γ for air equals 1.402, the density of mercury is 13600 kg/m^3 and that of air 1.293 kg/m^3?

$$P_0 = \frac{760}{1000} \times 13600 = 10336 \text{ kg/m}^2$$

$$= 10336 \times 9.807 \text{ newtons/m}^2$$

(see Appendix 1 for the correction factor, also Sect.1.1.1)

$$\therefore \quad P_0 = 101365 \text{ N/m}^2$$

$$\rho_0 = 1.293 \text{ kg/m}^3$$

$$\gamma = 1.402,$$

then from Equation (10.1)

$$\text{velocity}, c = \sqrt{\frac{1.402 \times 101365}{1.293}} = 331.5 \text{ m/s}$$

The velocity of sound in liquids, especially sea water is of major importance in sonar systems (underwater "radar"). Unfortunately the equations become more involved because temperature, depth and (of sea water) salinity, all have an effect. However, a general formula which gives a reasonably close answer is:

$$c = \sqrt{\frac{B_a}{\rho_0}} \text{ m/s} \qquad (10.2)$$

where B_a is the *bulk modulus of elasticity* in N/m^2, an intimidating term except for those with experience of mechanics. It can be taken as having some relationship with *compressibility*, a term with which we are better acquainted.

For fresh water $B_a = 2.18 \times 10^9$ N/m^2 and for sea water a general figure of 2.28×10^9 will suffice.

EXAMPLE 10.2:
What is the approximate velocity of sound waves in sea water which has a density of 1025 kg/m^3?

$$\rho_0 = 1025 \text{ kg/m}^3 \qquad B_a = 2.28 \times 10^9 \text{ N/m}^2$$

From Equation (10.2)

$$\text{velocity}, c = \sqrt{\frac{2.28 \times 10^9}{1025}} = 1491 \text{ m/s}$$

thus sound waves travel considerably faster in sea than in air as they do generally in liquids compared with gases.

10.1.2 Sound Pressure

In engineering some means of measuring the magnitude of a quantity is basic to success in its utilization. One measure of a sound wave is by the *loudness* perceived by a listener but human beings are notoriously fickle in such assessments. A parameter more suited to scientific measurement is that of wave *pressure*, that is, the alternating pressure (N/m^2) above and below normal. Sound waves are seldom of constant amplitude, in fact they are usually highly erratic but for simple analysis we consider a constant amplitude wave as might be generated by a loudspeaker fed with a pure tone.

The range of audible pressures is large, from some 2×10^{-5} N/m^2, a sound which is just audible, to one a million times greater at 20 N/m^2, a level of auditory discomfort. However with atmospheric pressure around 10^5 N/m^2 (Example 10.1), even the maximum sound pressure represents little disturbance of the air in its path.

Such a wide range suggests the use of logarithms and because the ear is also logarithmically inclined in that sound pressure changes are interpreted as loudness changes of much lower magnitude, this is the type of scale generally used. A logarithmic scale must be quoted relative to something and in this case the choice has been made of 2×10^{-5} N/m^2, the pressure which at 1 kHz is just audible although not by everybody, it is technically known as the *threshold of audibility*. Sound pressures quoted in this manner are therefore almost invariably positive and are known as *sound pressure levels* (SPL) in decibels, then the sound pressure level,

$$L_p = 20 \log_{10} \frac{p}{2 \times 10^{-5}} \text{ dB} \qquad (10.3)$$

where p is the r.m.s. sound pressure in N/m^2. The 20 multiplier is used because sound pressure is analogous to voltage.

EXAMPLE 10.3:
What is the SPL range equivalent to that of audible pressures i.e. 2×10^{-5} N/m^2 to 20 N/m^2?

From Equation (10.3)

$$L_{p(min)} = 20 \log \frac{2 \times 10^{-5}}{2 \times 10^{-5}} \text{ dB} = 0 \text{ dB}$$

$$L_{p(max)} = 20 \log \frac{20}{2 \times 10^{-5}} = 20 \log_{10} 10^6 = 120 \text{ dB},$$

i.e. the sound pressure level range is 0 – 120 dB.

EXAMPLE 10.4:
Measurements on the sound emitted by a horn show that at a certain distance an SPL of 102 dB arises. What is the actual sound pressure at that point?

From Equation (10.3)

$$L_p = 20 \log \frac{p}{2 \times 10^{-5}} \text{ dB}$$

$$\therefore \quad \text{antilog} \frac{L_p}{20} = \frac{p}{2 \times 10^{-5}}$$

$$\therefore \quad p = \left(\text{antilog} \frac{102}{20}\right) \times 2 \times 10^{-5} = 2.52 \text{ N/m}^2 .$$

10.1.3 Sound Intensity

If it is accepted that sound pressure as the driving force is analogous to voltage in an electrical circuit, then *sound power flow* (strictly, the average rate of transfer of sound energy through unit area) is proportional to (sound pressure)^2. The power flow is called the *sound intensity* (SI). Also to complete the analogy a *characteristic impedance* (Sect.12.3.1) must be attributed to the transmission medium (usually an air-path). It can be shown that:

characteristic impedance of transmission medium

$$= \rho_0 c \qquad (10.4)$$

(ρ_0 = mean density in kg/m^3, c = velocity in m/s). From this the correspondence with $P = E^2/R$ becomes:

$$\text{Intensity}, I = \frac{p^2}{\rho_0 c} \ \text{W/m}^2 \qquad (10.5)$$

where p = r.m.s. sound pressure in N/m^2.

If we wish to talk in terms of intensity levels (IL) using logarithms then again a convenient reference level must first be chosen. If ρ_0 for air at 20°C is taken as 1.21 kg/m^3 and c as 343 m/s, then $\rho_0 c = 415\ N/m^3$ per second. So from Equation (10.5) the reference sound pressure of 2×10^{-5} N/m^2 becomes an intensity of

$$\frac{(2 \times 10^{-5})^2}{415} = 9.64 \times 10^{-13}\ \text{W/m}^2$$

In practice this is rounded up to 10^{-12} W/m^2 to become a convenient reference level on an intensity basis, so the reference levels are:

Reference sound pressure level,

$$0\ \text{dB SPL} = 2 \times 10^{-5}\ \text{N/m}^2$$

Reference sound intensity level,

$$0\ \text{dB IL} = 10^{-12}\ \text{W/m}^2$$

and these are so nearly equal in their effects that SPL and IL values can be interchanged with little loss of accuracy. Because intensity is in the nature of power, then for an intensity of I W/m^2, sound intensity level,

$$IL = 10 \log \frac{I}{10^{-12}} \text{ dB} \tag{10.6}$$

EXAMPLE 10.5:
A motor car creates a noise r.m.s. sound pressure of 0.06 N/m^2 at 5 metres away. What is the sound intensity at this point?

$$p = 0.06 \text{ N/m}^2 \qquad I = ?$$

We have two different approaches.

(i) Knowing that SPL and IL are approximately equal, from Equation (10.3):

$$SPL = 20 \log \frac{p}{2 \times 10^{-5}} = 20 \log \frac{0.06}{2 \times 10^{-5}}$$

$$= 69.5 \text{ dB}.$$

Then from Equation (10.6)

$$69.5 = 10 \log \frac{I}{10^{-12}}$$

$$\therefore \quad I = (\text{antilog } 6.95) \times 10^{-12} = 8.9 \, \mu\text{W/m}^2 .$$

(ii) Alternatively, from Equation (10.5), knowing that $\rho_0 c$ for air is equal to 415,

$$I = \frac{(0.06)^2}{415} = 8.67 \, \mu\text{W/m}^2 ,$$

showing the slight inaccuracy caused by equating the two reference levels – hardly of consequence considering the capricious nature of most audible quantities.

10.1.4 Practical Sound Levels

Some figures are given below to add a touch of realism to the various sound pressure and intensity levels in the preceding sections. Clearly they may be wide of the mark in any particular case and for speech which in fact falls near to 0 dB between words and sentences, the figure is a *long term mean*, i.e. over several seconds.

	SPL or IL, dB
Whispering at 1 metre	20
Average office	55
Conversation at 1 metre	70
Jet aircraft at 300 metres	110

10.1.5 Propagation

Consider a point source of acoustic output W watts, free to radiate in all directions in air, and positioned at the centre of an imaginary sphere of radius r metres. Because the area of the surface of a sphere is $4\pi r^2$, the power flow at the surface is W watts spread over this area, i.e. sound intensity,

$$I = \frac{W}{4\pi r^2} \text{ W/m}^2 \qquad (10.7)$$

hence the intensity varies as $1/r^2$, i.e. it follows the *inverse square law*. Therefore for intensities I_1 and I_2 of the same wave at distances r_1 and r_2 respectively

$$\frac{I_1}{I_2} = \frac{{r_2}^2}{{r_1}^2} \qquad (10.8)$$

We can also look at this in terms of the intensity level for from Equation (10.7), for any distance r

$$\text{IL} = 10 \log \frac{\frac{W}{4\pi r^2}}{10^{-12}} \text{ db} = 10 \log \left(\frac{W}{4\pi \times 10^{-12}} \times \frac{1}{r^2} \right)$$

$$= 10 \log \frac{W}{4\pi \times 10^{-12}} - 10 \log r^2$$

[from Eqn.(A2.13)]

$$\therefore \quad \text{IL} = 10 \log \frac{W}{4\pi \times 10^{-12}} - 20 \log r \text{ dB} \qquad (10.9)$$

[from Eqn.(A2.14)] .

The first term is constant for any particular value of W, the second term shows the variation in intensity level as the distance r varies, for example, if r is doubled, the intensity level is reduced by 20 log 2, i.e. 6.02 dB. The *attenuation*, α between two points r_1 and r_2 is therefore

$$\alpha = 20(\log r_2 - \log r_1) \text{ dB} \qquad (10.10)$$

These equations are only useful for estimating the intensity (or IL) at one point from a knowledge of its value at some other point and the nearer point must not be too close to the source.

EXAMPLE 10.6:

In the preceding section it is suggested that a jet aircraft 300 m high gives rise to an intensity level on the ground of 110 dB. What would the intensity and IL be at 1000 m?

Let $r_1 = 300$ m , $\quad r_2 = 1000$ m .

At 300 m, from Equation (10.6):

$$110 = 10 \log \frac{I_1}{10^{-12}}$$

$$\therefore \quad I_1 = (\text{antilog } 11) \times 10^{-12} = 0.1 \text{ W/m}^2 .$$

Let I_2 be the intensity at 1000 m, then from Equation (10.8):

$$\frac{0.1}{I_2} = \frac{1000^2}{300^2} \quad \therefore I_2 = \frac{300^2 \times 0.1}{1000^2} = 0.009 \text{ W/m}^2 .$$

For the intensity level, Equation (10.6) can be used again:

$$\therefore \quad \text{IL at 1000 m} = 10 \log \frac{0.009}{10^{-12}} = 99.5 \text{ dB},$$

or we could use Equation (10.1C) instead:

$\therefore$ attenuation between 300 and 1000 m

$$= 20(\log 1000 - \log 300)$$

$$= 20(3.0 - 2.477) = 10.46 \text{ dB}$$

$$\therefore \quad \text{IL at 1000 m} = 110 - 10.46 = 99.5 \text{ dB} .$$

EXAMPLE 10.7:
The intensity level measured at 10 m from an electric motor is 15 dB. How far away will the sound be inaudible?

This can be calculated on the basis that an additional attenuation of 15 dB will reduce the IL to 0 dB, the threshold of audibility. From Equation (10.10):

$$20(\log r_2 - \log 10) = 15$$

$$\therefore \quad \log r_2 = 0.75 + 1$$

$\therefore \quad r_2 = 56.2 \text{ m}$.

Generally we humans, especially men, suffer a hearing loss as age progresses from some 5 dB by age 40 to perhaps 10–15 dB by 60. This is equivalent to raising the threshold of audibility accordingly. Also near to the ground absorptions and reflexions appreciably modify sound wave calculations so the above examples represent little more than intelligent guesses, but at least the elementary principles are revealed.

10.2 THE MUSICAL SCALE

It is only fair to the musicians amongst us that we should find at least one formula pertaining to music. The one below gives the frequency of any note in the musical scale from a knowledge of the key to it all, note A. This note usually but not necessarily has a frequency of 440 Hz but there are also two other standard pitches in use.

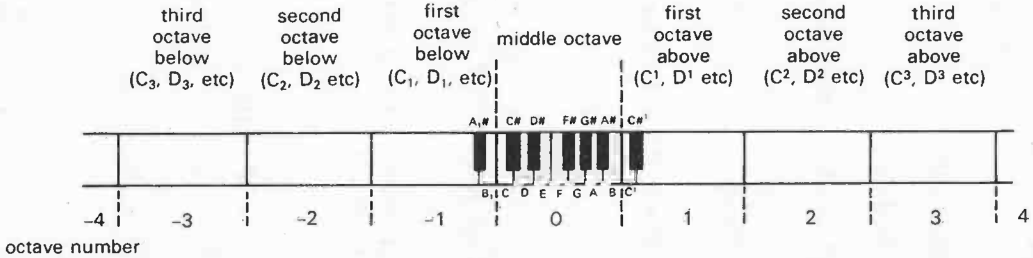
An octave rise doubles the frequency so an octave lower halves it, each of the 12 notes in an octave has a frequency $2^{1/12}$ (1.059) times that of the note below it and $2^{-1/12}$ (0.944) times that of the note above it. Hence labelling the octave and note numbers as in Figure 10.1 which uses the piano keyboard for demonstration, gives the formula:

frequency of any note

$$= f_A \times 2^x \qquad (10.11)$$

where f_A is the stated frequency of A and x = (octave number + note number/12).

EXAMPLE 10.8:
What are the frequencies of C^3 (i.e. note C in the 3rd octave above the middle), $F_2\#$ (i.e. F# or G^b in the second octave below the middle) and middle C given that the frequency of A is 440 Hz?



note	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
note number	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2

Fig. 10.1 Musical scale

The octave number for C^3 is 3, the note number is −9.

$$\therefore \quad x = 3 - \frac{9}{12} = 2.25$$

From Equation (10.11), frequency of

$$C^3 = 440 \times 2^{2.25} = 2093 \text{ Hz}.$$

The octave number for $F_2\#$ is −2, the note number is −3.

$$\therefore \quad x = \left(-2 - \frac{3}{12}\right) = -2.25$$

$$\therefore \quad \text{frequency of } F_2\# = 440 \times 2^{-2.25} = 92.5 \text{ Hz}.$$

The octave number for middle C is 0, the note number is −9.

$$\therefore \quad x = -0.75$$

and frequency of middle C $= 440 \times 2^{-0.75} = 261.6$ Hz .

EXAMPLE 10.9:
If the International Pitch A = 435, what is the range of frequencies on a piano from A_4 to C^4?

for A_4, $x = -4$

$$\therefore \quad \text{frequency} = 435 \times 2^{-4} = 27{,}19 \text{ Hz}$$

for C^4, $x = 3.25$

$$\therefore \quad \text{frequency} = 435 \times 2^{3.25} = 4138.4 \text{ Hz}.$$

10.3 SUMMARY OF KEY FORMULAE

B_a = bulk modulus of elasticity (N/m^2)

f_A = frequency of note A

P_0 = mean pressure (N/m^2)

p = r.m.s. pressure (N/m^2)

r = radius or distance

W = acoustic power

x = (octave number + 1/12th of note number – see Fig.10.1)

ρ_0 = mean density (kg/m^3)

γ = correction factor for adiabatic conditions.

QUANTITY	FORMULA	UNIT	SECTION
Velocity of sound wave in a gas	$c = \sqrt{\frac{\gamma P_0}{\rho_0}}$	m/s	10.1.1
Velocity of sound wave in a liquid	$c = \sqrt{\frac{B_a}{\rho_0}}$	m/s	10.1.1
Sound pressure level	$L_p = 20 \log \frac{p}{2 \times 10^{-5}}$	dB	10.1.2
Sound intensity	$I = \frac{p^2}{\rho_0 c}$	W/m^2	10.1.3
Sound intensity	$I = \frac{W}{4\pi r^2}$	W/m^2	10.1.5
Sound intensity level	$IL = 10 \log \frac{I}{10^{-12}}$	dB	10.1.3

QUANTITY	FORMULA	UNIT	SECTION
Sound intensity level	$\mathrm{IL} = 10 \log \dfrac{W}{4\pi \times 10^{-12}} - 20 \log r$	dB	10.1.5
Sound attenuation between r_1 and r_2	$\alpha = 20(\log r_2 - \log r_1)$	dB	10.1.5
Frequency of musical note	$f = f_A \times 2^x$	Hz	10.2

11. RADIO SYSTEMS

Whereas with physical circuits signal currents are constrained within wires or a tube, with radio the link is "wireless" and the signal is either *broadcast* or transmitted *point-to-point.* Never is the signal completely confined so privacy cannot always be ensured.

The existence of radio waves was first demonstrated by Hertz(A3) and subsequently developed commercially by Marconi(A3) around the turn of the century.

11.1 THE ELECTROMAGNETIC WAVE

The basis on which radio communication rests is the *electromagnetic wave.* Such waves are generated by moving charges. Consider an elementary aerial system ("antenna" in some countries) as in Figure 11.1 and that at a given instant the generator terminals have the polarity shown. The electron flow for this polarity is indicated and as far as the vertical wires are concerned, it is equivalent to an electron current flowing from top to bottom, reversing when the generator polarity reverses. The combined charges of the electrons creates an electric field while the fact they are moving in one direction and therefore constitute a current creates a magnetic field. A few "lines" of each are sketched in the Figure and the two fields at a reasonable distance away are perpendicular to each other. An important feature of the electromagnetic wave is that by continual transfer of energy the varying electrical field supports the magnetic field and vice versa, the two fields of the wave are inseparable.

A finite time is required for the fields around the wires to collapse hence if the generator polarity reverses before this is completed, all of the field energies cannot return to the wires because new and opposite fields are being created. The original fields are therefore forced outwards, that is, *propagated.* That part of the field which does return to the supply is known as the *induction* field and takes no part in radiation.

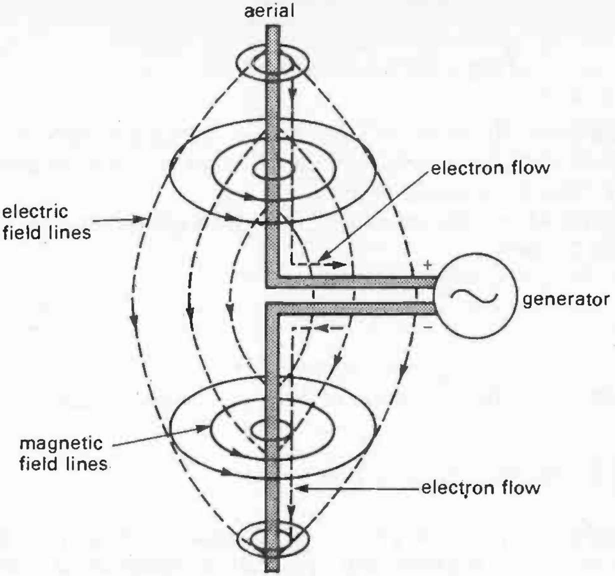


Fig. 11.1 Radiating system

11.1.1 Speed of Propagation

It can be shown that the speed of propagation (ν) of electromagnetic waves in a medium (i.e. cable, waveguide, fibre or free-space) is related to the electromagnetic constants (μ and ξ) of the medium by:

$$\nu = \sqrt{\frac{1}{\mu\xi}} \tag{11.1}$$

where $\mu = \mu_0\mu_r$ (Sect.3.2.3) and $\xi = \xi_0\xi_r$ (Sect.2.3). For free-space μ_r and $\xi_r = 1$, hence $\mu = \mu_0$ and $\xi = \xi_0$, therefore

$$\nu = \sqrt{\frac{1}{4\pi \times 10^{-7} \times 8.854 \times 10^{-12}}} = 2.998 \times 10^{8} \text{ m/s}$$

which is usually taken as 3×10^8 m/s and because light is an electromagnetic wave, its own speed in free space is the same. Light velocity is usually designed by c.

Knowing ν, the *wavelength*, λ can be calculated from the simple formula:

$$\lambda = \frac{\nu}{f} = \frac{3 \times 10^8}{f} \text{ metres} \tag{11.2}$$

where f is the frequency of the wave under consideration.

11.1.2 The Impedance of Free Space

Impedance can be expressed by a phasor which may be resolved into in-phase and quadrature components (Sect 4.2.1). Free-space has no quadrature component (i.e. it is non-reactive) so it causes no phase-change. It can be shown that:

$$\text{impedance of free-space, } Z_0 = R_0 = \sqrt{\frac{\mu_0}{\xi_0}} \tag{11.3}$$

$$\therefore \quad Z_0 = \sqrt{\frac{4\pi \times 10^{-7}}{8.854 \times 10^{-12}}} = 376.7\ \Omega$$

[the calculation is simplified by using $(10^{-9})/(36\pi)$ for ξ_0 (Sect.2.3), resulting in $Z_0 = 120\pi\Omega$ (= 377 Ω)].

Z_0 is also known as the *intrinsic* (belonging naturally) impedance of free-space. The impedance is also defined as the ratio of the electric field strength to the magnetic field strength (Sects.2.1 and 3.2.2), i.e.

$$Z_0 = \frac{E}{H}\ \Omega \tag{11.4}$$

where E is in V/m and H is in A/m. This appears reasonable by comparing with Ohm's Law and remembering that H is associated with current flow.

11.1.3 Propagation in Free Space

Mathematically propagation through any medium is expressed by a propagation constant γ which is complex and therefore subdivided into two components $\alpha + j\beta$. What matters in practice is the strength of a signal received at a distance as would be indicated by the attenuation constant α. This factor is expressed in decibels per metre and for any medium the basic equation is:

$$\alpha \approx \frac{\sigma}{2}\sqrt{\frac{\mu}{\xi}} \tag{11.5}$$

where σ is a factor known as the *electric conductivity* of the medium. For free space $\sigma = 0$, hence $\alpha = 0$, so indicating that an electromagnetic wave is not attenuated in its passage through space, for which we must truly be grateful. Note however that this is for *free* space, it is hardly free when rain, mist or certain gases are around but these have little effect until the radiated frequency exceeds some 10 GHz.

11.1.4 The Plane Wave

The simplest electromagnetic wave is the *plane wave* and in it the electric and magnetic fields are in phase and they and the direction of propagation are mutually perpendicular. Hence at some distance away the imaginary flux lines of an oncoming wave would appear to an observer as a vertical sheet of linear graph paper with the electric lines vertical and the magnetic, horizontal. This is a *vertically polarized* wave but if the lines of electric flux are horizontal with the magnetic vertical, then the wave is said to be *horizontally polarized.*

Because the attenuation of free space can generally be taken as zero, then if an *isotropic* (from Greek – equal in all

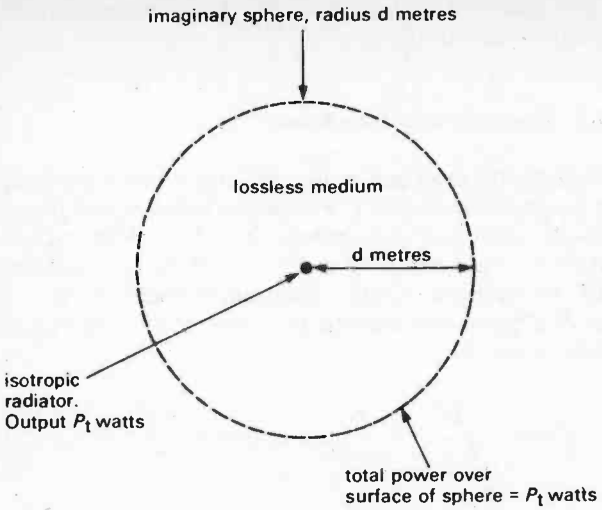


Fig. 11.2 Power flow from an isotropic radiator

directions) radiator is considered to be transmitting P_t watts as shown in Figure 11.2, since the power at any distance, d metres is equally spread over the surface of an imaginary sphere of radius d as shown, then the power flux received per unit area:

$$P_r = \frac{P_t}{4\pi d^2} \text{ watts per sq. m (W/m}^2\text{)} \qquad (11.6)$$

($4\pi d^2$ is the area of the surface of a sphere of radius d). Also calculation of the energy flow from first principles leads to

$$P_r = E \times H \text{ W/m}^2 \qquad (11.7)$$

hence since from Equation (11.4),

$$Z_0 = \frac{E}{H} = 120\pi\ \Omega$$

then

$$H = \frac{E}{Z_0}$$

and

$$P_r = \frac{E^2}{Z_0} \text{ watts}$$

$$\therefore \quad E^2 = P_r \times Z_0 = \frac{P_t}{4\pi d^2} \times 120\pi = \frac{30P_t}{d^2}$$

$$\therefore \quad E = \frac{\sqrt{30P_t}}{d} \text{ volts per metre} \qquad (11.8)$$

thus given the distance, d metres from an isotropic transmitting aerial of radiated power P_t watts, the field strength, E of an electromagnetic wave can be calculated.

EXAMPLE 11.1:
An isotropic aerial radiates a signal at a power of 83.3 watts. What are the electric and magnetic field strengths at a distance of 10 km?

$$P_t = 83.3 \text{ W} \qquad d = 10^4 \text{ m}$$

From Equation (11.8):

$$E = \frac{\sqrt{30 \times 83.3}}{10^4} \text{ V/m} = 5 \text{ mV/m}$$

and

$$H = \frac{E}{120\pi} = \frac{0.005}{120\pi} = 1.33 \times 10^{-5} \text{ A/m} .$$

11.2 THE RADIO FREQUENCY SPECTRUM

The fact that radio transmissions are designated by both frequency and wavelength and that there is a preference for frequency at the lower end of the range but for wavelength at the higher end means that we frequently need to convert from one to the other. A reminder is that Mega = 10^6 and Giga = 10^9.

11.2.1 Frequency-Wavelength Conversion

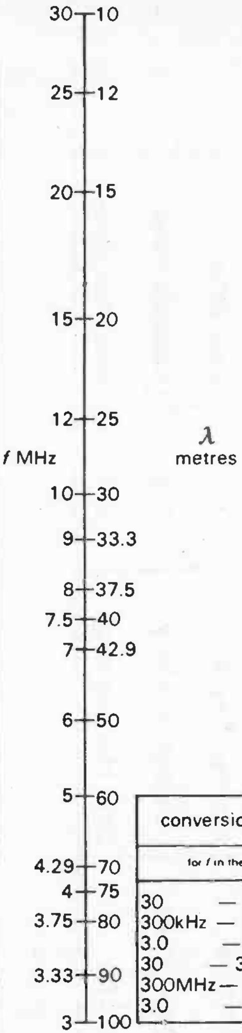
Section 11.1 developes the general formula relating frequency and wavelength [Eqn.(11.2)]. From this a simple chart as in Figure 11.3 enables quick conversions to be made. The following formulae are also relevant:

$$\text{wavelength in metres} = \frac{300}{\text{frequency in MHz}} \qquad (11.9)$$

$$\text{wavelength in centimetres} = \frac{30}{\text{frequency in GHz}} \qquad (11.10)$$

11.2.2 The Frequency Bands

For convenience the whole frequency spectrum is divided into bands as shown for most of the radio range in Table 11.1. The band number is the exponent used in converting the range 0.3 to 3 Hz to that required, i.e. band number n covers the range 0.3×10^n to 3×10^n Hz. These band numbers should not be confused with those used specifically for television.



conversion *f* into λ		conversion λ into *f*	
for *f* in the range	multiply λ by	for λ in the range	multiply *f* by
30 — 300kHz	100	1 — 10cm	1000
300kHz — 3.0MHz	10	10cm — 1.0m	100
3.0 — 30MHz	1.0	1.0 — 10m	10
30 — 300MHz	0.1	10 — 100m	1.0
300MHz — 3.0GHz	0.01	100 — 1000m	0.1
3.0 — 30GHz	0.001	1000 — 10,000m	0.01

Fig. 11.3 Frequency-wavelength conversion

Table 11.1: THE FREQUENCY SPECTRUM

Frequency Range	Description	Letter Designation	Metric Title	Band Number
30 – 300 kHz	Low frequency	LF	Kilometric waves	5
300 kHz – 3 MHz	Medium frequency	MF	Hectometric waves	6
3 – 30 MHz	High frequency	HF	Decametric waves	7
30 – 300 MHz	Very high frequency	VHF	Metric waves	8
300 MHz – 3 GHz	Ultra high frequency	UHF	Decimetric waves	9
3 – 30 GHz	Super high frequency	SHF	Centimetric waves	10

11.3 AERIALS

Section 11.1 introduces the concept of the isotropic aerial. It is a theoretical idea for such an aerial is not practicable, none can truly radiate equally in all directions. A practical basic aerial in common use is the *dipole* and calculations on this aerial help greatly in the understanding of more complex aerial systems. Dipoles are usually about half a wavelength long at the radiated frequency in which case they are described as *half-wave* and signified by $\lambda/2$. The form is as in Figure 11.1, that is, a vertical rod fed at the centre.

Most of what follows in this Section concerns transmitting aerials but generally the conclusions apply equally to receiving aerial in a reciprocal manner, the main difference between the two types is in the power-handling capacity, from watts to kilowatts for transmitting down to microwatts for receiving.

11.3.1 The Transmitting Dipole

The resonant circuit is discussed in Book 1 (Sect.4.7) and the dipole can be so considered because the two wires or rods have self-inductance with capacitance between them, it is therefore known as a *resonant* aerial. If a generator is connected to a dipole at its resonant frequency maximum power is radiated and from transmission line theory (Chapter 12) the aerial can be shown to be resonant when its length is equal to approximately half the wavelength of the radiated frequency. Although described as half-wave, the length is not exactly the expected $\lambda/2$ but slightly shorter because the length of the wave on the aerial is less than it is in free space. Returning to Equation (11.1), because μ_r and ξ_r exceed unity for the aerial rods, ν , the velocity of propagation is lower, hence from Equation (11.2), λ is less.

11.3.1.1 Directional Characteristics

All practical aerials radiate better in some directions than in others. To show the directivity performance of an aerial a *polar diagram* is constructed, as in Figure 11.4 where the distance from the point representing the aerial to the curve in any direction represents the power radiated in that direc-

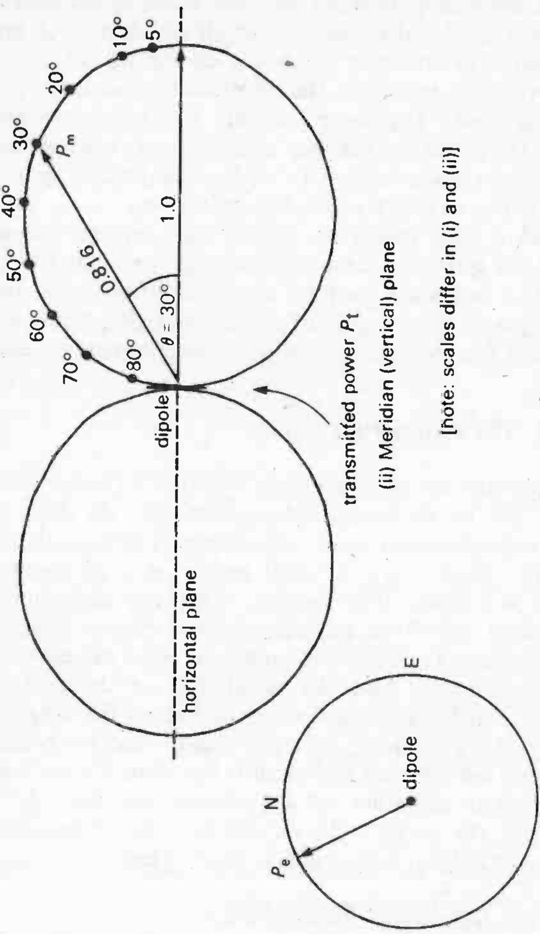


Fig. 11.4 Directivity patterns for vertical dipole

tion. Generally the directivity is shown for both the horizontal or *equatorial* and the vertical or *meridian* planes. The Figure shows these characteristics for a vertical half-wave dipole. In the horizontal plane the radiation towards all points of the compass is the same and its magnitude is represented by the vector P_e. Directivity in the vertical plane varies with the angle with the horizontal and the variation is accounted for by the fact that the aerial has finite length and both its halves contribute to the radiation together. Because the path lengths from them differ, their contributions are not in phase except at the horizontal where therefore the radiated power, P_m is maximum.

The derivation of the many equations analyzing aerial directivity is complicated but the single one for the vertical dipole gives some insight into what is involved. If $\theta°$ is the angle with respect to the plane perpendicular to the wire (the angle relative to the horizontal in the case of the vertical dipole) and ℓ the length of the dipole:

$$P \text{ (or } E \text{ or } H) \propto \frac{\cos(\pi\, \ell/\lambda \times \sin\theta)}{\cos\theta}$$

and because in this case

$$\frac{\ell}{\lambda} \approx \frac{1}{2}$$

$$P \propto \frac{\cos(\pi/2 \times \sin\theta)}{\cos\theta} \qquad (11.11)$$

Alternatively θ may be stated with respect to the wire axis whereupon the formula changes to:

$$P \propto \frac{\cos(\pi\, \ell/\lambda \times \cos\theta)}{\sin\theta} \qquad (11.12)$$

Note that π is in radians so because $\pi/2$ radians = 90 degrees, then if θ is in degrees:

$$P \propto \frac{\cos(90 \times \sin\theta)}{\cos\theta} \qquad (11.13)$$

(conversion factors between radians and degrees are given in Appendix 1).

EXAMPLE 11.2:
Construct a polar diagram for a vertical half-wave dipole in the vertical plane.

Using Equation (11.13):

$\theta°$	$\sin\theta$	$\cos\theta$	$(90\sin\theta)$	$\cos(90\sin\theta)$	$\frac{\cos(90\sin\theta)}{\cos\theta}$
0	0	1.0	0	1.0	1.000
5	0.087	0.996	7.844	0.991	0.994
10	0.174	0.985	15.630	0.963	0.978
20	0.342	0.940	30.780	0.859	0.914
30	0.500	0.866	45.0	0.707	0.816
↓					
90	1.0	0	90.0	0	0

[With a scientific calculator this table is unnecessary and especially so if the calculator is programmable.]

The results are in fact plotted in Figure 11.4(ii).

From this example it is evident that a vertical dipole radiates maximum power in a horizontal direction, tailing off to zero vertically and correspondingly for the horizontal dipole which is simply a vertical one rotated through 90°.

In the direction of maximum radiation it has been established that Equation (11.6) is modified to:

$$P_r = 1.64 \frac{P_t}{4\pi d^2} \text{ W/m}^2 \qquad (11.14)$$

and since $E = \sqrt{P_r Z_0}$, from Equation (11.3):

$$E_r = \frac{\sqrt{49.2 P_t}}{d} \text{ V/m} \qquad (11.15)$$

showing a gain of 1.64 (2.15 dB) over the isotropic aerial.

Where greater directivity is required two or more dipoles may be used spaced apart so that their signals at a distant point add or subtract to give the directivity required. It is instructive to examine this technique in a simple form because the effect of phase differences becomes apparent so leading to a better understanding of more complex aerial systems.

Suppose two vertical dipoles D1 and D2 are spaced a distance d apart and that the power radiated by each is the same, then working in, say, electric field strength:

$$E \propto \cos \frac{180d \times \cos\theta}{\lambda} \text{ V/m} \qquad (11.16)$$

and if there is a phase difference ϕ° between the currents in the two dipoles:

$$E \propto \cos \left\{ \frac{180d \times \cos\theta}{\lambda} \pm \frac{\phi}{2} \right\} \text{ V/m} \qquad (11.17)$$

(the ± sign allows for D1 leading or lagging D2 – we need not be too concerned which to use because it becomes evident from the polar diagram).

EXAMPLE 11.3:
Two vertical dipoles spaced one quarter of a wavelength apart are fed 90° out of phase. Plot the polar diagram.

From Equation (11.17), if $d = \lambda/4$, then

$$E \propto \cos\{(45 \cos\theta) + 45\}$$

Calculations for a few points are shown:

$\theta°$	$\cos\theta$	$45\cos\theta$	$(45\cos\theta) + 45$	$\cos\{(45\cos\theta) + 45\}$
80	0.174	7.814	52.81	0.604
140	–0.766	–34.47	10.53	0.983
210	–0.866	–38.97	6.03	0.994

The pattern is drawn in Figure 11.5 with the above points shown by vectors. In the direction D2 → D1 radiation is cancelled because energy travelling from D2 over a distance $\lambda/4$ to D1 arrives 90° later. D2 already lags on D1 by 90° hence radiation from D2 at D1 is 180° out of phase and cancels. In the opposite direction radiation from D1 arrives at D2 in phase.

Although this is a simple example embracing two dipoles only, it does show what can be done in estimating aerial performance. Directivity is further discussed in Section 11.3.2.3.

11.3.1.2 Radiation Resistance

Current flows into an aerial resulting in the radiation of power. It is often expedient to relate these two quantities by a fictitious resistance value known as the *radiation resistance*. This is defined as that value of resistance which when multiplied by the mean square value of the current (I^2_{rms}) flowing at the point of measurement, gives the power radiated. Hence

$$\text{Power radiated, } P_r = I^2 R_r \qquad (11.18)$$

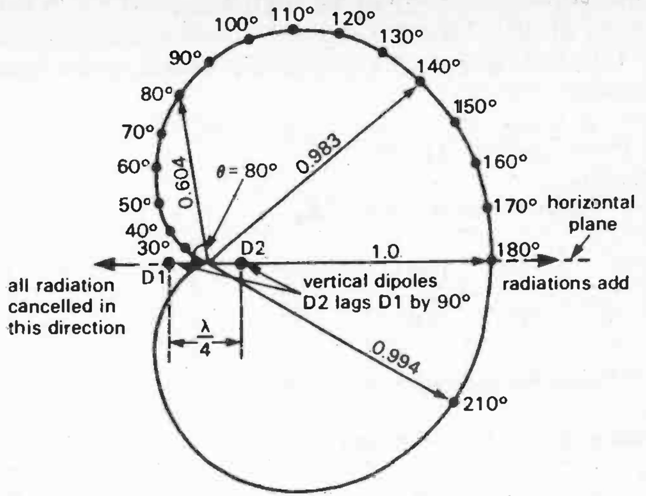


Fig. 11.5 Directivity pattern for 2 dipoles spaced $\frac{\lambda}{4}$ with 90° phase difference

where I is the aerial current and R_r the radiation resistance.

An aerial also has power losses (resistance, dielectric loss etc.) which are wasteful so aerial efficiency η is expressed by the ratio of power radiated to total power absorbed:

$$\text{aerial efficiency, } \eta = \frac{I^2 R_r}{I^2 R_r + I^2 R_L} \times 100\%$$

$$= \frac{100 R_r}{R_r + R_L} \qquad (11.19)$$

where R_L is a total loss resistance.

EXAMPLE 11.4:
A transmitting aerial has a total loss resistance of 1.2 Ω. If a current of 40 A fed to the aerial results in a radiated power of 1.28 kW, what are the radiation resistance, power input and aerial efficiency?

From Equation (11.18):

$$\text{Radiated power, } P_r = I^2 R_r$$

$$\therefore \quad R_r = \frac{P_r}{I^2} = \frac{1280}{40^2}$$

$$\therefore \quad \text{radiation resistance} = 0.8\ \Omega\,.$$

Total power input from supply,

$$P_i = I^2(R_r + R_L) = 40^2(0.8 + 1.2)\ \text{W} = 3.2\ \text{kW}\,.$$

Aerial efficiency,

$$\eta = \frac{P_r}{P_i} = \frac{1.28}{3.2} \times 100\% = 40\%$$

or from Equation (11.19),

$$\eta = \frac{100R_r}{R_r + R_L} = \frac{100 \times 0.8}{0.8 + 1.2} = 40\%\,.$$

11.3.1.3 Aerial Gain

Gain of an aerial system is defined as the ratio of the field strength at a distant point generated by the aerial under test to that which would be produced by a reference aerial (e.g. isotropic or half-wave dipole). The distant point is frequently in the direction of maximum radiation, hence gain and directivity of an aerial are closely related. Gain is usually quoted in decibels.

EXAMPLE 11.5:
An aerial system under test is fed with a power of 500 W and the field strength at a distant point is measured. The aerial is substituted by a $\lambda/2$ dipole which is found to require 3.15 kW to produce the same field strength at the same place. What is the test aerial gain in the particular direction? What is the gain relative to an isotropic radiator?

$$\text{Aerial gain relative to } \frac{\lambda}{2} \text{ dipole} = 10 \log \frac{3.15}{0.5} = 8 \text{ dB}.$$

The $\lambda/2$ dipole itself has a gain of 2.15 dB relative to an isotropic radiator (see Sect.11.3.1.1). Therefore aerial gain relative to isotropic radiator is equal to

$$8.0 + 2.15 = 10.15 \text{ dB}.$$

Equation (11.8) can now be modified for a practical transmitting aerial with gain G_t relative to the isotropic:

$$\text{field strength, } E = \frac{\sqrt{30 G_t \times P_t}}{d} \text{ V/m} \qquad (11.20)$$

11.3.2 The Receiving Dipole

Together the alternating fields of an electromagnetic wave are capable of forcing electrons in a wire into a state of oscillation at the wave frequency. The wire is then acting as a receiving aerial and for a vertical aerial in the path of a vertically polarized wave, the e.m.f. induced is maximum but in a horizontally polarized wave it is theoretically zero. Both aerials and wave-fronts may have angles to the horizontal other than 0° or 90° (say, $\theta°$), in calculations this is usually accounted for by multiplying the result by $\sin\theta$ or $\cos\theta$ as appropriate. The formula for e.m.f. generated in a simple aerial embodies a parameter known as the *effective length*, this is therefore considered next.

11.3.2.1 Effective Length

This is defined as the length of a hypothetical aerial in which the current is *uniform* throughout the length and equal to the maximum value in the actual aerial. A centre-connected aerial such as the dipole has maximum current at the centre because this is its supply point and zero at the two ends where the metallic circuit of the aerial ceases as shown in Figure 11.6. If the aerial is excited sinusoidally the current distribution as represented by the dotted curve in the Figure is also sinusoidal.

The average value of a sine wave $= 2/\pi \times I_{max}$ so the area under the current curve in the Figure is represented by the rectangle on base $\lambda/2$ and height $2/\pi \times I_{max}$ as shown. The second rectangle of equivalent area and height I_{max} has therefore a base shorter than $\lambda/2$ and this is the effective length of the aerial, ℓ_{eff}. Then for the $\lambda/2$ dipole:

$$\frac{2}{\pi} \times I_{max} \times \frac{\lambda}{2} = I_{max} \times \ell_{eff}$$

$$\therefore \quad \ell_{eff} = \frac{2}{\pi} \times \frac{\lambda}{2} = \frac{\lambda}{\pi} \tag{11.21}$$

The effective length of a receiving aerial can also be defined as:

$$\ell_{eff} = \frac{\text{open circuit voltage delivered by aerial } (e_a)}{\text{field strength } (E_r)}$$

and this indicates a practical way of measurement. Then

$$e_a = E_r \times \ell_{eff} \tag{11.22}$$

The definition of ℓ_{eff} enables us to prove Equation (11.22) from first principles by using Equation (3.16) for the e.m.f. induced in a conductor when magnetic flux is cut:

$$e = B \times \ell_{eff} \times v$$

where B is the wave magnetic flux density and v is the wave velocity.

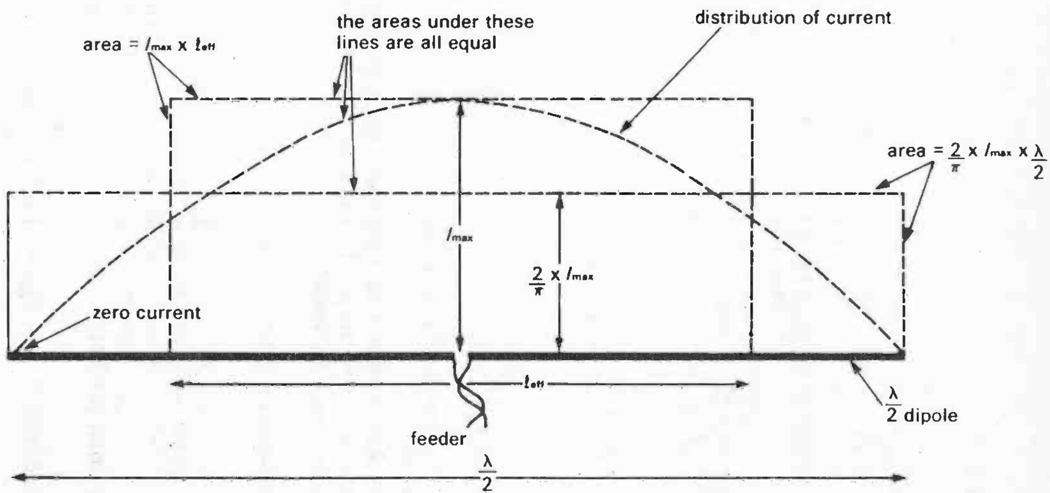


Fig. 11.6 Effective length of half-wave dipole

From Equation (3.4)

$$e = \mu_0 H \times \ell_{eff} \times \nu$$

and from Equation (11.1)

$$\nu = \frac{1}{\sqrt{\mu_0 \xi_0}}$$

and from Equations (11.3) and (11.4)

$$H = E_r \sqrt{\frac{\xi_0}{\mu_0}}$$

$$\therefore \quad e = \mu_0 \; E_r \frac{\sqrt{\xi_0}}{\sqrt{\mu_0}} \times \ell_{eff} \times \frac{1}{\sqrt{\mu_0} \times \sqrt{\xi_0}}$$

$$= E_r \times \ell_{eff} \,.$$

EXAMPLE 11.6:
A horizontal $\lambda/2$ dipole 1 m long picks up a horizontally polarized signal of field strength 15.7 mV/m. What is the p.d. generated across a matched load?

From Equation (11.21):

$$\ell_{eff} \text{ of dipole} = \frac{2}{\pi} \times 1 \text{ m} = \frac{2}{\pi} \text{ m}$$

and from Equation (11.22):

$$\text{e.m.f. induced} = E_r \times \ell_{eff} = 15.7 \times \frac{2}{\pi} \text{ mV}$$

$$= 10 \text{ mV} \,.$$

$\therefore$ p.d. across matched load $= \dfrac{10}{2}$ mV $= 5$ mV

[the p.d. across a matched load is equal to half the generator e.m.f. – see Fig.11.7(i)].

11.3.2.2 Aerial Signal Strength

For any receiving aerial system a figure known as the *effective receiving area* is first calculated. Not unexpectedly this involves the gain, G_r of the aerial (Sect.11.3.1.3) relative to an isotropic radiator which itself has an effective area of $\lambda^2/4\pi$ and the formula is:

$$\text{receiving effective area, } A_r = \frac{G_r\lambda^2}{4\pi} \qquad (11.23)$$

The wave power density at a distance d from the transmitter becomes [from Equation (11.6)]:

$$P_r = \frac{P_t}{4\pi d^2} \times G_t \qquad (11.24)$$

where G_t is the transmitting aerial gain relative to the isotropic. Multiplying by the receiving aerial effective area:

$$\text{Power received by aerial, } P_a = \frac{P_t \times G_t}{4\pi d^2} \times \frac{G_r\lambda^2}{4\pi}$$

$$= \frac{P_tG_tG_r\lambda^2}{(4\pi d)^2} \text{ watts} \qquad (11.25)$$

which can be converted to voltage from a knowledge of the aerial impedance. Thus if $\lambda/2$ dipoles are used for both transmitter and receiver:

$G_t = 1.64$, $\quad G_r = 1.64$

and

$$P_a = \frac{1.64^2 \, P_t\lambda^2}{(4\pi d)^2} = \frac{0.017 \, P_t\lambda^2}{d^2} \qquad (11.26)$$

EXAMPLE 11.7:
A radio link has a 1000 W transmitter with an aerial gain of 5 at a frequency of 300 MHz. The transmission is received on a half-wave dipole aerial 12 km away. What is the receiving aerial p.d.? (The aerial impedance can be taken as 73 Ω.)

We can answer this question (i) by using Equation (11.25) which gives the received aerial power directly or (ii) by first calculating the field strength at the receiver and then from this the voltage developed in the aerial. Both methods are shown below to give a measure of confidence in the formulae so far developed:

$$\lambda = \frac{v}{f} = \frac{3 \times 10^8}{300 \times 10^6} = 1 \text{ m},$$

$d = 12000$ m, $P_t = 1000$ W, $G_t = 5$, $G_r = 1.64$ and $Z_a = 73$ Ω.

(i) From Equation (11.25)

$$\text{Receiving aerial power, } P_a = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2} \text{ watts}$$

$$= \frac{1000 \times 5 \times 1.64 \times 1}{(4\pi \times 12000)^2} = 0.361 \, \mu\text{W}$$

$$\therefore \quad \text{Aerial p.d., } E_a = \sqrt{P_a Z_a} = \sqrt{0.361 \times 10^{-6} \times 73}$$

$$= 5.13 \text{ mV}$$

(ii) From Equation (11.24)

$$\text{Wave power density at receiver, } P_r = \frac{P_t G_t}{4\pi d^2}$$

$$= \frac{1000 \times 5}{4\pi \times 12000^2} = 2.76\ \mu\text{W/m}^2\ ,$$

then because field strength at receiver $E_r = \sqrt{P_r \times Z_0}$ (where Z_0 is the impedance of free space)

$$\therefore \quad E_r = \sqrt{2.76 \times 10^{-6} \times 120\pi} = 0.0323\ \text{V/m}\ .$$

Then for a dipole aerial, from Equations (11.21) and (11.22):

$$\text{induced e.m.f., } e_a = E_r \times \ell_{eff}$$

$$= 0.0323 \times 0.5 \times \frac{2}{\pi}\ \text{V} = 10.28\ \text{mV}$$

$$\therefore \quad \text{aerial p.d., } E_a = \frac{10.28}{2} = 5.14\ \text{mV}$$

[see Fig.11.7(i)].

The various quantities are shown pictorially in Figure 11.7(ii).

11.3.2.3 The Yagi Array

A particularly interesting receiving aerial system which has wide use in television because of its good directivity was developed by two Japanese engineers, H. Yagi and S. Uda although the aerial is generally known as a *Yagi array*. This has a $\lambda/2$ dipole assisted by a *reflector* behind it and one or more *directors* in front as shown in Figure 11.8. These are simply rods which are excited by the oncoming wave and

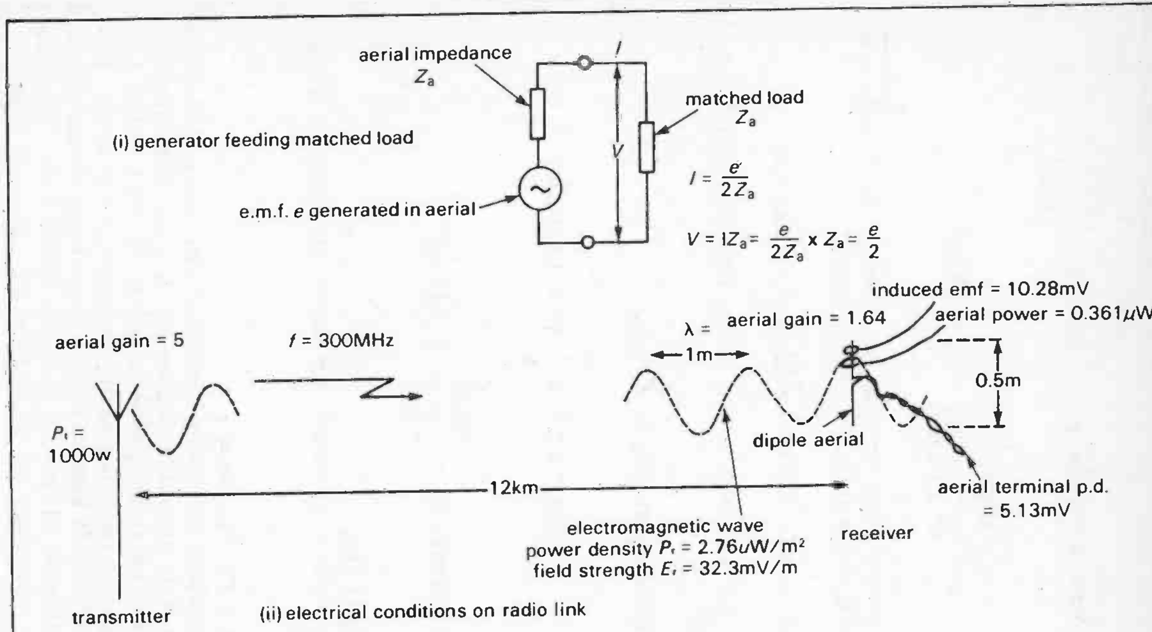


Fig. 11.7 Calculation of receiving aerial P.D.

their spacings from the dipole are such that they re-radiate energy to arrive in the desired phase. The array dimensions are a compromise because not only do they affect directivity but also aerial bandwidth and resistance. Nevertheless we can design a simple Yagi receiving aerial by using a set of ratios in common use and based on a 5% longer reflector and 5% shorter director, (one director only, not two as in the Figure) resulting in:

dipole length $= 0.5\ \lambda$

reflector length $= 0.525\ \lambda$,

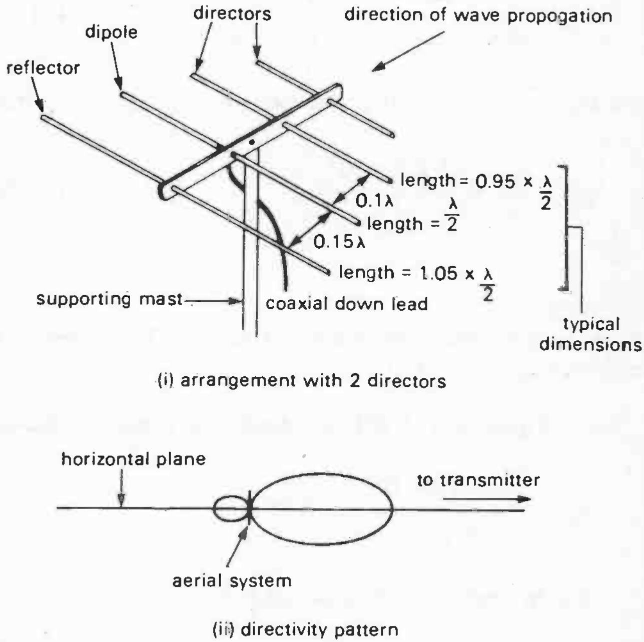


Fig. 11.8 Yagi aerial

spacing from dipole $= 0.15\ \lambda$

director length $= 0.475\ \lambda$,

spacing from dipole $= 0.1\ \lambda$

as shown in Figure 11.8(i).

A factor which must be taken into consideration arises from the fact that waves travel more slowly in the aerial conductors than in free space (see also Sects. 11.3.1 and 12.3.3). It is known as the *velocity factor* and is determined from the length/diameter ratio of the aerial conductors: velocity factor,

$$k = \frac{\text{velocity of waves in aerial}}{\text{velocity of waves in free-space } (c)} \tag{11.27}$$

then for the aerial conductors, Equation (11.2) is modified to

$$\lambda_a = \frac{kv}{f} = \frac{k \times 3 \times 10^8}{f}\ \text{m} \tag{11.28}$$

EXAMPLE 11.8:
Design a Yagi-type television aerial for 500 MHz given that the velocity factor is 0.95.

From Equation (11.28), wavelength in aerial conductors,

$$\lambda_a = \frac{0.95 \times 3 \times 10^8}{500 \times 10^6} = 0.57\ \text{m}.$$

$\therefore$ dipole length $= 0.5\ \lambda_a = 28.5$ cm

reflector length $= 0.525\ \lambda_a = 29.9$ cm

director length $= 0.475\ \lambda_a = 27.1$ cm.

The element spacings are in air hence we revert to Equation (11.2) or Equation (11.28) with $k = 1$

$$\therefore \quad \lambda = \frac{3 \times 10^8}{500 \times 10^6} = 0.6 \text{ m}$$

and reflector spacing $= 0.15\,\lambda = 0.15 \times 0.6$ m $= 9$ cm ,

also director spacing $= 0.1\,\lambda = 0.1 \times 0.6$ m $= 6$ cm .

11.4 THE SUPERHETERODYNE RECEIVER

Born out of the need to achieve high gain and selectivity at r.f. without instability the *superheterodyne* principle is now employed in the majority of radio receivers. We get the word *heterodyne* from the Greek, meaning different "forces", or in this case, "frequencies". The original label was "supersonic heterodyne", now abbreviated to superheterodyne or even to *superhet.* Super (Latin for "above") and sonic (Latin, relating to sound) combine to put the device in the r.f. spectrum rather than the audio. It is essentially an arrangement by which most of the r.f. amplification is carried out at a single fixed frequency (the *intermediate frequency* or i.f.) irrespective of the input signal frequency. The technique therefore involves frequency changing from signal to i.f. as outlined in Figure 11.9(i). The oscillator and mixer combine to form a *frequency changer* following the principles of modulation outlined in Section 6.5.1.3 in which the signal frequency (f_s) and oscillator frequency (f_0) are applied together to a non-linear device (the *mixer*). Sum and difference frequencies ($f_0 + f_s$) and ($f_0 - f_s$) are generated and one of these is filtered off as the i.f. and from the earlier section it is evident that this contains the modulation of the signal. The i.f. is then amplified and demodulated (Sect.6.5.1.4) with the resulting audio frequency amplified and fed to a loudspeaker. The i.f. amplifiers employ double-tuned coupled circuits having a pass band to suit the modulating frequency (Book 1, Sect.4.7.5).

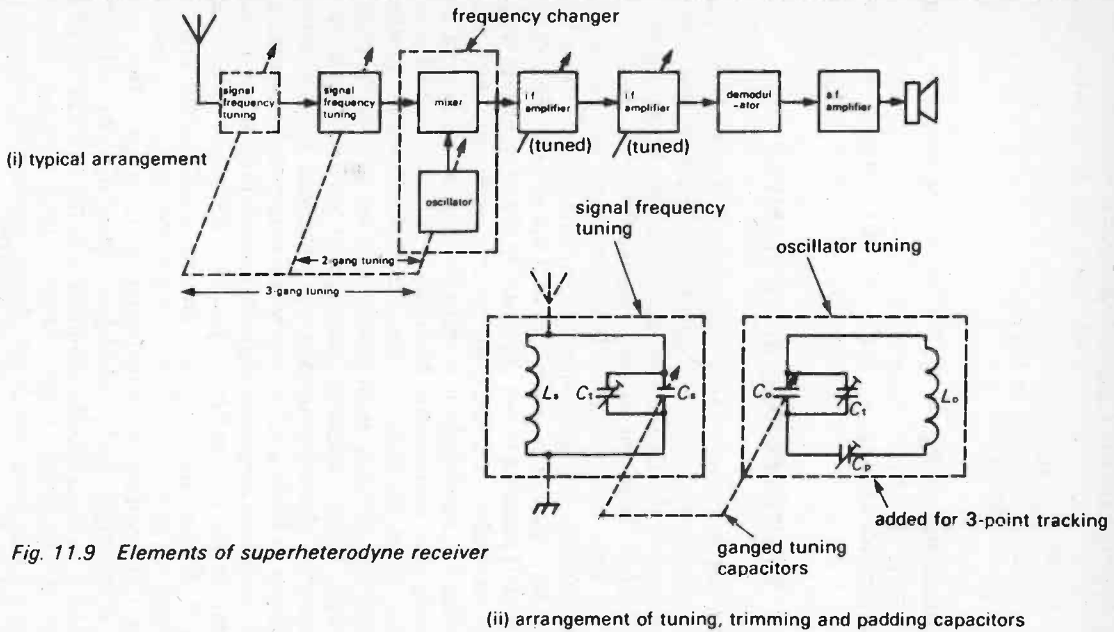


Fig. 11.9 Elements of superheterodyne receiver

The question arises as to the best values for the intermediate frequency (f_{if}) and f_0. Most modern receivers use an i.f. of 450–470 kHz and it can be shown that f_0 must be higher than f_s, hence

$$f_{if} = (f_0 - f_s) \qquad (11.29)$$

11.4.1 Image Channel Interference

One complication arising from the technique is that a second incoming signal is capable of producing the i.f. It is known as the *image frequency* (f_{im}) and when mixed with f_0 gives ($f_{im} + f_0$) and ($f_{im} - f_0$). Trouble arises when the difference frequency ($f_{im} - f_0$) is equal to the i.f. for then the image signal modulation is carried by the i.f. amplifiers, as follows:

since $f_{if} = f_{im} - f_0$

and from Equation (11.29),

$$f_0 = f_{if} + f_s,$$

then, $f_{if} = f_{im} - f_{if} - f_s$

$$\therefore \quad f_{im} = f_s + 2f_{if} \qquad (11.30)$$

showing that any incoming signal higher in frequency than the wanted signal by twice the i.f. will cause what is termed *image channel* (or *second channel*) *interference*. Rejection of the image signal must therefore be accomplished before the mixing stage.

EXAMPLE 11.9:
A superhet receiver with 465 kHz i.f. is tuned to a signal at 1.0 MHz. What are the frequencies of the oscillator and the image signal?

$$f_{if} = 465 \text{ kHz} \qquad f_s = 1.0 \text{ MHz}$$

From Equation (11.29), oscillator frequency,

$$f_o = f_{if} + f_s = 1465 \text{ kHz} .$$

From Equation (11.30), image signal frequency,

$$f_{im} = f_s + 2f_{if} = 1000 + (2 \times 465) \text{ kHz} = 1930 \text{ kHz}$$

11.4.2 Components of the Tuned Circuits

Although the most recent techniques employ varicap diodes for tuning, the standard 2- or 3-gang variable air capacitor is still very much in evidence. These are available with certain capacitance ranges only so the tuning inductors are wound to suit. Figure 11.9(ii) shows an arrangement in which C_s and C_o are the two sections of a ganged tuning capacitor controlling the signal frequency tuning circuit and the oscillator frequency at the same time. A 3-gang capacitor is used when an additional stage of r.f. amplification is provided as shown dotted in (i) of the Figure. Ignore for the present the capacitor C_p . Whatever the position of the tuning control, C_s is equal to C_o and it is essential that, whatever their values, the two circuits should have a frequency difference equal to the i.f. in accordance with Equation (11.29).

Firstly the values of L_s and L_o must be calculated. Consider the signal-frequency tuning circuit of Figure 11.9(ii) and let C_t represent not only the pre-set capacitance of a *trimmer* capacitor but in addition the total stray capacitance affecting the circuit. If the maximum and minimum values of C_s are designated $C_{s(max)}$ and $C_{s(min)}$ then the highest frequency tunable is:

$$f_h = \frac{1}{2\pi\sqrt{L_s(C_{s(min)} + C_t)}} \text{ Hz} \qquad (11.31)$$

and the lowest:

$$f_\ell = \frac{1}{2\pi\sqrt{L_s(C_{s(max)} + C_t)}} \text{ Hz} \qquad (11.32)$$

Then

$$L_s C_{s(max)} + L_s C_t = \frac{1}{4\pi^2 f_\ell^2}$$

$$L_s C_{s(min)} + L_s C_t = \frac{1}{4\pi^2 f_h^2}$$

and subtracting:

$$L_s(C_{s(max)} - C_{s(min)}) = \frac{1}{4\pi^2}\left(\frac{1}{f_\ell^2} - \frac{1}{f_h^2}\right)$$

$$= \frac{1}{4\pi^2}\left(\frac{f_h^2 - f_\ell^2}{f_h^2 \times f_\ell^2}\right)$$

$$\therefore \quad L_s = \frac{f_h^2 - f_\ell^2}{4\pi^2 \times f_h^2 \times f_\ell^2 \, (C_{s(max)} - C_{s(min)})} \qquad (11.33)$$

Also from Equation (11.31)

$$C_t = \frac{1}{4\pi^2 f_h^2 \times L_s} - C_{s(min)} \qquad (11.34)$$

or from Equation (11.32)

$$C_t = \frac{1}{4\pi^2 f_\ell^2 \times L_s} - C_{s(max)} \qquad (11.35)$$

hence, given the value of C_s, both L_s and C_t can be determined for any frequency range. L_0, C_0 and its trimmer, C_t are calculated from the same formulae.

EXAMPLE 11.10:
What values of inductor and trimmer capacitor are required

for tuning over the medium wave band (550–1600 kHz) with a 15–365 pF variable capacitor?

$$C_{s(max)} = 365 \text{ pF}$$

$$C_{s(min)} = 15 \text{ pF}$$

$$f_\ell = 550 \text{ kHz}$$

$$f_h = 1600 \text{ kHz}$$

From Equation (11.33)

$$L_s = \frac{(1600^2 - 550^2) \times 10^6 \times 10^6}{4\pi^2 \times 1600^2 \times 550^2 \times (365 - 15)} \mu\text{H} = 211 \,\mu\text{H}$$

From Equation (11.34)

$$C_t = \frac{10^{12}}{4\pi^2 \times 1600^2 \times 10^6 \times 211 \times 10^{-6}} - 15 \text{ pF}$$

$$= 49.6 - 15 = 31.9 \text{ pF},$$

or equally from Equation (11.35). A few pF, (say, 5) are taken up by stray capacitances, therefore the trimmer capacitor must be variable around 27 pF, i.e. a 33 pF would be suitable.

EXAMPLE 11.11:
If the tuned circuit in Example 11.10 is the signal frequency circuit of a superheterodyne receiver having an intermediate frequency of 460 kHz, calculate the values of the inductor and the trimmer capacitor required for the oscillator.

$$C_o = C_s = 15 - 365 \text{ pF},$$

$$f_\ell = 550 + 460 = 1010 \text{ kHz},$$

$$f_h = 1600 + 460 = 2060 \text{ kHz}.$$

From Equation (11.33):

$$L_o = \frac{(2060^2 - 1010^2) \times 10^6}{4\pi^2 \times 2060^2 \times 1010^2 \times 10^{12} \times 350 \times 10^{-12}}$$

$$\times 10^6\ \mu H = 53.9\ \mu H,$$

and from Equation (11.34):

$$C_t = \frac{10^{12}}{4\pi^2 \times (2060 \times 10^3)^2 \times 53.9 \times 10^{-6}} - 15\ pF$$

$$= 110.7 - 15 = 95.7\ pF.$$

Again allowing, say, 5 pF for stray capacitances, the trimmer must be variable around 91 pF, i.e. a variable capacitor of maximum value 100 or 120 pF would be suitable.

11.4.3 Tracking

All is not necessarily well, in the two examples above we have designed the signal frequency and oscillator circuits for the extreme ends of the frequency range only. We have yet to see how well each circuit tunes to the desired frequency at other points within the range, in superhet terms, how good is the *tracking*. Let us first look at the centre of the range, that is when C_s and C_0 both have a value of say, 180 pF. Then for the signal frequency circuit, the total capacitance = $C_s + C_t$ = 180 + 32 = 212 pF, and for the oscillator, 180 + 96 = 276 pF, and

$$f_{sig} = \frac{1}{2\pi\sqrt{L_s(C_s + C_t)}}$$

$$= \frac{1}{2\pi\sqrt{211 \times 10^{-6} \times 212 \times 10^{-12}}}\ Hz$$

$$= \frac{10^6}{2\pi\sqrt{211 \times 212}} \text{ kHz} = 753 \text{ kHz}$$

$$f_{osc} = \frac{10^6}{2\pi\sqrt{53.9 \times 276}} \text{ kHz} = 1305 \text{ kHz}$$

which will mix with a signal of 1305 – 406 = 845 kHz to produce an i.f. of 460 kHz. Thus the tuning by the signal frequency circuit is of no help. Clearly to receive a signal

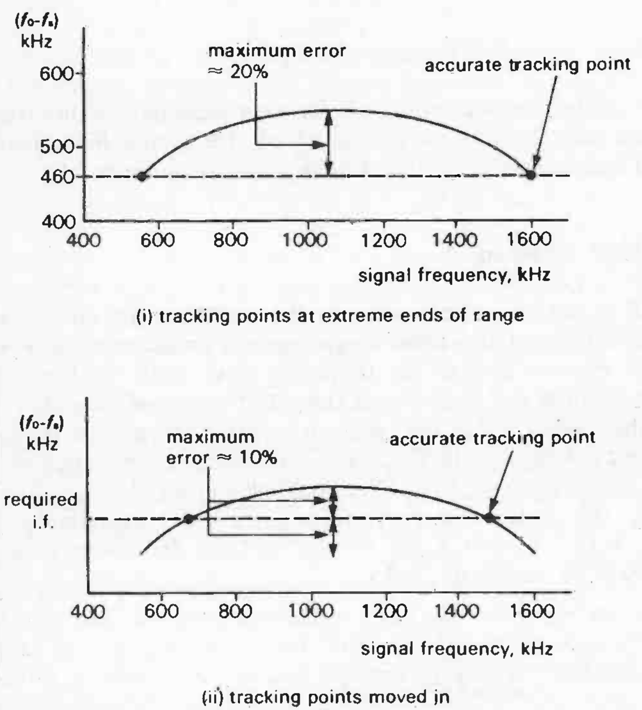


Fig. 11.10 Superhet two-point tracking

at 753 kHz the oscillator should tune to 753 + 460 = 1213 kHz instead of 1305. The difficulty arises from the fact that similar variable capacitors are being expected to cover different frequency *ratios*, for example, in the medium wave case we are examining, the signal frequency range is 1600/550 = 2.9 , whereas for the oscillator it is 2060/1010 = 2.0 . Examining the difference frequency over the whole range as in Figure 11.10(i) brings the obvious conclusion that it is wrong to design for accurate alignment at the ends of the range, there is too great an error at the centre. In fact the maximum error can be approximately halved by designing for alignment at two points within the range as indicated in Figure 11.10(ii). This is known as *two-point tracking.*

Better tracking still is obtained if the alignment is made correct at three points (*three-point tracking*) and it is achieved by reducing C_0 by means of a series *padding* capacitor C_p [see Fig.11.9(ii)]. This is the technique most commonly employed and the receiver is aligned by injection of a modulated r.f. test signal at the aerial terminal, then the two trimmers, and padder are adjusted for maximum overall sensitivity at various frequencies over the range. Because so many variables are present, mathematical solution is tedious but to give us confidence that the system can work efficiently the next example takes a typical commercial circuit and shows the calculations necessary to determine its tracking efficiency.

EXAMPLE 11.12:
Figure 11.11(i) shows the medium wave section of the frequency-changer circuit of a superhet receiver. The i.f. is 460 kHz and when the set has been aligned, measurement of the variable capacitors gives the values quoted in (ii). Plot the tracking characteristic.

The required characteristic can be obtained by assigning various values to C_s and C_0 , calculating the total capacitance in both resonant circuits, then calculating f_s , f_0 and finally $(f_0 - f_s)$, truly a job for a scientific calculator. Steps in each calculation might be as follows and note that for f_s and f_0 it may be worthwhile reducing the formula (as shown) to make repeated calculations less tedious. Take as a single example the calculations involved when C_s and C_0 have

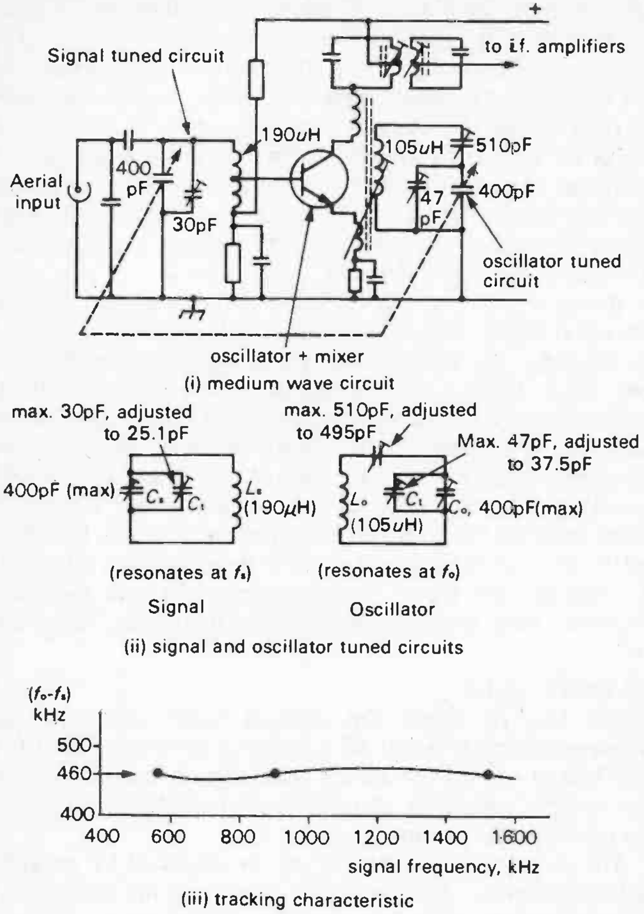


Fig. 11.11 Superhet three-point tracking

capacitances of 150 pF, then:

$$C_s \text{ and } C_0 = 150 \text{ pF} \qquad \text{(i)}$$

Total capacitance in signal circuit $=$ (i) $+$ 25.1

$$[175.1 \text{ pF}] \qquad \text{(ii)}$$

Total capacitance in oscillator circuit without C_p $=$

$$\text{(i)} + 37.5 \qquad [187.5 \text{ pF}] \qquad \text{(iii)}$$

Total capacitance in oscillator circuit with C_p $=$

$$\frac{495 \times \text{(iii)}}{495 + \text{(iii)}} \qquad [136 \text{ pF}] \qquad \text{(iv)}$$

Then

$$f_s = \frac{1}{2\pi\sqrt{190 \times 10^{-6} \times \text{(ii)} \times 10^{-12}}}$$

$$= \frac{10^9}{2\pi\sqrt{190 \times \text{(ii)}}} \text{ Hz} = \frac{159155}{\sqrt{190 \times \text{(ii)}}} \text{ kHz}$$

$$[872.6 \text{ kHz}] \qquad \text{(v)}$$

similarly

$$f_0 = \frac{159155}{\sqrt{105 \times \text{(iv)}}} \text{ kHz} \qquad [1331.9 \text{ kHz}] \qquad \text{(vi)}$$

and finally

$$(f_0 - f_s) = \text{(vi)} - \text{(v)} \qquad [459 \text{ kHz}] ,$$

the tracking curve can then be plotted on a basis of C_s/C_0 or f_s. The latter is chosen for Figure 11.11(iii) so that the order of improvement over the simple methods of tracking as in (i) and (ii) can be judged. In fact the addition of the padding capacitor means that at the worst the signal frequency circuit is less than 0.7% off tune (at 1154 kHz it is out of alignment by 8 kHz). Even better results may be possible by adjusting the i.f. amplifiers to slightly above or below the nominal frequency to suit.

11.5 SUMMARY OF KEY FORMULAE

C = capacity

d = distance (m)

E = electric field strength

f = frequency

G_r = receiving aerial gain

G_t = transmitting aerial gain

H = magnetic field strength

P = power

P_t = transmitted power

R_L = loss resistance

R_r = radiation resistance

ξ = permittivity

ξ_0 = permittivity of free space

μ = permeability

μ_0 = permeability of free space

θ = angle (degrees)

σ = electric conductivity

QUANTITY	FORMULA	UNIT	SECTION
Speed of propagation of electro-magnetic waves	$v = \sqrt{\frac{1}{\mu\xi}}$	m/s	11.1.1
Wavelength	$\lambda = \frac{v}{f}$	m	11.1.1
Wavelength in metres	$\lambda = \frac{300}{f \text{ in MHz}}$	m	11.2
Wavelength in centimetres	$\lambda = \frac{30}{f \text{ in GHz}}$	cm	11.2
Impedance of free space	$Z_0 = \sqrt{\frac{\mu_0}{\xi_0}} = \frac{E}{H} \approx 120\pi$	Ω	11.1.2
Wave power at distance from isotropic radiator	$P_r = \frac{P_t}{4\pi d^2} = E \times H$	W/m²	11.1.4

QUANTITY	FORMULA	UNIT	SECTION
Field strength at distance from isotropic radiator	$E = \frac{\sqrt{30P_t}}{d}$	V/m	11.1.4
Field strength, aerial gain G_t	$E = \frac{\sqrt{30G_tP_t}}{d}$	V/m	11.3.1.3
Aerial directivity	P or E or $H \propto \frac{\cos(90 \times \sin\theta)}{\cos\theta}$	W/m V/m A/m	11.3.1.1
Dipole radiation	$E_r = \frac{\sqrt{49.2\,P_t}}{d}$	V/m	11.3.1.1
Two dipoles spaced d with phase difference ϕ	$E \propto \cos\left(\frac{180d\cos\theta}{\lambda} \pm \frac{\phi}{2}\right)$	V/m	11.3.1.1

QUANTITY	FORMULA	UNIT	SECTION
Aerial efficiency	$\eta = \dfrac{100R_r}{R_r + R_L}$	%	11.3.1.2
Effective length of λ/2 dipole	$\ell_{eff} = \dfrac{\lambda}{\pi}$	m	11.3.2.1
Power delivered by aerial	$P_a = \dfrac{P_t G_t G_r \lambda^2}{(4\pi d)^2}$	W	11.3.2.2
Superheterodyne principle	$f_{if} = (f_o - f_s)$	Hz	11.4
Image channel interference	$f_{im} = (f_s + 2f_{if})$	Hz	11.4.1
Superhet tuned circuit inductance [Fig. 11.11(ii)]	L_s or L_o $= \dfrac{f_h^2 - f_\ell^2}{4\pi^2 f_h^2 \times f_\ell^2 (C_{s(max)} - C_{s(min)})}$	H	11.4.2
Trimmer (+ stray) capacitance	$C_t = \dfrac{1}{4\pi^2 f_h^2 L_s} - C_{s(min)}$	F	11.4.2

12. TRANSMISSION LINES

The need to transmit information over a distance brings two options, use of the free electromagnetic wave as outlined in Section 11.1 or alternatively by enclosing the signal within the confines of a *transmission line.* It is profitable first to gain a visual impression of how a wave is projected along a line.

12.1 PRACTICAL LINES

If a p.d. is applied to one end of a pair of wires, its appearance at the other end is not instantaneous simply because the disturbance created by the applied voltage needs time to "travel". We start with the proposition that there will be a wavelength, λ , following the normal radio formula,

$$\lambda = \frac{\nu}{f} = \nu T$$

where ν is the velocity of propagation and T is the wave period. Consider as an example, a frequency of 100 kHz at 1 V maximum to be applied to the sending-end of a loss-free line (no series or parallel resistance) as in Figure 12.1(i). At $t = 0$ the applied voltage is the maximum of 1 V and 10 μs later it will have travelled one wavelength. (The wave velocity varies according to the electrical characteristics of the line hence we cannot yet determine the actual wavelength.) Thus, using the line as the x-axis of a graph a point (i) can be plotted above it at length λ as shown. Point (iii) on the waveform (at -1 V) is not applied to the line until 5 μs later so it will only have travelled half as far and is therefore plotted at $\lambda/2$ away from the sending end. Points (ii) and (iv) on the waveform represent zero voltage and although we cannot talk in terms of zero flowing down a line, it is obvious that point (ii) occurs 7.5 μs later at $3\lambda/4$ and point (iv) at $\lambda/4$.

In (ii) of the Figure are shown the conditions 2.5 μs later, the +1 V will have progressed to $5\lambda/4$ and all other points will have similarly moved on by $\lambda/4$. It is now evident that the

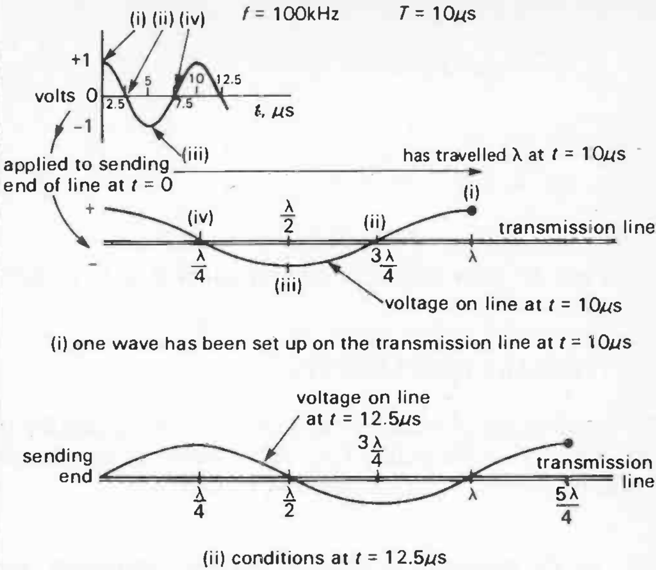


Fig. 12.1 Progress of a wave on a transmission line

waveform is travelling down the line and where at λ in (i) of the Figure there was originally 1 V, now, 2.5 μs later the voltage is zero and so at any point the voltage rises and falls according to the applied frequency but with a delay relative to the sending end. In terms of the angle through which the applied frequency has progressed, the phase difference must be 2π radians per wavelength. The change in phase per unit transmission line length is known as the *phase change coefficient*, generally represented by β and so:

$$\beta = \frac{2\pi}{\lambda} \text{ radians/m} \tag{12.1}$$

and since the velocity of the wave,

$$v = \frac{\lambda}{T} \text{ and also } \lambda f,$$

$$v = \frac{2\pi f}{\beta} \text{ m/s} \qquad (12.2)$$

We get to grips with the calculation of β and v later.

12.2 PRIMARY COEFFICIENTS

The main electrical characteristics of a line are stated by its four *primary coefficients.* They are quoted per unit length, nowadays the metre but the mile still persists:

(i) R, the *resistance* in ohms of the two conductors, also known as the *loop resistance.* Skin effect (Book 1, Sect.4.9) causes the value to rise with frequency.

(ii) G, the *leakance* in siemens arising from the conductance of the insulating material separating the two conductors.

(iii) C, the *capacitance* in farads is simply that between the two wires. The formula for the self-capacitance of two parallel wires with spacing D metres between their centres and each having a radius of r metres, for $D >> r$ is:

$$C = \frac{\pi\xi}{\log_e((D - r)/r)} \text{ farad/metre} \qquad (12.3)$$

where $\xi = \xi_0\xi_r$ (Sect.2.3), (the proof involves mathematical calculus so is not given here).

For the coaxial line:

$$C = \frac{2\pi\xi}{\log_e(R/r)} \text{ F/m} \qquad (12.4)$$

where R is the inner radius of the outer conductor and r is the outer radius of the inner conductor. So the capacitance falls as the tube size increases.

(iv) L, the *inductance* in henries for the two-wire loop is given by the formula:

$$L = \frac{\mu}{\pi} \log_e\left(\frac{D-r}{r}\right) \text{ henry/metre} \qquad (12.5)$$

where $\mu = \mu_0\mu_r$ (Sect.3.2.3), so as the spacing increases the inductance does also, hence an open-wire line has higher inductance than a cable pair.

For the coaxial line:

$$L = \frac{\mu}{2\pi} \log_e(R/r) \text{ henry/metre} \qquad (12.6)$$

EXAMPLE 12.1:
Calculate the primary coefficients (except the leakance which is negligible) for a 2.64 mm copper open-wire line with 30 cm spacing. (Resistivity, ρ of copper at 20°C = 1.725×10^{-8} ohm-metres, $\xi_0 = 8.854 \times 10^{-12}$ F/m, $\mu_0 = 4\pi \times 10^{-7}$ H/m. ξ_r and μ_r can be taken as 1.)

$D = 0.3$ m $\qquad r = 1.32$ mm.

Cross-sectional area of wire,

$$a = \pi \times 1.32^2 \text{ sq. mm} = 5.474 \times 10^{-6} \text{ sq. m}$$

Resistance of 1 m of wire is equal to

$$\frac{\rho}{a} = \frac{1.725 \times 10^{-8}}{5.474 \times 10^{-6}} \,\Omega = 0.00315 \,\Omega$$

$\therefore$ resistance of 1 m loop $= 0.00315 \times 2 = 0.0063 \,\Omega$.

From Equation (12.3)

$$C = \frac{\pi \times 8.854 \times 10^{-12}}{\log_e \dfrac{300 - 1.32}{1.32}} \text{ F/m}$$

$$= \frac{\pi \times 8.854}{5.42} \text{ pF/m} = 5.13 \text{ pF/m}.$$

From Equation (12.5)

$$L = \frac{4\pi \times 10^{-7}}{\pi} \times 5.42 \text{ H/m} = 0.4 \times 5.42 \,\mu\text{H/m}$$

$$= 2.17 \,\mu\text{H/m}.$$

The primary constants are therefore:

$R = 0.0063 \,\Omega/\text{m}$,

$G = 0$,

$C = 5.13$ pF/m,

$L = 2.17 \,\mu$H/m.

EXAMPLE 12.2:
The capacitance of a pair of paper-insulated wires of 1.3 mm diameter in an underground cable is 37 pF/m. What is the approximate wire spacing? ($\xi_0 = (10^{-9})/(36\pi)$ F/m

and the relative permittivity, ε_r of the paper and air between the conductors can be taken as 1.5.)

radius of wires, $r = 0.65 \times 10^{-3}$ m .

Let D = wire spacing, then from Equation (12.3)

$$37 \times 10^{-12} = \frac{\pi \times 10^{-9} \times 1.5}{36\pi \times \log_e \dfrac{D - (0.65 \times 10^{-3})}{0.65 \times 10^{-3}}}$$

$$\therefore \quad \log_e \frac{D - (0.65 \times 10^{-3})}{0.65 \times 10^{-3}} = \frac{10^{-9} \times 1.5}{36 \times 37 \times 10^{-12}} = 1.126$$

$$\therefore \quad \frac{D - (0.65 \times 10^{-3})}{0.65 \times 10^{-3}} = \text{antilog}_e\, 1.126 = 3.08$$

$$\therefore \quad D - (0.65 \times 10^{-3}) = 2.002 \times 10^{-3}$$

$$\therefore \quad D = 2.65 \times 10^{-3} \text{ m} = 2.65 \text{ mm},$$

or, if natural logarithms are not available, from Equation (A2.17):

$$37 \times 10^{-12} = \frac{\pi \times 10^{-9} \times 1.5}{36\pi \times 2.3026 \log_{10} \dfrac{D - (0.65 \times 10^{-3})}{0.65 \times 10^{-3}}}$$

which again results in $D = 2.65$ mm.

12.3 SECONDARY COEFFICIENTS

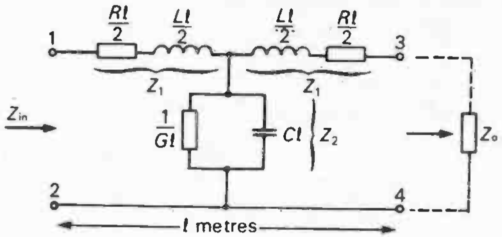
The four primary coefficients lead directly to the calculation of the *secondary coefficients* from which the efficiency of the line can be assessed. These coefficients are the *characteristic impedance*, Z_0, the *propagation coefficient*, p, which contains information on both the attenuation and phase change and finally the *velocity of propagation*, v.

12.3.1 Characteristic Impedance

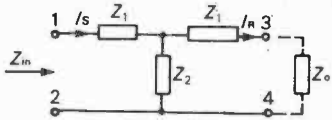
The impedance looking into one end of a transmission line varies according to the length, hence the latter must be quoted in any definition. As line length increases, the impedance tends to stabilize leading to the natural conclusion that for a definition of characteristic impedance the line should be infinitely long. Although the concept of infinity is purely theoretical, such a definition does not preclude practical use. This is because when a section is cut off from an infinite line, the remainder is still infinitely long hence the length cut off can be terminated by the characteristic impedance and it will continue to behave as though it were infinitely long.

To determine the value of the characteristic impedance, Z_0, from the primary coefficients we first consider a very short length of line, ℓ metres, so short in fact that the parameters can be considered as being lumped instead of distributed evenly. Remembering that the primary coefficients are quoted per metre length (or per mile) then the length ℓ can be drawn as an unbalanced T-section as in Figure 12.2(i). For most lines the balanced T-section is more appropriate but this only complicates the issue and does not affect the conclusions. The leakance, G which is a conductance has been changed to a resistance, $1/G$ so that we can work entirely on an impedance basis. Because leakance per metre is G, then for a length, ℓ it is $G\ell$ and the equivalent resistance is $1/G\ell$.

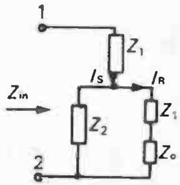
As shown above the T-section can be considered as being terminated in Z_0 and the problem is to calculate the impedance looking into terminals 1 and 2. By considering first an equivalent generalized T network as in Figure 12.2(ii), a simplified expression for Z_{in} can be obtained. Operator j is invaluable here and Chapter 4 contains its whole story. In



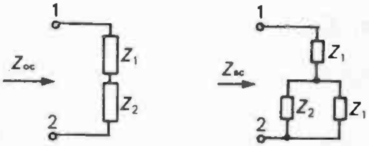
(i) T-network of short length of line



(ii) equivalent generalized network



(iii) equivalent network re-drawn



(terminals 3 & 4 o/c) (terminals 3 & 4 s/c)

(iv) the T-network under open and short-circuit conditions

Fig. 12.2 Calculation of characteristic impedance this case:

$$Z_1 = (R + j\omega L) \times \frac{\ell}{2}$$

$$Z_2 = \frac{\frac{1}{G\ell} \times \frac{1}{j\omega C\ell}}{\frac{1}{G\ell} + \frac{1}{j\omega C\ell}} = \frac{\frac{1}{j\omega CG\ell^2}}{\frac{G\ell + j\omega C\ell}{j\omega CG\ell^2}} = \frac{1}{(G + j\omega C)\ell}$$

then from (iii) of the Figure:

$$Z_{in} = Z_1 + \frac{Z_2(Z_1 + Z_0)}{Z_1 + Z_2 + Z_0}$$

$$= \frac{Z_1{}^2 + Z_1 Z_0 + 2Z_1 Z_2 + Z_2 Z_0}{Z_1 + Z_2 + Z_0}$$

and because by definition $Z_{in} = Z_0$

$$Z_1 Z_0 + Z_2 Z_0 + Z_0{}^2 = Z_1{}^2 + Z_1 Z_0 + 2Z_1 Z_2 + Z_2 Z_0$$

$$\therefore \quad Z_0{}^2 = Z_1{}^2 + 2Z_1 Z_2 \qquad (12.7)$$

$$\therefore \quad Z_0 = \sqrt{Z_1{}^2 + 2Z_1 Z_2}$$

and next by substituting for Z_1 and Z_2 :

$$Z_0 = \sqrt{\frac{(R + j\omega L)^2 \ell^2}{4} + \frac{(R + j\omega L)}{(G + j\omega C)}}$$

This is the answer for a finite length of line. No matter how short this length is, we only arrive at the true answer for Z_0 when the length approaches zero for then the primary coefficients are evenly distributed as they are in practice. Accordingly as ℓ tends to zero, the first term under the root sign gets less and at the limit disappears, leaving:

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \ \Omega \qquad (12.8)$$

EXAMPLE 12.3:
A 6.8 mm diameter coaxial cable has the following primary coefficients at 100 kHz:

$$R = 0.02 \ \Omega/\text{m}$$

$$G = 0.03 \ \mu\text{S/m}$$

$$C = 48 \ \text{pF/m}$$

$$L = 0.3 \ \mu\text{H/m}$$

What is its characteristic impedance at this frequency?

From Equation (12.8)

$$Z_0 = \sqrt{\frac{0.02 + j(2\pi \times 10^5 \times 0.3 \times 10^{-6})}{0.03 \times 10^{-6} + j(2\pi \times 10^5 \times 48 \times 10^{-12})}}$$

$$= \sqrt{\frac{0.02 + j0.1885}{0.03 \times 10^{-6} + j30.16 \times 10^{-6}}}$$

$$= 10^3 \sqrt{\frac{0.02 + j0.1885}{0.03 + j30.16}}$$

$$= 10^3 \sqrt{\frac{0.1896 \angle 83.9°}{30.16 \angle 89.9°}}$$

(note the change from cartesian to polar coordinates to simplify division and square root)

$$\therefore \quad Z_0 = 10^3(0.079\angle{-3°}) = 79\angle{-3°}\ \Omega$$

(only the modulus is multiplied by 10^3, not the angle).

Throughout this calculation it is evident that in each complex quantity the real component is small compared with the imaginary. This is frequently a feature of coaxial cables working at the higher frequencies and allows us to calculate Z_0 approximately by:

$$Z_0 \approx \sqrt{\frac{j\omega L}{j\omega C}} = \sqrt{\frac{L}{C}} \qquad (12.9)$$

The coaxial cable of the example above demonstrates this point for with $C = 48$ pF/m and $L = 0.3\ \mu$H/m

$$Z_0 \approx \sqrt{\frac{0.3 \times 10^{-6}}{48 \times 10^{-12}}} = \sqrt{0.00625 \times 10^6} = 79\angle 0°$$

also showing that for most coaxial cables the characteristic impedance is almost entirely resistive.

12.3.1.1 Calculation from Physical Dimensions

From the above approximation, i.e. $Z_0 \approx \sqrt{L/C}$ we can obtain a formula for the calculation of Z_0 from the line dimensions instead of from the primary coefficients. For the open-wire line with wire radius r and spacing D and for which μ_r and ξ_r both = 1, since

$$\mu_0 = 4\pi \times 10^{-7}\ \text{H/m}$$

and

$$\xi_0 = \frac{1}{36\pi \times 10^9}\ \text{F/m},$$

then from Equation (12.5):

$$L = 4 \times 10^{-7} \log_e \left(\frac{D-r}{r}\right)$$

and from Equation (12.3):

$$C = \frac{1}{36 \times 10^9} \times \frac{1}{\log_e \left(\frac{D-r}{r}\right)}$$

$$\therefore \quad \frac{L}{C} = 4 \times 10^{-7} \times 36 \times 10^9 \left[\log_e \left(\frac{D-r}{r}\right)\right]^2$$

$$\therefore \quad \sqrt{\frac{L}{C}} \approx Z_0 = 120 \log_e \left(\frac{D-r}{r}\right)$$

$$= 120 \times 2.3026 \log_{10} \left(\frac{D-r}{r}\right)$$

i.e. $$Z_0 \approx 276 \log_{10} \left(\frac{D-r}{r}\right) \Omega \qquad (12.10)$$

For the coaxial cable with mainly air dielectric, using Equations (12.6) and (12.4) and similar reasoning:

$$Z_0 \approx 138 \log_{10} \frac{R}{r} \ \Omega \qquad (12.11)$$

and when a solid dielectric is employed:

$$Z_0 \approx \frac{138}{\sqrt{\xi_r}} \times \log_{10} \frac{R}{r} \ \Omega \qquad (12.12)$$

EXAMPLE 12.4:
A flexible 6.8 mm coaxial cable has an inner conductor of 1.07 mm diameter. The polythene dielectric has a relative permittivity of 2.25. What is the approximate characteristic impedance?

$$R = \frac{6.8}{2} \text{ mm}, \qquad r = \frac{1.07}{2} \text{ mm}, \qquad \xi_r = 2.25$$

From Equation (12.12):

$$Z_0 \approx \frac{138}{\sqrt{2.25}} \log_{10} \frac{6.8}{1.07} = 92 \times 0.803 = 74 \ \Omega$$

12.3.1.2 Practical Measurement
Although the characteristic impedance of a long transmission line is given fairly accurately by a simple measurement of impedance at one end, this does not apply to relatively short lengths. However, with cooperation at the distant end if required to open-circuit the line and then to short-circuit it, two impedance measurements only are required for accurate calculation instead. Consider the line as a T-network as in Figure 12.2(ii) and that measurements are being made at terminals 1 and 2, then if terminals 3 and 4 are open-circuited and short-circuited the electrical conditions are as in (iv) and

$$Z_{oc} = Z_1 + Z_2$$

$$Z_{sc} = Z_1 + \frac{Z_1 Z_2}{Z_1 + Z_2}$$

$$\therefore \quad Z_{sc} = \frac{Z_1^2 + 2Z_1 Z_2}{Z_1 + Z_2}$$

which from Equation (12.7) becomes

$$\frac{Z_0^2}{Z_1 + Z_2}$$

i.e. $$Z_{sc} = \frac{Z_0^2}{Z_{oc}}$$

or $$Z_0 = \sqrt{Z_{oc} \times Z_{sc}} \qquad (12.13)$$

EXAMPLE 12.5:
Impedance measurements at 10 kHz are made at one end of a pair of wires in a 10 metre length of 2.6 mm underground cable when the opposite end is open-circuited (Z_{oc}) and then short-circuited (Z_{sc}). The results are:

$$Z_{oc} = 42.3\ \text{k}\Omega \angle{-89°}$$

$$Z_{sc} = 0.51\ \Omega \angle 57°$$

What is the characteristic impedance of the cable pair?

From Equation (12.13) the characteristic impedance,

$$Z_0 = \sqrt{Z_{oc} \times Z_{sc}} = \sqrt{42300 \angle{-89°} \times 0.51 \angle 57°}$$

$$= \sqrt{21573 \angle{-32°}} = 147 \angle{-16°}\ \Omega$$

[or in cartesian coordinates: $147 \cos 16° - \text{j}147 \sin 16° = 141 - \text{j}40.5\ \Omega$.

12.3.2 Propagation Coefficient

This has a dual role and is in complex form. Perhaps the more important part of its duality is the *attenuation coefficient* (α) for this real part expresses the loss of the transmission line per unit length, an essential factor in the judgment of line efficiency. The imaginary part is known as the *phase-change coefficient* (β) which we met briefly in Section 12.1.

The propagation coefficient is labelled p (γ is also used) and by definition it is the natural logarithm of the ratio of the currents (or voltages) at two points on the line, it is complex because the currents are not in phase (Sect.12.1). Thus if the sent and received currents at the two points are I_S and I_R as in Figure 12.2(ii):

$$p = \log_e \frac{I_S}{I_R} = \alpha + j\beta \qquad (12.14)$$

and taking the antilog of both sides:

$$e^p = \frac{I_S}{I_R} \qquad (12.15)$$

Figure 12.2(iii) shows a little more clearly the division of current within the generalized network and it follows that

$$\frac{I_R}{I_S} = \frac{Z_2}{Z_1 + Z_2 + Z_0}$$

or

$$\frac{I_S}{I_R} = 1 + \frac{Z_1}{Z_2} + \frac{Z_0}{Z_2}$$

Now, for a short length of line, ℓ, we have found that

$$Z_1 = \frac{(R + j\omega L)\ell}{2}$$

$$Z_2 = \frac{1}{(G + j\omega C)\ell}$$

and the current ratio is $e^{p\ell}$ (not $e^p \times \ell$ as might be expected at first). Take as an example, 1 m of line with a ratio e^p, for 2 m it is $e^p \times e^p = e^{2p}$, hence for ℓm, $e^{p\ell}$) therefore:

$$e^{p\ell} = 1 + \frac{(R + j\omega L)(G + j\omega C)\ell^2}{2}$$

$$+ \sqrt{\frac{R + j\omega L}{G + j\omega C}} \times (G + j\omega C)\ell$$

$$= 1 + \sqrt{(R + j\omega L)(G + j\omega C)}\ell$$

$$+ \frac{(R + j\omega L)(G + j\omega C)\ell^2}{2}$$

Next from Equation (A2.25):

$$e^{p\ell} = 1 + p\ell + \frac{p^2\ell^2}{2!} + \frac{p^3\ell^3}{3!} + \ldots \text{ to } \infty$$

now ℓ is very small and as it approaches zero, terms containing ℓ^2 and higher orders can be neglected so comparing the two expressions above for $e^{p\ell}$ it is evident that:

$$p = \sqrt{(R + j\omega L)(G + j\omega C)} \qquad (12.16)$$

[note – we have to be content with this mathematical proof, a better one is given using *hyperbolic functions* but to gain an understanding of these first would make this Chapter disproportionately large.]

The following example shows how the expression is handled and the information which can be gained from it.

EXAMPLE 12.6:
A 0.91 mm paper-insulated trunk telephone cable has the following primary coefficients at 100 kHz:

$$R = 0.085\ \Omega/\text{m}$$

$$G = 0.28\ \mu\text{S/m}$$

$$C = 37.2\ \text{pF/m}$$

$$L = 0.63\ \mu\text{H/m}\ .$$

Calculate the characteristic impedance, Z_0 and propagation coefficient, p . What is the cable pair attenuation in dB/km?

$$R + j\omega L = 0.085 + j2\pi \times 10^5 \times 0.63 \times 10^{-6}$$

$$= 0.085 + j0.396 = 0.405 \angle 77.9^\circ$$

$$G + j\omega C = 0.28 \times 10^{-6} + j2\pi \times 10^5 \times 37.2 \times 10^{-12}$$

$$= 10^{-6}(0.28 + j23.37) = 10^{-6}(23.37 \angle 89.3^\circ)$$

From Equation (12.8):

$$Z_0 = 10^3 \sqrt{\frac{0.405 \angle 77.9^\circ}{23.37 \angle 89.3^\circ}} = 10^3\sqrt{0.0173 \angle -11.4^\circ}$$

$$= 132 \angle -5.7^\circ$$

or, in cartesian form:

$$132(\cos -5.7^\circ + j \sin -5.7^\circ) = 131 - j13$$

From Equation (12.16):

$$p = 10^{-3}\sqrt{(0.405 \angle 77.9^\circ)(23.37 \angle 89.3^\circ)}$$

$$= 10^{-3}\sqrt{9.46 \angle 167.2^\circ} = 10^{-3}(3.08 \angle 83.6^\circ)$$

So far we have calculated p in polar notation. In this form it tells us very little, so we reveal the components by expressing it in cartesian form:

$$p = \alpha + j\beta = 10^{-3} \times 3.08 (\cos 83.6^\circ + j \sin 83.6^\circ)$$

$$= 10^{-3}(0.343 + j3.06),$$

i.e. the attenuation coefficient, $\alpha = 0.343 \times 10^{-3}$ nepers/metre and the phase-change coefficient, $\beta = 3{,}06 \times 10^{-3}$ radians/metre.

In more practical units (most transmission loss measurements are made in decibels, not nepers),

$$\alpha = 0.343 \times 8.686 = 2.98 \text{ dB/km}$$

(one neper is equivalent to 8.686 decibels – see Appendix 1).

12.3.3 Velocity of Propagation

As a by-product in the example above we calculated the value of β as 3.06×10^{-3} radians/metre. Using this as an example, from Equation (12.2) the velocity of propagation at 100 kHz:

$$v = \frac{\omega}{\beta} = \frac{2\pi \times 10^5}{3.06 \times 10^{-3}} = 2.05 \times 10^8 \text{ m/s},$$

considerably lower than the speed of electromagnetic waves in free space (Sect.11.1.1). The wavelength, λ follows from Equation (12.1):

$$\lambda = \frac{2\pi}{\beta} = \frac{2\pi}{3.06 \times 10^{-3}} \text{ m} = 2053 \text{ metres}$$

(at 100 kHz), shorter than for an electromagnetic wave in free space because of the lower velocity.

12.4 SUMMARY OF KEY FORMULAE

C = capacitance

D = spacing between wires

f = frequency

G = leakance (S)

L = inductance

R = inner radius

R = resistance

r = radius of conductor

Z_{oc} = line impedance measured with distant-end open-circuited

Z_{sc} = line impedance measured with distant-end short-circuited

ξ = permittivity

μ = permeability

λ = wavelength

QUANTITY	FORMULA	UNIT	SECTION
Transmission Lines:			
Capacitance (parallel wires)	$C = \frac{\pi\xi}{\log_e((D-r)/r)}$	F/m	12.2
Capacitance (coaxial)	$C = \frac{2\pi\xi}{\log_e(R/r)}$	F/m	12.2
Inductance (parallel wires)	$L = \frac{\mu}{\pi} \log_e((D-r)/r)$	H/m	12.2
Inductance (coaxial)	$L = \frac{\mu}{2\pi} . \log_e(R/r)$	H/m	12.2
Characteristic impedance	$Z_0 = \sqrt{Z_{oc} \times Z_{sc}}$	Ω	12.3.1.2
	$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \approx \sqrt{\frac{L}{C}}$	Ω	12.3.1

QUANTITY	FORMULA	UNIT	SECTION
Characteristic impedance (contd.)			
(parallel wires)	$Z_0 = 276 \log_{10}((D - r)/r)$	Ω	12.3.1.1
(coaxial)	$Z_0 = 138 \log_{10}(R/r)$	Ω	12.3.1.1
(coaxial with solid dielectric)	$Z_0 = \frac{138}{\sqrt{\xi_r}} \log_{10}(R/r)$	Ω	12.3.1.1
Propagation coefficient	$p = \sqrt{(R + j\omega L)(G + j\omega C)}$		12.3.2
Phase-change coefficient	$\beta = \frac{2\pi}{\lambda}$	radians/m	12.1
Velocity of propagation	$\nu = \frac{2\pi f}{\beta}$	m/s	12.1

13. DIGITAL LOGIC

Digital logic embraces discrete signals switched "on" or "off" and the original digital logic device is the electromagnetic relay. It is superseded nowadays in data processing by the semiconductor "gate", thousands of which can be accommodated within a single I.C. The analysis of logic was first proposed by George Boole(A3) in the mid-eighteen-hundreds and it has led to what is now called *Boolean algebra.* This is especially useful in the study of information flow in a complex switching network (such as employed in a computer), usually with the aim of minimizing the number of gates required for a particular function.

The two binary states may be described by "on" and "off", "high" and "low", "true" and "false" etc., but generally in digital logic considerations as "logic 1" and "logic 0" or perhaps most frequently by simply "1" and "0".

13.1 BINARY NUMBERS

A number system is characterized by its *base* or *radix* (from Latin, root) and it requires a range of symbols equal in number to the radix. As an example our everyday system has a radix of 10 and 10 symbols, 0 to 9. For electronic systems the near certainty of being able to decide between two states only (e.g. ON or OFF, pulse or no pulse etc.) requires a number system with a radix of 2 (binary). Two symbols only are therefore required, the now familiar 1 and 0.

In general for any number system an integer (whole number) N comprising n digits is expressed by:

$$N = a_{n-1}r^{n-1} + a_{n-2}r^{n-2} + \ldots$$

$$a_2 r^2 + a_1 r^1 + a_0 r^0 \qquad (13.1)$$

where a_0 , a_1 etc. are the digit values and r is the radix.

Taking a decimal number, say 205 as an example, then $a_0 = 5$, $a_1 = 0$, $a_2 = 2$ and $r = 10$. The number is analysed by Equation (13.1) as:

$$(2 \times 10^2) + (0 \times 10^1) + (5 \times 10^0) \qquad (10^0 = 1)$$

Taking a binary number, say 1010010, $n = 7$, $r = 2$ and working this time from left to right,

$$a_6 = 1, a_5 = 0, a_4 = 1, a_3 = 0, a_2 = 0, a_1 = 1, a_0 = 0$$

the number is therefore expressed in the terms of Equation (13.1) as

$$N = (1 \times 2^6) + (0 \times 2^5) + (1 \times 2^4) + (0 \times 2^3)$$

$$+ (0 \times 2^2) + (1 \times 2^1) + (0 \times 2^0)$$

which in the decimal system is 64 + 0 + 16 + 0 + 0 + 2 + 0 = 82 (Table 13.1 gives powers of 2). Each of the digits in a binary number is called a *bit* (from *b*inary dig*it*).

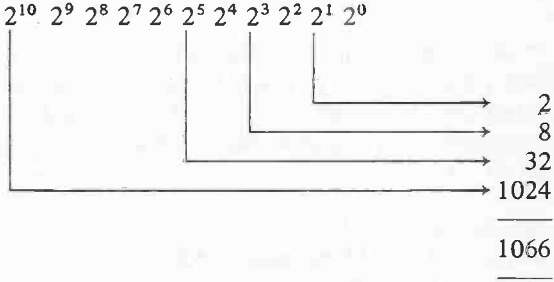
Table 13.1: POWERS OF 2

n	2^n	n	2^n
0	1	10	1024
1	2	11	2048
2	4	12	4096
3	8	13	8192
4	16	14	16,384
5	32	15	32,768
6	64	16	65,536
7	128	17	131,072
8	256	18	262,144
9	512	19	524,288
		20	1,048,576

13.1.1 Conversion from Binary to Decimal

This follows from the analysis of the binary number above. Take as an example the number $N = 10000101010$. It has 11 digits $\therefore n = 11$. From Equation (13.1) the left-hand or most significant digit is $a_{n-1}r^{n-1}$, i.e. $a_{10}r^{10} = 1 \times 2^{10}$ which from Table 13.1 = 1024. We need only be concerned with the 1's in the number so the next is $(1 \times 2^5) = 32$, followed by $(1 \times 2^3) = 8$ and $(1 \times 2^1) = 2$. The total is 1066, so binary 10000101010 = decimal 1066 and it is worthy of note that more digits are required to express a number in a lower radix. The same method but arranged more methodically is as follows:

$$N = 1\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0$$



Consideration is given to the "1" bits only.

13.1.2 Conversion from Decimal to Binary

Table 13.1 is again helpful for this process, a search is first made for the largest power of 2 which is smaller than the number N. The power of 2 is subtracted from N and the process is repeated with the remainder until the latter is exhausted. The binary number is then constructed with a 1 in each position corresponding to the power of 2 used for it and 0's in all other positions. As an example, conversion of say, the decimal number 107 to binary might be set out as follows:

		2^6	2^5	2^4	2^3	2^2	2^1	2^0	
Number N =	107	1	1	0	1	0	1	1	← binary
subtract 2^6 =	64	↑							
	43								
subtract 2^5 =	32		↑						
	11								
subtract 2^3 =	8				↑				
	3								
subtract 2^1 =	2						↑		
	1								
subtract 2^0 =	1							↑	
	0								

A second method successively divides the number by 2 and a note is made of the remainder. An even number has no remainder when divided by 2 but an odd number has a remainder of 1. Converting the decimal number 107 again gives:

Number N =	107		1	1	0	1	0	1	1	← binary
divide by 2 =	53,	remainder	↑							
,, ,, ,, =	26	,,		↑						
,, ,, ,, =	13	,,			↑					
,, ,, ,, =	6	,,				↑				
,, ,, ,, =	3	,,					↑			
,, ,, ,, =	1	,,						↑		
,, ,, ,, =	0	,,							↑	

13.2 LOGIC GATES

The function of a *gate* is to change its output signal when certain conditions are fulfilled at the input. Here we consider the three basic gates known as AND, OR and NOT (the significance of these somewhat peculiar titles is made evident later) and three further ones built on these basic forms. One device is examined first in some detail for experience with the term "gate" and also for an introduction to the special algebra.

13.2.1 The AND Gate

The simplest form has two inputs and only when both are at logic 1 does the output go to 1. The function could be achieved by using two electromagnetic switches or "relays", each with a single "make" contact and the two contacts connected in series. Both relays must be energised for current to flow through their contacts as shown in Figure 13.1(i). In detail, a voltage applied only to input A operates relay A but this has no effect on the output f which therefore remains at logic 0 (no voltage), neither does it if relay B on its own is operated. But when both relays A AND B are operated, +5V (the logic 1) appears at f.

A simple diode AND gate circuit is shown at (ii), this having three inputs and demonstrating that any number of inputs can be provided. When any input logic signal is at 0, (e.g. B in the Figure) current flows through that particular diode, maintaining it at low resistance and creating a voltage drop across R sufficient to keep the output terminal f near 0 V. Only when all input logic signals are at 1 is the heavy current through R cut off with the potential at f allowed to rise to +V, i.e. logic 1.

The mode of operation of any gate is described by its *truth table* as shown in (iv) for a two-input gate. The table shows all possible input combinations with the effect of each on the output. In this particular example f goes positive in switching to 1, this is known as *positive logic*, *negative logic* has the opposite polarity.

To express the operation of the gate algebraically certain operators are used which are not the same as in conventional algebra and until one gets used to them, they can be misleading. The logical equation for a three or more input AND gate is

$$f = A \, . \, B \, . \, C \qquad (13.2)$$

the dot or full-stop indicating the AND function. The equa-

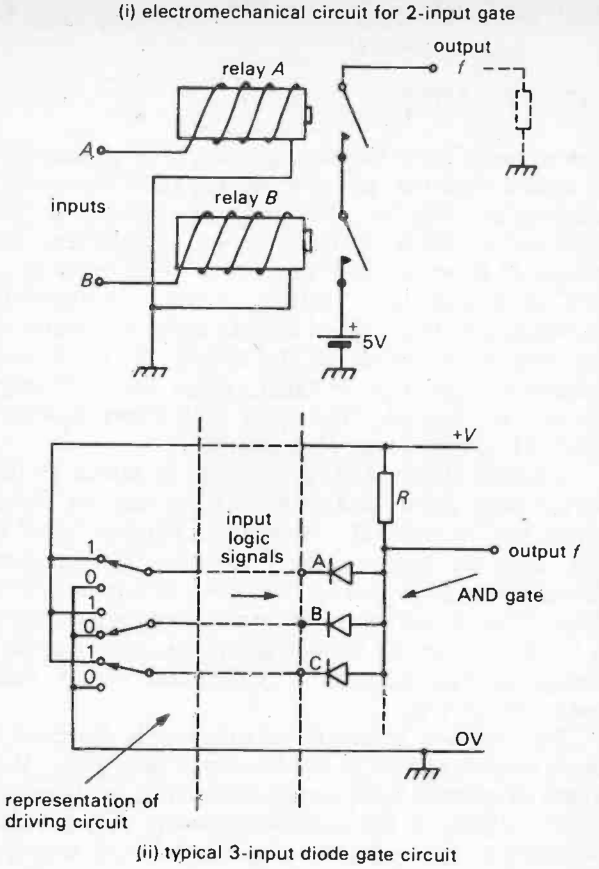
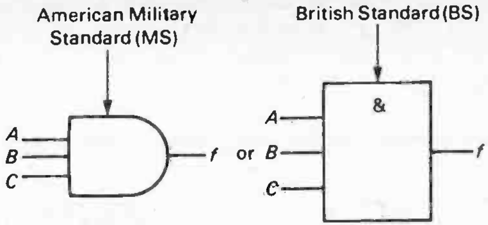


Fig. 13.1 The AND gate

tion therefore reads that *f* goes to 1 when *A* and *B* and *C* (and any additional inputs) are at 1.

The symbols in general use for an AND gate are shown in Figure 13.1(iii).



(iii) symbols

inputs		output
A	B	f
0	0	0
0	1	0
1	0	0
1	1	1

(iv) truth table for a 2-input gate

Fig. 13.1 (contd)

13.2.2 The OR Gate

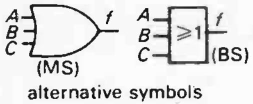
A two-input OR gate can be represented by two relay contacts in parallel so that the circuit is switched through when either one OR the other contact closes. The symbols in use are shown in Figure 13.2(i) together with the truth table. The OR function is represented in logic algebra by a + sign, again, not to be confused with the + of arithmetic, hence for a three or more input OR gate:

$$f = A + B + C + \qquad (13.3)$$

13.2.3 The NOT Gate

The simplest of all, it changes 0 to 1 and 1 to 0, generally termed *inverting* or *complementing*. The function is indicated by use of a horizontal bar which is extended over all the inputs affected, i.e. NOT A is shown as $\bar{A}$, hence

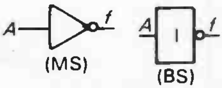
$$f = \bar{A} \tag{13.4}$$



Inputs		output
A	B	f
0	0	0
0	1	1
1	0	1
1	1	1

truth table

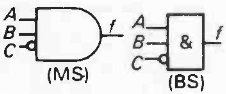
(i) OR gate

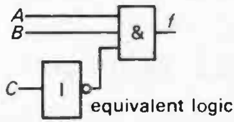


Input	Output
A	f
0	I
I	0

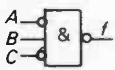
truth table

(ii) NOT gate

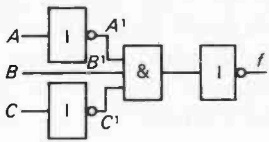




(iii) AND gate with one inverted input



(a) combined symbol (BS)

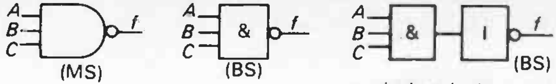


(b) equivalent logic (BS)

(iv) a gate with inversion at both input and output

[(MS) and (BS) — see Fig. 13.1]

Fig. 13.2 Logic gates



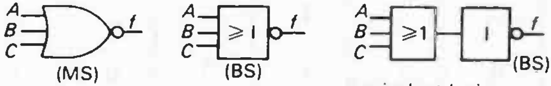
equivalent logic

alternative symbols

Inputs		Output
A	B	f
0	0	1
0	1	1
1	0	1
1	1	0

truth table

(v) NAND gate



equivalent logic

alternative symbols

Inputs		Output
A	B	f
0	0	1
0	1	0
1	0	0
1	1	0

truth table

(vi) NOR gate

Fig. 13.2 (contd)

The symbols and truth table are shown in Figure 13.2(ii). The small circle on a symbol is a general indication of the NOT function, inversion or *negation*, it can be at the input, output or both as shown in (iii) where the C input is inverted, hence $f = A \,.\, B \,.\, \bar{C}$, meaning that f goes to logic 1 only when A and B are 1 and C is 0. In the following example some of the rules are put into practice.

EXAMPLE 13.1:
The drawing in Figure 13.2(iv) shows a type of gate at (a) and its constituents at (b). Develop the logic formula and truth table.

As far as the AND gate in (b) is concerned, from Equation (13.2), its formula is $f = A \,.\, B \,.\, C$. However, the A and C inputs are inverted:

$$\therefore \quad f = \bar{A} \,.\, B \,.\, \bar{C}$$

and also because the output is inverted

$$f = \overline{\bar{A} . B . \bar{C}}$$

The truth table follows:

Inputs			Inputs to AND gate			Output f
A	B	C	A^1	B^1	C^1	
0	0	0	1	0	1	1
0	0	1	1	0	0	1
0	1	0	1	1	1	0
0	1	1	1	1	0	1
1	0	0	0	0	1	1
1	0	1	0	0	0	1
1	1	0	0	1	1	1
1	1	1	0	1	0	1

(Columns 4–6 are not normally included in a truth table but are added here for clarification.)

Hence the output f is normally at 1 and changes to 0 only when inputs A and C are at 0 and B is at 1.

13.2.4 The NAND Gate

This is short for NOT AND and the device is simply an AND gate followed by a NOT gate as shown in Figure 13.2(v) with the truth table having an opposite output to that of the AND gate. The logical formula for a 3-input gate is

$$f = \overline{A . B . C} \qquad (13.5)$$

13.2.5 The NOR Gate

In this case, an OR gate followed by a NOT gate, in other words, a negated OR. The symbols and truth table are in (vi) of the Figure, the formula for a 3-input gate being:

$$f = \overline{A + B + C} \tag{13.6}$$

(a reminder, the + stands for OR).

13.2.6 The Exclusive-OR Gate

This is a 2-input gate with the same truth table as for the OR except that the condition of A and B both at 1 is excluded from changing the output to 1 as shown in Figure 13.3(ii). It is also of interest that f only goes to 1 when there is an odd number of 1's in the inputs and this condition persists no matter how many inputs are used. The exclusive-OR function has its own special algebraic symbol $\oplus$ hence

$$f = A \oplus B \tag{13.7}$$

The equivalent logic diagram contains 5 elementary gates as developed in the next example.

EXAMPLE 13.2:
The truth table for a 2-input exclusive-OR gate is given in Figure 13.3(ii). Develop a logic arrangement for this facility.

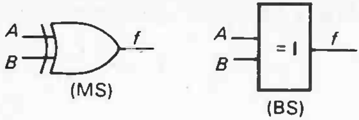
From the truth table f changes to 1 when

(a) $A = 0$, $B = 1$ or equally $\bar{A} = 1$, $B = 1$

(for if $A = 0$ then $\bar{A} = 1$).

OR

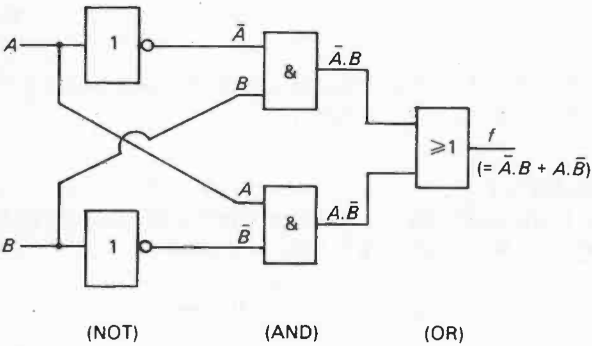
(b) $A = 1$, $B = 0$ or equally $A = 1$, $\bar{B} = 1$,



(i) alternative symbols

Inputs		Output
A	B	f
0	0	0
0	1	1
1	0	1
1	1	0

(ii) truth table



(iii) equivalent logic diagram

Fig. 13.3 The exclusive-OR gate

so we can write the algebraic expression as

$$f = \bar{A} . B + A . \bar{B} .$$

The horizontal bars indicate that NOT gates are required (2), the stops are for AND gates (2) and the + is for an OR gate (1). A suitable arrangement is given in Figure 13.3(iii) where $\bar{A} . B$ is produced at the output of the upper AND gate and $A . \bar{B}$ at the lower. Both are then applied to the single OR gate. The complete operation is developed in the Table on page 390.

13.3 LOGIC ALGEBRA

Through this algebra, systems are designed as in the above example, also the number of gates can be *minimized* for maximum efficiency. All operations are realized from the three basic functions AND, OR, NOT and in the computer and other data processing systems the practical outcome is most likely to be as semiconductor logic circuits, usually in IC form. In the algebra brackets have the same significance in the order of performing operations as with conventional algebra, e.g.

$$A . B + C = (A . B) + C \text{, not } A . (B + C) .$$

The AND operation is therefore completed first.

As with electronics generally laws and theorems have been developed to aid the solution of equations.

13.3.1 Laws

These are of considerable help in disentangling complicated expressions:

(i) *The Associative Law* states that the order in which *identical* operations are performed in an equation is irrelevant, hence

Inputs		$\bar{A}$	$\bar{B}$	Input to upper AND gate		Input to lower AND gate		Upper AND gate output	Lower AND gate output	Output f
A	B			$\bar{A}$	B	A	$\bar{B}$			
0	0	1	1	1	0	0	1	0	0	0
0	1	1	0	1	1	0	0	1	0	1
1	0	0	1	0	0	1	1	0	1	1
1	1	0	0	0	1	1	0	0	0	0

$$A \cdot B \cdot C = (A \cdot B) \cdot C = A \cdot (B \cdot C) \qquad (13.8)$$

so that if first, for example, the result of A AND B is found to be equal to X then subsequently the outcome of X AND C gives the final answer. Also

$$A + B + C = (A + B) + C = A + (B + C) \qquad (13.9)$$

(ii) *The Commutative Law* – so called by a title derived from Latin, "to exchange" this law states that the order in which terms are arranged in an equation is irrelevant, hence

$$A \cdot B = B \cdot A \qquad (13.10)$$

and

$$A + B = B + A \qquad (13.11)$$

this perhaps being obvious from our knowledge of the operation of AND and OR gates.

(iii) *The Distributive Law* – This refers to two operations which have similarities with conventional algebra.

$$A \cdot (B + C) = A \cdot B + A \cdot C \qquad (13.12)$$

$$A + (B \cdot C) = (A + B) \cdot (A + C) \qquad (13.13)$$

EXAMPLE 13.3:
Prove Equation (13.12).

This can be done by construction of a table showing all combinations of A, B and C and the outputs of both sides of the equation, these must obviously be the same. Any other law or theorem can be proved similarly (see Table on page 393). Columns 5 and 8 which represent the two sides of Equation (13.12) are identical for all combinations of A, B and C, thus proving the equality.

13.3.2 Theorems

There are several rules or theorems, some obvious, others perhaps not so obvious but all can easily be checked as in the example above. Relative to the AND function and based on the fact that $0 . 0 = 0$, $1 . 0 = 0 . 1 = 0$, $1 . 1 = 1$ then

$$A . 0 = 0 \tag{13.14}$$

$$A . 1 = A \tag{13.15}$$

$$A . A = A \tag{13.16}$$

Relative to the OR function and based on the fact that

$$0 + 0 = 0, \quad 0 + 1 = 1 + 0 = 1, \quad 1 + 1 = 1$$

$$A + 0 = A \tag{13.17}$$

$$A + 1 = 1 \tag{13.18}$$

$$A + A = A \tag{13.19}$$

and next including the NOT function

$$A . \overline{A} = 0 \tag{13.20}$$

$$A + \overline{A} = 1 \tag{13.21}$$

$$\overline{\overline{A}} = A \tag{13.22}$$

There are several more, two useful ones are

$$A + A . B = A \tag{13.23}$$

$$A . \overline{B} + B = A + B \tag{13.24}$$

13.3.2.1 de Morgan's Theorem

This has been developed for interpretation of the complement

A	B	C	$(B+C)$	$A.(B+C)$	$A.B$	$A.C$	$A.B+A.C$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	0	0	0	0	0	0	0
0	0	1	1	0	0	0	0
0	1	0	1	0	0	0	0
0	1	1	1	0	0	0	0
1	0	0	0	0	0	0	0
1	0	1	1	1	0	1	1
1	1	0	1	1	1	0	1
1	1	1	1	1	1	1	1

or inversion of an expression such as $A \,.\, B$ or $A + B$, these for example are shown to be equivalent to

$$\overline{A \,.\, B} = \bar{A} + \bar{B} \tag{13.25}$$

$$\overline{A + B} = \bar{A} \,.\, \bar{B} \tag{13.26}$$

and it will be seen that each term on the left-hand side of the expression is separately inverted and the AND function changed to OR and vice versa. Any number of inputs can be so treated, e.g.

$$\overline{A \,.\, B \,.\, C} = \bar{A} + \bar{B} + \bar{C}$$

Practical use of the laws and theorems follows in the next Section.

13.3.3 Algebraic Minimization

This is most easily illustrated by means of examples. The technique relies on the use of the various laws and theorems to reduce the number of terms in an expression and therefore the number of gates to be employed.

EXAMPLE 13.4:
Three relays, A, B and C on a robot control the starting of a motor which occurs when

(i) A only is operated

(ii) B only is operated

(iii) A and C are operated

(iv) B and C are operated.

No other sequence of relay operation must be effective. Draw the appropriate logic diagram and then minimize.

The conditions (i) to (iv) can be assembled into an equation as follows:

$$f = A.\overline{B}.\overline{C}+\overline{A}.B.\overline{C}+A.\overline{B}.C+\overline{A}.B.C$$

which could be achieved by four AND gates feeding into one OR gate as shown in Figure 13.4(i). The commutative law allows us to rearrange the order.

$$\therefore \quad f = A.\overline{B}.C+A.\overline{B}.\overline{C}+\overline{A}.B.C+\overline{A}.B.\overline{C}$$

and from the distributive law:

$$f = A.\overline{B}(C+\overline{C})+\overline{A}.B(C+\overline{C})$$

and from Equation (13.21), $(C + \overline{C}) = 1$, hence the system reduces to

$$f = A.\overline{B}+\overline{A}.B$$

so there is a possibility that the facility can be provided with two relays only and using 5 gates instead of 8 as in Figure 13.4(ii). It is a worthwhile exercise at the design stage but it should be appreciated that in this particular case the minimal solution may not be feasible electrically or operationally.

The next example demonstrates the use of de Morgan's theorem.

EXAMPLE 13.5:
Simplify

$$f = D+\overline{A}+\overline{B}+\overline{C}+A.C$$

After rearrangement of the terms, de Morgan's theorem can be used to convert $\overline{A}+\overline{C}$ into $\overline{A.C}$ [Eqn.(13.25)], i.e.

$$f = D+\overline{B}+A.C+(\overline{A}+\overline{C})$$

$$= D+\overline{B}+A.C+\overline{A.C}$$

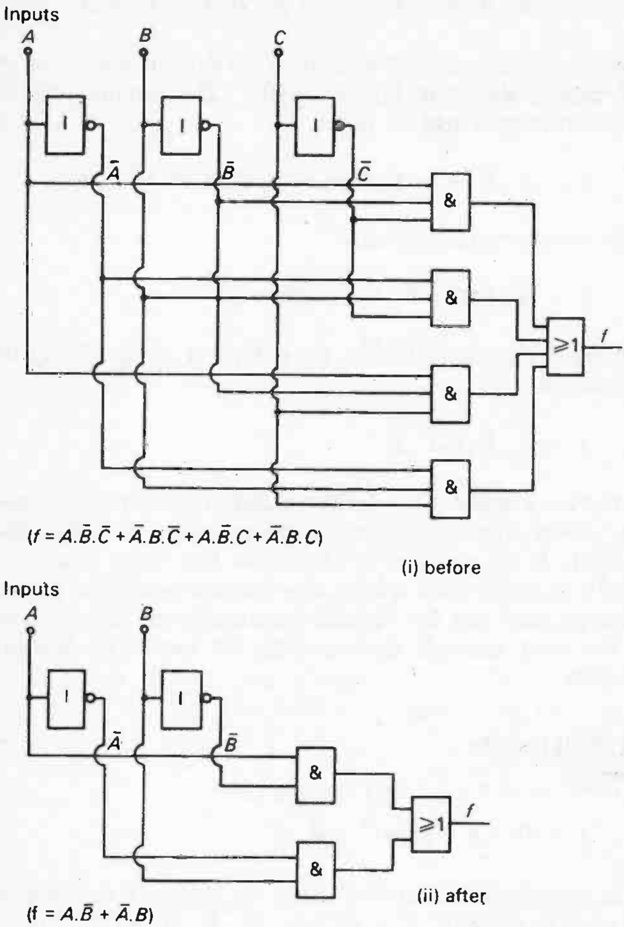


Fig. 13.4 An example of minimization

Next from Equation (13.21)

$$A \cdot C + \overline{A \cdot C} = 1$$

$$\therefore \quad f = \overline{B} + D + 1$$

and from Equation (13.18)

$$D + 1 = 1$$

$$\therefore \quad f = \overline{B} + 1$$

and again from Equation (13.18),

$$\overline{B} + 1 = 1$$

(actually anything + 1 = 1),

$$\therefore \quad f = 1$$

So whatever the logical values of A, B, C and D, f is always equal to 1, hence the system can serve no practical purpose.

These are examples of manipulating logic arrangements using boolean algebra and its known laws and theorems. For further practice logic equations can be made up and then minimized, safe in the knowledge that the answer can always be checked by drawing up the truth table.

13.4 SUMMARY OF KEY FORMULAE

A, B, C etc. = gate inputs

f = gate output

n = number of digits

r = radix

QUANTITY	FORMULA	SECTION
Expression of any number	$N = a_{n-1}r^{n-1} + a_{n-2}r^{n-2} + \ldots a_1 r^1 + a_0 r^0$	13.1
Associative law	$A \, . \, B \, . \, C = (A \, . \, B) \, . \, C = A \, . \, (B \, . \, C)$	13.3.1
	$A + B + C = (A + B) + C = A + (B + C)$	13.3.1
Commutative law	$A \, . \, B = B \, . \, A \qquad A + B = B + A$	13.3.1
Distributive law	$A \, . \, (B + C) = A \, . \, B + A \, . \, C$	13.3.1
	$A + (B \, . \, C) = (A + B) \, . \, (A + C)$	13.3.1
Theorems: (AND function)	$A \, . \, 0 = 0 \quad A \, . \, 1 = A \quad A \, . \, A = A$	13.3.2
(OR function)	$A + 0 = A \quad A + 1 = 1 \quad A + A = A$	13.3.2
(NOT function)	$A \, . \, \bar{A} = 0 \quad A + \bar{A} = 1 \quad \bar{\bar{A}} = A$	13.3.2
(de Morgan's)	$\overline{A \, . \, B} = \bar{A} + \bar{B}, \quad \overline{A + B} = \bar{A} \, . \, \bar{B}$	13.3.2.1

14. POWER SUPPLIES

Power for electronic equipment is almost inevitably in d.c. form and is available from primary batteries, especially for small and portable equipments, from rectification of the a.c. mains or from solar cells. It is mainly with rectification that we get involved in calculations.

14.1 THE SEMICONDUCTOR DIODE

The unidirectional qualities of a semiconductor diode leave the device with little competition for a.c. rectification. It is not perfect though for it has some resistance in the forward direction and less than infinite resistance in reverse. A typical forward characteristic (i.e. in the conducting direction) is given in Figure 14.1(i) and in (ii) the basic principle of rectification is illustrated.

14.1.1 Half-Wave Rectification

The single diode of Figure 14.1(ii) is shown in (iii) connected to a mains transformer with a turns ratio such as to give a voltage V_S capable of producing the required load voltage V_L. The output current is unidirectional but not *smooth*, such a pulsating current is far removed from what is normally required. Improvement may be gained by adding a filter, this is discussed in Section 14.2.

From Fourier analysis of the output wave (Book 1, Sect. 5.2.1) and for convenience, assuming a perfect diode:

$$v = \frac{2V_{max}}{\pi}\left(0.5 + \frac{\pi}{4}\sin\omega t - \frac{\cos 2\omega t}{3} - \frac{\cos 4\omega t}{15} - \frac{\cos 6\omega t}{35}\right)\ldots \qquad (14.1)$$

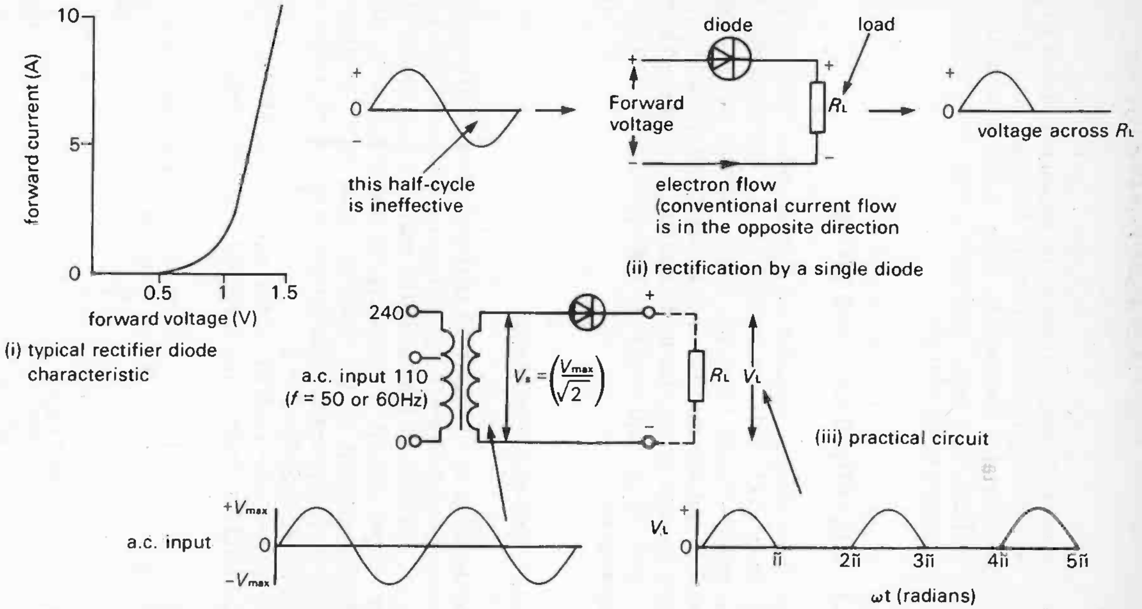


Fig. 14.1 The semiconductor rectifier

where V_{max} is the maximum applied voltage and $\omega = 2\pi f$.

Examining the terms in Equation (14.1) shows that the only component independent of frequency is the first, i.e.

$$\frac{2V_{max}}{\pi} \times 0.5 .$$

This we label V_L (the load voltage), hence:

$$V_L = \frac{V_{max}}{\pi} = 0.318\, V_{max} \qquad (14.2)$$

and if the r.m.s. supply voltage is V_s, then $V_{max} = \sqrt{2}\, V_s$

$$\therefore \quad V_L = \frac{\sqrt{2}\, V_s}{\pi} = 0.45\, V_s \qquad (14.3)$$

The remaining terms of Equation (14.1) comprise a component at the supply frequency, f and even harmonics theoretically to infinity but rapidly reducing in magnitude as the order (n) increases.

The efficiency, defined as

$$\frac{\text{output power}}{\text{input power}} \times 100\%$$

is calculated as follows:

From Equation (14.3), output power is equal to

$$\frac{V_L{}^2}{R_L} = \left(\frac{\sqrt{2}\, V_s}{\pi}\right)^2 \times \frac{1}{R_L}$$

and since power is taken from the supply for only half the time

$$\text{input power} = \frac{V_s^2}{2R_L}$$

$$\therefore \quad \text{efficiency} = \frac{\text{output power}}{\text{input power}} \times 100\%$$

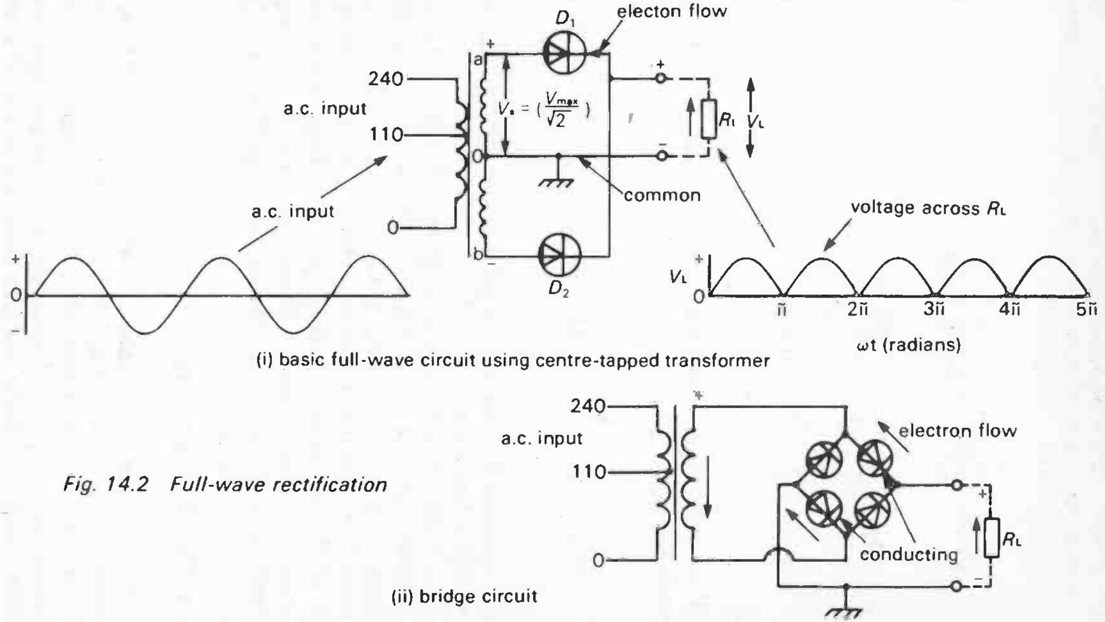
$$= \frac{2V_s^2}{\pi^2} \times \frac{1}{R_L} \times \frac{2R_L}{V_s^2} \times 100$$

$$= \frac{400}{\pi^2} = 40.5\% \tag{14.4}$$

14.1.2 Full-Wave Rectification

The main disadvantage of the half-wave rectifier is in the amount of filtering or smoothing required because of the relatively long periods between the current pulses. During the times of no current the filter circuit itself must take over, seemingly a job for a large capacitor but this may have disadvantages. The full-wave rectifier does not dispense with alternate half-cycles of the supply but rectifies and inverts them as shown in Figure 14.2(i). It requires two diodes and a centre-tapped secondary winding on the mains transformer. On the half-cycles when point a is positive to the common line, $D1$ conducts. Point b is at negative potential so $D2$ does not conduct. Current flows through the load as shown. On the opposite half-cycle of input $D2$ conducts whereas $D1$ does not and the current flow through R_L is in the same direction as before. From Fourier analysis:

$$v = \frac{2V_{max}}{\pi} - \frac{4V_{max}}{\pi}\left\{\frac{\cos 2\omega t}{3} + \frac{\cos 4\omega t}{15} + \frac{\cos 6\omega t}{35} + \ldots\right\} \tag{14.5}$$



(i) basic full-wave circuit using centre-tapped transformer

(ii) bridge circuit

Fig. 14.2 Full-wave rectification

giving a d.c. component $2V_{max}/\pi$, double that for the half-wave circuit. There is no component present at mains frequency, only second and higher harmonics, hence filtering is less complicated. Then

$$V_L = \frac{2V_{max}}{\pi} = 0.637\ V_{max} \tag{14.6}$$

and substituting $V_{max} = \sqrt{2}\,V_s$

$$V_L = \frac{2\sqrt{2}\,V_s}{\pi} = 0.9\ V_s \tag{14.7}$$

and the efficiency is obviously twice that for the half-wave circuit, i.e.

$$\text{efficiency} = \frac{800}{\pi^2} = 81\% \tag{14.8}$$

being appreciably less than 100% because of the generation of unusable harmonics.

The bridge rectifier circuit as shown in Figure 14.2(ii) is a full-wave method which needs no centre-tap on the transformer winding but requires 4 diodes instead of 2. Electron current flow is shown for the input half-cycle which makes the top of the secondary winding positive. When the polarity reverses the opposite pair of diodes conducts but resulting in current flow through R_L in the same direction as before.

14.2 FILTER CIRCUITS

A relatively large value capacitor connected across the load is inexpensive (an electrolytic is suitable) and moderately effective in reducing the *ripple* on the load current. The filter performance is improved markedly if the single capacitor is backed up by a simple low-pass filter consisting of a series inductor followed by a second capacitor as shown in Figure

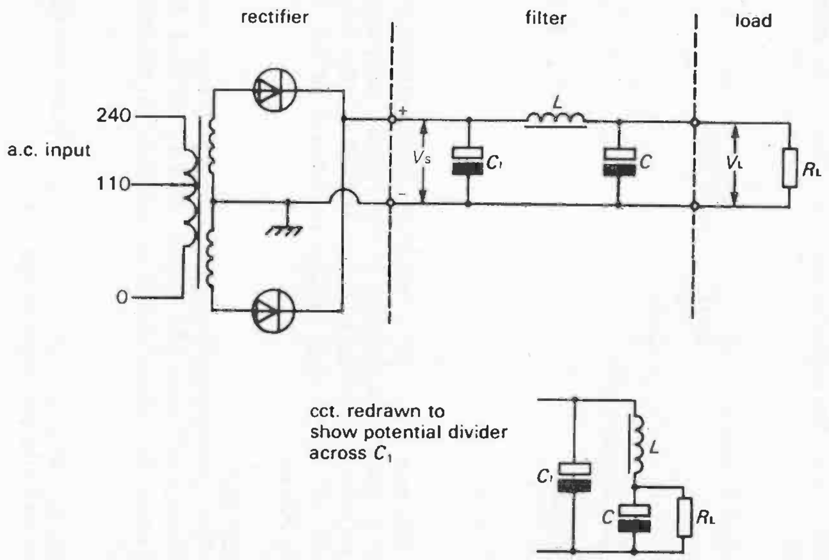


Fig. 14.3 Rectifying and filtering system

14.3 using the full-wave circuit as an example. The effect of L and C may be estimated by considering them as a potential divider connected across C_1, the output across C being fed to R_L. If we accept that by design the reactance of L must be many times that of C then we can simplify the calculation of the degree of reduction in ripple to X_C/X_L where X_C is the reactance of the capacitor and X_L that of the inductance, hence

$$\text{reduction in ripple} = \frac{\dfrac{1}{2\pi f C}}{2\pi f L} = \frac{1}{4\pi^2 f^2 LC}$$

(C in farads, L in henries)

$$= \frac{25330}{f^2 LC} \qquad (14.9)$$

(C in μF, L in H)

f in this case is the *ripple* frequency, i.e. the lowest (and therefore of greatest magnitude) of the Fourier components other than the d.c. For the half-wave rectifier it is the fundamental, 50 or 60 Hz but for the full-wave it is the 2nd harmonic, 100 Hz.

EXAMPLE 14.1:
What is the approximate reduction in ripple voltage for a filter circuit as in Figure 14.3 where $C = 47$ μF and $L =$ 20 H (i) when the rectifier circuit is half-wave, (ii) when it is full-wave?

(i) for the half-wave source, $f = 50$ Hz. Therefore from Equation (14.9) the reduction in ripple is

$$\frac{V_L}{V_s} = \frac{25330}{f^2 LC} = \frac{25330}{50^2 \times 20 \times 47} = 0.0108 \quad (\tfrac{1}{93})$$

(ii) for the full-wave source, $f = 100$ Hz. Therefore the reduction in ripple is

$$\frac{25330}{100^2 \times 20 \times 47} = 0.0027 \quad (\tfrac{1}{371})$$

so demonstrating the advantage of full over half-wave rectification, i.e. that ripple is reduced 4 times as much for the same filter circuit.

EXAMPLE 14.2:
In Example 14.1 a 20 H inductor is used. What value would be required for a full-wave circuit for the same ripple reduction as given by the half-wave circuit?

From Equation (14.9)

$$0.0108 = \frac{25330}{f^2 LC}$$

$$\therefore \quad L = \frac{25330}{0.0108 \times f^2 \times C}\ \text{H} = \frac{25330}{0.0108 \times 100^2 \times 47}$$

$$\therefore \quad L = 5\ \text{H}$$

(perhaps obvious because $L \propto 1/f^2$).

14.3 VOLTAGE STABILIZATION

Apart from the reduction of ripple in the load voltage, there is the problem of maintaining its mean value steady. Variation arises from:

(i) mains voltage changes

(ii) load current changes which create voltage drop variations across diode and transformer windings.

Both effects can be largely overcome by use of *stabilizer* circuits.

14.3.1 Zener Diode Stabilization

Whereas reverse breakdown is to be avoided with rectifying diodes, in the *zener diode* (after C. M. Zener, an American physicist), this is the desirable feature. A typical characteristic is given in Figure 14.4(i) and although the forward direction is included for completeness, it is not used. In the breakdown region relatively large changes in current can occur with little change in reverse voltage. That shown in the Figure is 7.5 V but devices are manufactured over a range of values up to some 200 V or more. The device therefore has the property of voltage stabilization or regulation as in the simple circuit of Figure 14.4(ii). V_s is the source voltage as supplied by the rectifier system, V_L the load voltage. The series resistance R reduces V_s to V_L since $V_L = V_s - IR$. I is made up of I_Z, the current through the diode and I_L, that in the load hence:

$$V_s = V_L + IR = V_L + (I_Z + I_L)R$$

$$\therefore \quad V_s = I_L R_L + (I_Z + I_L)R \qquad (14.10)$$

Thus should V_L try to increase, I_Z increases and creates a greater voltage drop across R to compensate and vice versa. Also should V_s increase, again the diode takes a sufficiently greater current to maintain V_L and vice versa. From Equation (14.10):

$$V_s - I_L R_L = (I_Z + I_L)R$$

$$\therefore \quad R = \frac{V_s - I_L R_L}{I_Z + I_L} = \frac{V_s - V_L}{I_Z + I_L} \qquad (14.11)$$

which assumes a constant value of V_s. Note that when I_L is maximum I_Z must be minimum so that lower values

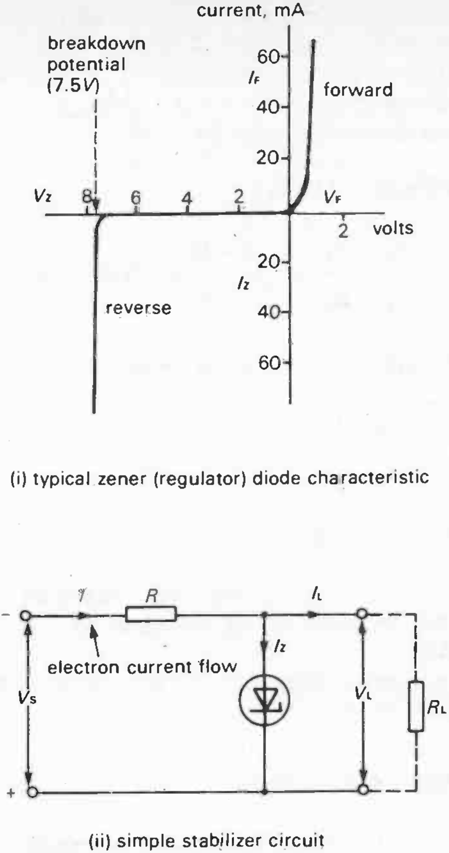


Fig. 14.4 Zener diode stabilization

of I_L can be compensated by increases in I_Z. I_Z can be calculated from Equation (14.10) giving:

$$I_Z = \frac{V_s - V_L}{R} - I_L \qquad (14.12)$$

EXAMPLE 14.3:
A stabilization circuit as in Figure 14.4(ii) requires V_L to remain as near as possible to 6 V over a load current range 6 – 18 mA. The supply voltage V_S is 14 V and the zener diode minimum current is 2 mA. Calculate the value of R and the maximum dissipation in the diode.

From Equation (14.11):

$$R = \frac{V_S - V_L}{I_Z + I_L} = \frac{14 - 6}{2 + 18} \times 1000 = 400\ \Omega$$

I_Z is maximum when I_L is minimum so from Equation (14.12):

$$I_Z = \frac{V_S - V_L}{R} - I_L = \frac{14 - 6}{400} - 0.006$$

$$= 0.014\ \text{A} = 14\ \text{mA}$$

(i.e. when I_L falls 12 mA from maximum to minimum, this current is taken up by the diode so I_Z rises from 2 to 14 mA),

then maximum dissipation in zener diode equals $V_L \times I_Z$ $= 6 \times 0.014$ W $=$ 84 mW.

14.3.2 Series Stabilization

The zener diode method above relies on creation of a voltage drop across a *fixed* series resistance. A more efficient system changes the value of the resistance instead. The basic idea of the method is illustrated in Figure 14.5(i) in which the voltage V_L and a reference voltage V_r are both fed to a *comparator* which produces an *error signal* according to the difference between them. The error signal causes the control element to reduce or increase the load current accordingly so that V_L remains more or less constant. The reference voltage in its simplest form is a zener (reference) diode making use of the fact that its voltage remains relatively constant irrespective of

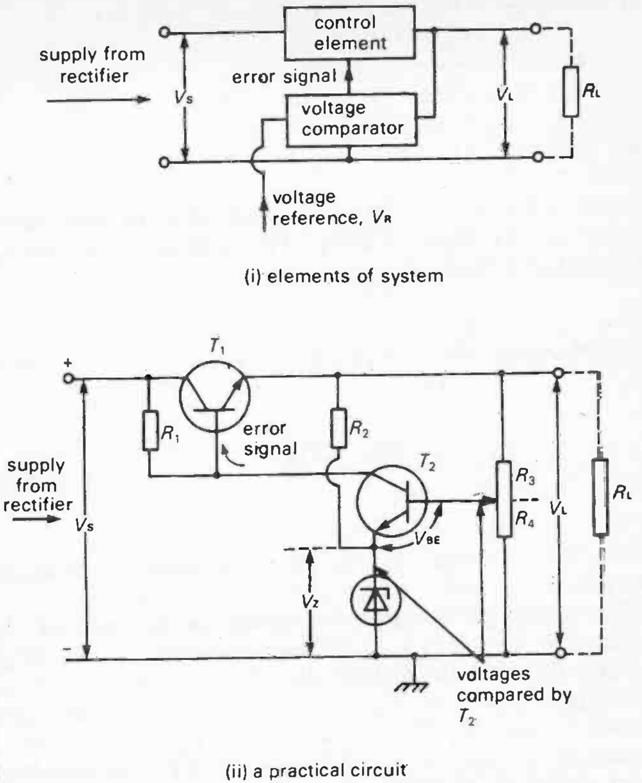


Fig. 14.5 Series stabilization

the current through it [Fig.14.4(i)]. The main elements of a simple series stabilizer are in (ii) of the Figure which shows that the two voltages are compared by T_2, the collector current of which is dependent on their difference. This is applied to T_1 to control its main current in order to bring the difference between the compared voltages to minimum.

The voltages being compared are:

(i) $V_L \times \dfrac{R_4}{R_3 + R_4}$, and

(ii) $(V_Z + V_{BE})$

i.e. the zener diode reference voltage plus the base-emitter voltage of T_2 (about 0.6 V). For stabilization (i) and (ii) are equal, therefore:

$$V_L \times \frac{R_4}{R_3 + R_4} = (V_Z + V_{BE})$$

$$\therefore \quad V_L = \frac{(R_3 + R_4)(V_Z + V_{BE})}{R_4} \qquad (14.13)$$

from which, knowing V_L the appropriate zener diode can be selected.

The *degree* of stabilization is given by the ratio of the variation in V_L (δV_L) to the original variation in V_S (δV_S) i.e. $\delta V_L / \delta V_S$. If the voltage gain of T_2 is A_{v2}, then the error signal applied to T_1 is $\delta V_L \times A_{v2}$. The net base-emitter voltage of T_1 is therefore $(\delta V_L \times A_{v2} - \delta V_L)$ so producing a voltage $A_{v1}(\delta V_L \times A_{v2} - \delta V_L)$ between base and collector terminals of T_1 which is assumed to have a voltage gain A_{v1}. The voltage across these terminals is also approximately $(\delta V_S - \delta V_L)$ hence:

$$\delta V_S - \delta V_L = A_{v1}(\delta V_L \times A_{v2} - \delta V_L)$$

$$= \delta V_L(A_{v1} \times A_{v2} - A_{v1})$$

$$\therefore \quad \delta V_S = \delta V_L(A_{v1} \times A_{v2} - A_{v1} + 1)$$

$$\therefore \quad \frac{\delta V_L}{\delta V_S} = \frac{1}{A_{v1}(A_{v2} - 1) + 1} \qquad (14.14)$$

EXAMPLE 14.4:

In a circuit such as Figure 14.5(ii), T_2 has a voltage gain of −110 and T_1 a voltage gain of −9. What is the degree of stabilization? Also, if the supply has an output resistance of 150 Ω and the load current increases by 50 mA, by how much do the input and output voltages of the circuit change?

$$A_{v1} = -9 \qquad A_{v2} = -110$$

From Equation (14.14):

$$\frac{\delta V_L}{\delta V_S} = \frac{1}{A_{v1}(A_{v2} - 1) + 1} = \frac{1}{-9(-111) + 1}$$

$$= \frac{1}{1000} = 0.001$$

showing that the circuit has a 1000:1 voltage variation reduction. Voltage drop in supply for 50 mA load current = $50 \times 10^{-3} \times 150$ = 7.5 V so when δV_S = 7.5 V , since $\delta V_L/\delta V_S$ = 0.001, δV_L = 0.0075 V = 7.5 mV, i.e. when the load current increases by 50 mA, the input voltage falls by 7.5 V but the load voltage falls by only 7.5 mV.

A practical difficulty with series stabilization is that the full load current flows through the series element thus transistor T_1 must have a sufficiently high power rating. More sophisticated circuits are therefore available, for example, shunt stabilizers and the inevitable IC specially designed for the purpose.

14.4 SUMMARY OF KEY FORMULAE

A_{v1} , A_{v2} = transistor voltage gains

C = capacitance (F)

I_L = load current

I_Z = zener diode current

L = inductance (H)

n = series of even integers to ∞ (i.e. 2, 4, 6, . . .)

R_L = load resistance

V_{BE} = base-emitter d.c. voltage

V_S = r.m.s. supply voltage to rectifying system

QUANTITY	FORMULA	UNIT	SECTION
Half-wave rectification:			
Fourier analysis of waveform	$v = \frac{2V_{max}}{\pi}\left(0.5 + \frac{\pi}{4}\sin \omega t - \left\{\frac{\cos n\omega t}{(n+1)(n-1)}\right\}\right)$	V	14.1.1
load voltage	$V_L = \frac{\sqrt{2}\,V_S}{\pi} = 0.45\ V_S$	V	14.1.1
efficiency	$\eta = \frac{400}{\pi^2} = 40.5$	%	14.1.1
Full-wave rectification:			
Fourier analysis of waveform	$v = \frac{2V_{max}}{\pi} - \frac{4V_{max}}{\pi} \times \left\{\frac{\cos n\omega t}{(n+1)(n-1)}\right\}$	V	14.1.2
load voltage	$V_L = \frac{2\sqrt{2}\,V_S}{\pi} = 0.9\ V_S$	%	14.1.2

QUANTITY	FORMULA	UNIT	SECTION
Full-wave rectification (contd.):			
efficiency	$\eta = \frac{800}{\pi^2} = 81$	%	14.1.2
LC filter – reduction in ripple	$\approx \frac{1}{4\pi^2 f^2 LC}$		14.2
Zener diode stabilization (Fig.14.4):			
supply voltage	$V_s = I_L R_L + (I_Z + I_L) R$	V	14.3.1
series resistance	$R = \frac{V_S - V_L}{I_Z + I_L}$	Ω	14.3.1
diode current	$I_Z = \frac{V_S - V_L}{R} - I_L$	A	14.3.1

QUANTITY	FORMULA	UNIT	SECTION
Series stabilization (Fig.14.5):			
load voltage	$V_L = \frac{(R_3 + R_4)(V_Z + V_{BE})}{R_4}$	V	14.3.2
stabilization	$\frac{\delta V_L}{\delta V_S} = \frac{1}{A_{v1}(A_{v2} - 1) + 1}$		14.3.2

Appendix 1

CONVERSION FACTORS

Although standardization of units is proceeding apace in the form of the rationalized system of metric units (S.I.), there are many instances when earlier units or those for specialist subjects appear in calculations. The following Table therefore contains conversion factors as an aid. The list is not exhaustive but should cater for a large proportion of present day needs.

Some abbreviations are used where convenient:

/ = per	cm = centimetre(s)
cu = cubic	ft = foot
hrs = hours	kg = kilogramme(s)
lb = pound(s)	log = logarithm
m = metre(s)	sq. = square

To convert	into	multiply by
ampere-hours	coulombs	3600
amperes/metre	oersteds	1.257×10^{-2}
angstroms	metres	10^{-10}
angstroms	nanometers	0.1
atmospheres	newtons/sq. m.	1.013×10^{5}
bars	newtons/sq. m.	10^{5}
bits (reliability)	fits	10

To convert	into	multiply by:
bits (reliability)	failures/1000 hrs	10^{-5}
bits (reliability)	% failures/1000 hrs	10^{-3}
BTU	joules	1.055×10^{3}
BTU	kilowatt-hours	2.93×10^{-4}
calories	joules	4.187
Celsius, °C (Centigrade)	Fahrenheit, °F	(°C x 9/5) + 32
centimetres	inches	3.937×10^{-1}
centimetres	feet	3.281×10^{-2}
coulombs	ampere-hours	2.778×10^{-4}
cubic cm	cubic inches	6.102×10^{-2}
cubic feet	litres	28.32
cubic inches	cubic cm	16.39
cubic inches	litres	1.639×10^{-2}
decibels	nepers	0.1151
degrees	radians	1.745×10^{-2}
dynes	grammes	1.020×10^{-3}
dynes	newtons	10^{-5}
dynes	poundals	7.233×10^{-5}
dynes	pounds	2.248×10^{-6}

To convert	into	multiply by:
dynes/sq. cm	kg/sq. m.	1.020×10^{-2}
dynes/sq. cm	newtons/sq. m.	10^{-1}
dynes/sq. cm	pounds/sq. inch	1.45×10^{-5}
electron-volts	joules	1.602×10^{-19}
ergs/second	watts	10^{-7}
Fahrenheit, °F	Celsius, °C (Centigrade)	(°F – 32) x 5/9
failures/1000 hrs	fits	10^{6}
failures/1000 hrs	bits	10^{5}
feet	centimetres	30.48
feet	kilometres	3.048×10^{-4}
fits	bits	10^{-1}
fits	failures/1000 hrs	10^{-6}
fits	% failures/1000 hrs	10^{-4}
foot-pounds	kilowatt-hrs	3.766×10^{-7}
foot-pounds/sec	horsepower	1.818×10^{-3}
gallons (U.K.)	litres	4.546
gallons (U.S.)	litres	3.785
gauss	lines/sq. inch	6.452
gauss	teslas	10^{-4}
gauss	webers/sq. m.	10^{-4}

To convert	into	multiply by:
grammes	dynes	980.7
grammes/cu cm	kilogrammes/cu m	10^3
horsepower	ft-lbs/sec	550
horsepower	watts	745.7
inches	centimetres	2.54
inches	mils	1000
joules	BTU	9.478×10^{-4}
joules	calories	0.239
joules	electron-volts	6.24×10^{18}
joules	kilowatt-hours	2.778×10^{-7}
kilogrammes	dynes	9.807×10^5
kilogrammes	newtons	9.807
kilogrammes	pounds	2.205
kilogrammes	tonnes	10^{-3}
kg/sq. m	dynes/sq. cm	98.07
kg/sq. m	newtons/sq. m	9.807
kg/sq. m	pounds/sq. inch	1.422×10^{-3}
kg/cu m	grammes/cu cm	10^{-3}
kilometres	feet	3281
kilowatt-hours	BTU	3413

To convert	into	multiply by:
kilowatt-hours	foot-pounds	2.655×10^6
kilowatt-hours	joules	3.6×10^6
lines/sq. inch	gauss	0.155
litres	cubic feet	3.53×10^{-2}
litres	cubic inches	61.02
litres	gallons (U.K.)	0.22
litres	gallons (U.S.)	0.2642
$\log_e$	$\log_{10}$	0.4343
$\log_{10}$	$\log_e$	2.3026
maxwells	webers	10^{-8}
metres	angstroms	10^{10}
metres	microns	10^6
metres	yards	1.094
microhms/cu cm	microhms/inch cube	0.3937
microhms/inch^3	microhms/cu cm	2.54
microns	metres	10^{-6}
mils	inches	10^{-3}
nanometres	angstroms	10
nepers	decibels	8.686

To convert	into	multiply by:
newtons	dynes	10^5
newtons	kilogrammes	0.102
newtons	poundals	7.233
newtons	pounds	0.2248
newtons/sq. m	atmospheres	9.869×10^{-6}
newtons/sq. m	bars	10^{-5}
newtons/sq. m	dynes/sq. cm	10
newtons/sq. m	kilogrammes/sq. m	0.102
newtons/sq. m	pounds/sq. inch	1.45×10^{-4}
oersteds	amperes/metre	79.58
pascals	pounds/sq. inch	1.45×10^{-4}
poundals	dynes	1.383×10^4
poundals	newtons	0.1383
pounds	dynes	4.448×10^5
pounds	kilogrammes	0.4536
pounds	newtons	4.448
pounds/sq. inch	dynes/sq. cm	6.895×10^4
pounds/sq. inch	newtons/sq. m (pascals)	6.895×10^3
pounds/sq. inch	kilogrammes/sq. m	7.031×10^2

To convert	into	multiply by:
radians	degrees	57.3
square inches	square cm	6.452
square cms	square inches	0.155
teslas	gauss	10^4
tonnes	kilogrammes	10^3
tonnes	tons	0.9842
tons	tonnes	1.016
webers	maxwells	10^8
webers/sq. m	gauss	10^4
watts	ergs/sec	10^7
watts	horsepower	1.341×10^{-3}
yards	metres	0.9144

Appendix 2

MATHEMATICAL FORMULAE

This Appendix is in essence a compendium of the mathematical formulae likely to be encountered in electronics calculations up to the level employed in the main text, i.e. calculus, and other more advanced tactics are excluded. It is assumed that the reader has some acquaintance with the formulae and only needs reminders. An exception is to be found in Section A2.1.4.1 which discusses the *power series* in some depth.

A2.1 ALGEBRA

In what follows + . . . or − . . . means "and terms in like manner to infinity".

! stands for the *factorial*, $n!$ (factorial n) being the product $1 \times 2 \times 3 \times 4 \times \ldots \times n$.

A2.1.1 Solution of Equations

One unknown:

If $ax + b = c$ then $x = \dfrac{c - b}{a}$ (A2.1)

Two unknowns:

Two equations are required, e.g. $a_1x + b_1y = c_1$ and $a_2x + b_2y = c_2$. They can be solved as simultaneous equations or directly from formulae:

$$x = \frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1} \qquad y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1} \qquad \text{(A2.2)}$$

Quadratic (involves squares but no higher order): If

$$ax^2 + bx + c = 0,$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad \text{(A2.3)}$$

when $b^2 < 4ac$ the roots are complex (i.e. they have real and imaginary components, see Chapter 4).

$$(a \pm b)^2 = a^2 \pm 2ab + b^2 \qquad \text{(A2.4)}$$

$$a^2 - b^2 = (a+b)(a-b) \qquad \text{(A2.5)}$$

A2.1.2 Exponents

An exponent is the raised number indicating what *power* of a factor is to be taken. The term *index* is also used.

$$a^0 = 1 \qquad a^n = a \times a \times a \ldots n \text{ times}$$

$$a^{-n} = \frac{1}{a^n} \qquad \text{(A2.6)}$$

$$a^n \times a^m = a^{n+m} \qquad \text{(A2.7)}$$

$$a^n \times a^{-m} = \frac{a^n}{a^m} = a^{n-m} \qquad \text{(A2.8)}$$

$$a^n \times b^n = (ab)^n \qquad \text{(A2.9)}$$

$$\frac{a^n}{b^n} = \left(\frac{a}{b}\right)^n \qquad \text{(A2.10)}$$

$$(a^n)^m = a^{nm} \qquad \text{(A2.11)}$$

A2.1.3 Logarithms

There are several forms of logarithms each having a different *base* which must be quoted. The logarithm of any number is the power to which the base must be raised to produce that number, thus if the base is a and $a^m = n$ then the logarithm to the base a of the number n is m, written as

$$m = \log_a n\ .$$

The base of *common* logarithms is 10, written log or $\log_{10}$ and of *natural* logarithms, e, written ℓn or $\log_e$.

$$\log_a (m \times n) = \log_a m + \log_a n \qquad \text{(A2.12)}$$

$$\log_a \frac{m}{n} = \log_a m - \log_a n \qquad \text{(A2.13)}$$

$$\log_a m^n = n \log_a m \qquad \text{(A2.14)}$$

$$\log_a \sqrt[n]{m} = \log_a m^{1/n} = \frac{1}{n} \log_a m \qquad \text{(A2.15)}$$

Change of base:

$$\log_b m = \log_a m \times \log_b a \qquad \text{(A2.16)}$$

and substituting 10 for a and e (= 2.71828) for b

$$\log_e m = \log_{10} m \times \log_e 10 = 2.3026 \times \log_{10} m \qquad \text{(A2.17)}$$

and conversely:

$$\log_{10} m = \log_e m \times \log_{10} e = 0.4343 \log_e m \qquad \text{(A2.18)}$$

A2.1.4 Series

Arithmetical progression

$$a + (a + d) + (a + 2d) + \ldots$$

$$n\text{th term} \quad \ell = a + (n-1)d \tag{A2.19}$$

Sum of n terms:

$$\frac{n}{2}(a + \ell) = \frac{n}{2}\left\{2a + (n-1)d\right\} \tag{A2.20}$$

Geometrical progression

$$a,\ ar,\ ar^2 + \ldots$$

$$n\text{th term} = ar^{n-1} \tag{A2.21}$$

Sum of n terms:

$$\frac{a(1 - r^n)}{1 - r} \tag{A2.22}$$

Binomial Series

$$(a + b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \frac{n(n-1)(n-2)}{3!}a^{n-3}b^3 + \ldots \tag{A2.23}$$

Exponential Series

$$e^{\pm x} = 1 \pm x + \frac{x^2}{2!} \pm \frac{x^3}{3!} + \ldots \qquad \text{(A2.24)}$$

$$(= 2.71828 \text{ when } x = 1)$$

$$e^{\pm kx} = 1 \pm kx + \frac{k^2x^2}{2!} \pm \frac{k^3x^3}{3!} + \ldots \qquad \text{(A2.25)}$$

Trigonometrical Series

(x in radians)

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots \qquad \text{(A2.26)}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots \qquad \text{(A2.27)}$$

A2.1.4.1 The Power Series

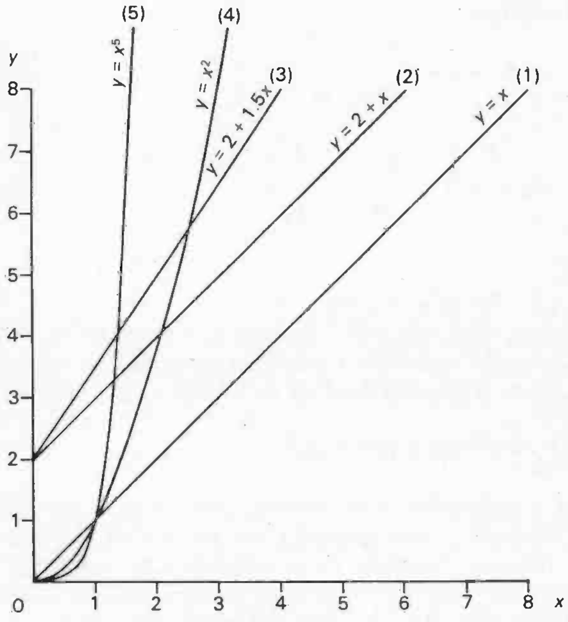
In dealing with distortion through non-linearity, the suggestion is usually made that a particular curve may be represented by a *power series*, the formula of which has the general form:

$$y = a + bx + cx^2 + dx^3 + \ldots$$

It is a complicated, never ending expression so nearly all the later terms must be neglected, but where does one draw the line? It may therefore be worthwhile to experiment with the formula to see not only what it can do for us but what its limitations are. To construct the graph the values of a, b, c etc. must be known, but firstly we can see what their effect is from the simple curves given in Figure A2.1(i). When a, c, d . . . are all equal to 0 and $b = 1$, the simplest relationship of all arises, $y = x$, a straight line passing through the origin

(1). For any other value of a, the line emanates from the y axis at that value and at (2) this is shown for $a = 2$, $b = 1$, $c, d \ldots = 0$. Thus a merely shifts the graph up or down, b alters the slope as shown at (3) for $a = 2$, $b = 1.5$, c etc. $= 0$. This is normally expressed in the form $y = mx + c$. Thus juggling with a and b only leaves us always with a straight line.

When a power of x is introduced the graph is no longer straight as (4) and (5) show for $y = x^2$ and x^5. Now the plot thickens for it is by using combinations of these and others that any smooth curve can be represented. They are

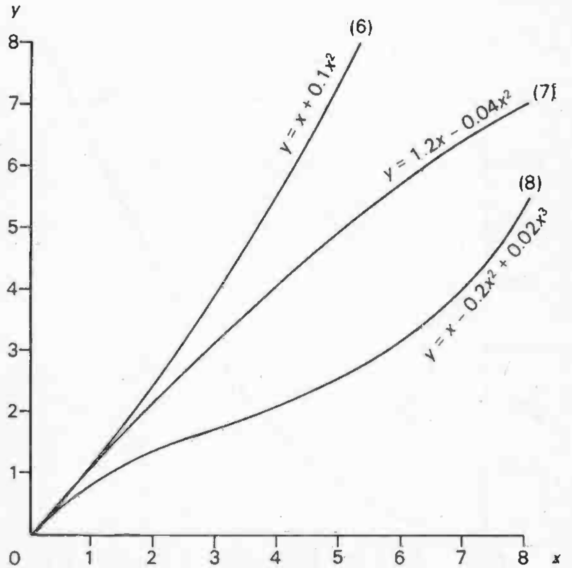


(i) straight lines and simple curves

Fig. A2.1 The power series

the ingredients from which a final mix is made, the amount of each being stated by the values of a, b, c etc. Taking a simple example of $a = 0$, $b = 1$, $c = 0.1$, d, etc. = 0 results in $y = x + 0.1x^2$ and this is plotted in Figure A2.1(ii) at (6), a rather gentle curve because only a little of x^2 has been added. This curve is said to follow a *square law* because one component is proportional to the square of x. An example of a curve bending in the opposite direction is at (7) for $y = 1.2x - 0.04x^2$.

So far we have examined how the simple mixture of $a + bx + cx^2$ can produce an infinite variety of square-law



(ii) examples of graphs of power series

Fig. A2.1 (contd)

curves. Adding in one more component, that of x^3 but in a small amount, ($d = 0.02$) shows at (8) how the curve direction can be changed at will.

Many smooth curves can therefore be represented reasonably well by the power series up to only the 4th term. Addition of higher orders of x for fine adjustment only involves low values of the coefficients, hence the conclusion so often seized upon that these can be neglected, for if not, the mathematics do really get out of hand.

Knowing that transistor input curves (e.g. Fig.5.4) are generally square-law and can be expressed by the first three terms of the power series, we look at such a curve as in Figure

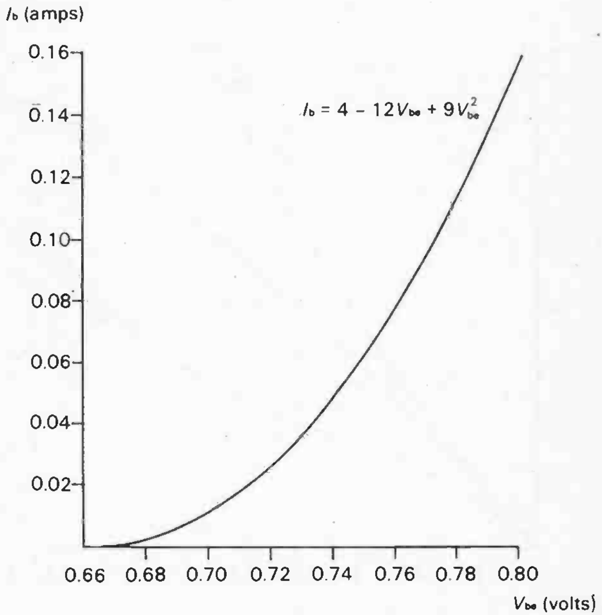


Fig. A2.2 Transistor input characteristic expressed as a power series

A2.2 and calculate its power series from three points only, e.g.

at $V_{be} = 0.7$ V, $I_b = 0.01$ A

at $V_{be} = 0.75$ V, $I_b = 0.0625$ A

at $V_{be} = 0.8$ V, $I_b = 0.160$ A

leading to three simultaneous equations:

$$0.01 = a + 0.7b + 0.7^2c$$

$$0.0625 = a + 0.75b + 0.75^2c$$

$$0.160 = a + 0.8b + 0.8^2c$$

the solution being

$$a = 4, \quad b = -12, \quad c = 9.$$

The curve can therefore be represented by the power series

$$I_b = a + bV_{be} + cV_{be}^2$$

where a, b and c have the values above.

A2.1.5 Equations of Curves

$$y = mx + c \qquad \text{(A2.28)}$$

(straight line, gradient m passing through $0, c$)

$$y = ax^2 + bx + c \qquad \text{(A2.29)}$$

(parabola, axis parallel to y axis)

$$x^2 + y^2 = a^2 \qquad \text{(A2.30)}$$

(circle, centre at origin, radius a)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad \text{(A2.31)}$$

(ellipse, semi-axes a and b).

A2.2 TRIGONOMETRY

In the right-angled triangle shown in Figure A2.3(i), θ is one of the two angles less than 90°. Labelling the sides of the triangle as shown:

$$\text{sine } \theta \ (\sin \theta) = \frac{a}{c},$$

$$\text{cosine } \theta \ (\cos \theta) = \frac{b}{c},$$

$$\text{tangent } \theta \ (\tan \theta) = \frac{a}{b},$$

$$\text{cosecant } \theta \ (\text{cosec } \theta) = \frac{1}{\sin \theta} = \frac{c}{a},$$

$$\text{secant } \theta \ (\sec \theta) = \frac{1}{\cos \theta} = \frac{c}{b},$$

$$\text{cotangent } \theta \ (\cot \theta) = \frac{1}{\tan \theta} = \frac{b}{a}.$$

From the above,

$$\sin^2 \theta = \frac{a^2}{c^2}, \qquad \cos^2 \theta = \frac{b^2}{c^2}$$

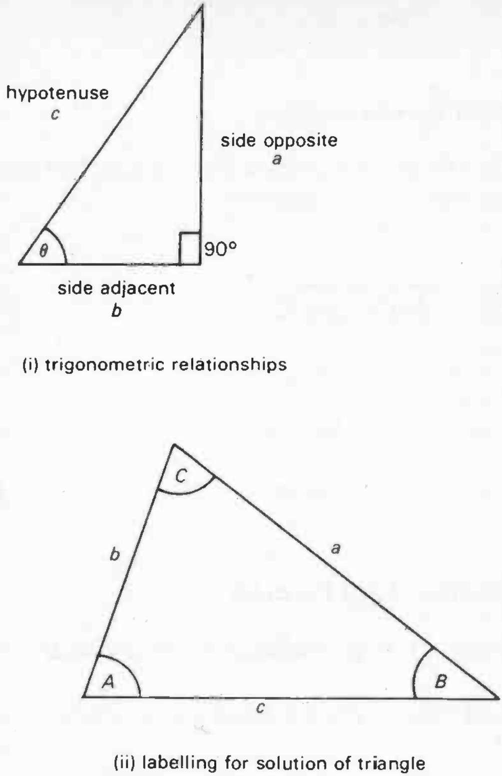


Fig. A2.3 Trigonometry

$$\therefore \quad \sin^2\theta + \cos^2\theta = \frac{a^2 + b^2}{c^2}$$

but $a^2 + b^2 = c^2$ (Pythagoras)

$$\therefore \quad \sin^2\theta + \cos^2\theta = 1 \qquad \text{(A2.32)}$$

$$\text{Also } \frac{\sin\theta}{\cos\theta} = \frac{a/c}{b/c} = \frac{a}{b} = \tan\theta \qquad \text{(A2.33)}$$

A2.2.1 Solution of Triangles

Figure A2.3(ii) shows a triangle (not necessarily right-angled) with angles and sides labelled; then

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} \qquad \text{(A2.34)}$$

$$a^2 = b^2 + c^2 - 2bc\cos A \qquad \text{(A2.35)}$$

$$b^2 = a^2 + c^2 - 2ac\cos B \qquad \text{(A2.36)}$$

$$c^2 = a^2 + b^2 - 2ab\cos C \qquad \text{(A2.37)}$$

A2.2.2 Multiple Angle Formulae

$$\sin(A \pm B) = \sin A \times \cos B \pm \cos A \times \sin B \qquad \text{(A2.38)}$$

$$\cos(A \pm B) = \cos A \times \cos B \mp \sin A \times \sin B \qquad \text{(A2.39)}$$

$$\tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \times \tan B} \qquad \text{(A2.40)}$$

$$\sin 2A = 2\sin A\cos A \qquad \text{(A2.41)}$$

$$\cos 2A = \cos^2 A - \sin^2 A = 2\cos^2 A - 1$$

$$= 1 - 2\sin^2 A \qquad \text{(A2.42)}$$

[from which $\sin^2 A = \frac{1}{2}(1 - \cos 2A)$ and $\cos^2 A = \frac{1}{2}(1 + \cos 2A)$]

$$\tan 2A = \frac{2 \tan A}{1 - \tan^2 A} \qquad \text{(A2.43)}$$

$$\sin A \cos B = \tfrac{1}{2}\{\sin(A+B) + \sin(A-B)\} \qquad \text{(A2.44)}$$

$$\cos A \sin B = \tfrac{1}{2}\{\sin(A+B) - \sin(A-B)\} \qquad \text{(A2.45)}$$

$$\sin A \sin B = \tfrac{1}{2}\{\cos(A-B) - \cos(A+B)\} \qquad \text{(A2.46)}$$

$$\cos A \cos B = \tfrac{1}{2}\{\cos(A+B) + \cos(A-B)\} \qquad \text{(A2.47)}$$

$$\sin A + \sin B = 2 \sin \frac{A-B}{2} \cos \frac{A+B}{2} \qquad \text{(A2.48)}$$

$$\sin A - \sin B = 2 \cos \frac{A+B}{2} \sin \frac{A-B}{2} \qquad \text{(A2.49)}$$

$$\cos A + \cos B = 2 \cos \frac{A+B}{2} \cos \frac{A-B}{2} \qquad \text{(A2.50)}$$

$$\cos A - \cos B = -2 \sin \frac{A+B}{2} \sin \frac{A-B}{2} \qquad \text{(A2.51)}$$

Appendix 3

HISTORICAL

A3.1 THE RENOWNED IN ENGINEERING (pre 1900)

AMPERE, Andre Marie	French mathematician and physicist	1775–1836
BAIRD, John Logie	Scottish inventor	1888–1946
BELL, Alexander Graham	Scottish-American inventor	1847–1922
BOLTZMANN, Ludwig	Austrian physicist	1844–1906
BOOLE, George	English schoolteacher and mathematician	1815–1864
COULOMB, Charles Augustin de	French engineer and physicist	1738–1806
DESCARTES, Renée	French mathematician and philosopher	1596–1650
de MORGAN, Augustus	English mathematician and logician	1806–1871
EINSTEIN, Albert	German-Swiss physicist	1879–1955
EDISON, Thomas Alva	American scientist and inventor	1847–1931
FARADAY, Michael	English chemist and physicist	1791–1867
FLEMING, John Ambrose	English electrical engineer and physicist	1849–1945

FOURIER, Jean-Baptiste Joseph,	French mathematician	1768–1830
FRANKLIN, Benjamin	American scientist and statesman	1706–1790
GALILEI, Galileo	Italian mathematician physicist and astronomer	1564–1642
GAUSS, Karl Friedrich	German mathematician	1777–1855
HENRY, Joseph	American physicist and scientific administrator	1797–1878
HERTZ, Henrich	German physicist	1857–1894
JOULE, James Prescott	English scientist	1818–1889
KELVIN, Lord William Thomson	Scottish physicist and and mathematician	1824–1907
KIRCHHOFF, Gustav Robert	German physicist	1824–1887
LAPLACE, Marquis de	French astronomer and scientist	1749–1827
LENZ, Heinrich Friedrich Emil	German-Russian physicist	1804–1865
MARCONI, Guglielmo	Italian electrical engineer	1874–1937
MAXWELL, James Clerk	Scottish physicist	1831–1879

MORSE, Samuel Finley Breese	American electrical engineer and artist	1791–1872
NAPIER, John	Scottish mathematician	1550–1617
NEWTON, Sir Isaac	English physicist and mathematician	1642–1727
OERSTED, Hans Christian	Danish physicist and chemist	1777–1851
OHM, Georg Simon	German experimental physicist	1787–1854
PASCAL, Blaise	French scientist and mathematician	1623–1662
PLANCK, Max	German theoretical physicist	1858–1947
RUTHERFORD, Lord, Ernest	New Zealand physicist	1871–1937
SIEMENS, Ernst Werner von	German electrical engineer	1816–1892
TESLA, Nikola	American (but Yugoslav born) electrical engineer	1856–1943
THOMSON, Sir Joseph John	English physicist	1856–1940
VOLTA, Alessandro	Italian physicist	1745–1827
WATT, James	Scottish engineer	1736–1819
WEBER, Wilhelm Eduard	German physicist	1804–1891

WHEATSTONE, Sir Charles	English physicist	1802–1875
WIEN, Wilhelm	German physicist	1864–1928

A3.2 EVENTS IN ENGINEERING

	1687	Newton published laws of motion and gravitation
	1799	Volta invented electric battery
	1800	
02		Young proved wave theory of light
07		Fourier proposed his Theorem
	1810	
19		Oersted discovered connection between current and magnetism
	1820	
22		Charles Babbage designed his "Difference Engine"
27		Ohm proposed his Law
	1830	
31		Faraday discovered electromagnetic induction
33		Gauss and Weber demonstrated electric telegraph
38		Morse produced his Code Cooke and Wheatstone developed first telegraph for public use
	1840	
43		Joule measured mechanical equivalent of heat
	1850	
	1860	
62		Maxwell proposed electromagnetic theory of light

1870	
76	Bell filed patent for telephone
77	Edison devised carbon microphone and phonograph
78	First use of selenium cell by Bell and Tainter. Swan produced incandescent electric lamp
1880	
86	Hertz demonstrated existence of radio waves
1890	
95	X-rays discovered by Röntgen
97	Thomson discovered electron
98	Poulsen developed magnetic recording
1900	Planck produced quantum theory of radiation
01	Marconi transmitted by radio across Atlantic
04	Fleming devised valve for reception of radio signals
05	Einstein developed Theory of Relativity
1910	Geiger invented counter
11	Rutherford proposed existence of atomic nuclei
1920	
23	Zworykin developed electronic scanning
26	Baird demonstrated television
1930	
32	Chadwick suggested existence of the neutron
1940	
47	Bardeen, Brattain and Shockley produced point-contact transistor

	1950	
51		Shockley patented junction transistor
58		First IC developed
	1960	Maiman demonstrated ruby laser

OTHER TITLES BY THE SAME AUTHOR

PRACTICAL ELECTRONICS CALCULATIONS AND FORMULAE **BP53**

F. A. Wilson, CGIA, CEng, FIEE, FIERE, FBIM

Written not for the family bookshelf but for the electronics enthusiast's workshop bench. Its aim is to bridge the gap between complicated technical theory, and "cut-and-tried" methods which may bring success in design but leave the experimenter unfulfilled.

There is, therefore, a strong practical bias – tedious and higher mathematics have been avoided where possible and many tables have been included, partly to save calculation and partly because actual figures bring a greater intimacy with the design process.

As a reference book, sections have been written to be as self-contained as possible. The book is divided into six basic sections: Units and Constants, Direct-current Circuits, Passive Components, Alternating-current Circuits, Networks and Theorems, Measurements.

0 900162 70 8 *256 pages* *1979* £2.95

THE SIMPLE ELECTRONIC CIRCUIT AND COMPONENTS **BP62**

ALTERNATING CURRENT THEORY **BP63**

SEMICONDUCTOR TECHNOLOGY **BP64**

(Elements of Electronics – Books 1, 2 and 3 respectively)

F. A. Wilson, CGIA, CEng, FIEE, FIERE, FBIM

The aim of this series of books can be stated quite simply – it is to provide an inexpensive introduction to modern electronics so that the reader will start on the right road by thoroughly understanding the fundamental principles involved.

That the books are inexpensive is perhaps obvious. By accepting their modest size the total information is that which in normal-size text books would cost some 2 to 3 times as much.

Although written especially for readers with no more than ordinary arithmetical skills, the use of mathematics is not avoided, and all mathematics required is taught as the reader progresses. There is little doubt that some skill in electronics and mathematics will be an asset to almost everybody in the future, destined as we are to be controlled more and more by computers, microprocessors and the like. There is also much to offer those who need revision or to be brought up-to-date with fundamentals, those with grown-up children asking awkward questions or preparing for entry to a college or even those who may find a different approach refreshing.

The course is not written to a specific syllabus but concentrates on the understanding of the important concepts central to electronics rather than continually digressing over the whole field on the basis that once the fundamentals are mastered, the technicalities of most other things are soon revealed. The author anticipates where the difficulties lie and hopefully guides the reader through them.

Each book is a complete treatise of a particular branch of the subject and, therefore, can be used on its own with one proviso, that the later books do not duplicate material from their predecessors, thus a working knowledge of the subjects covered by the earlier books is assumed.

BOOK 1: This book contains all the fundamental theory necessary to lead to a full understanding of the simple electronic circuit and its main components.

BOOK 2: This book continues with alternating current theory without which there can be no comprehension of speech, music, radio, television or even the electricity mains.

BOOK 3: Follows on semiconductor technology, leading up to transistors and integrated circuits.

Books 1–3 therefore constitute a complete but inexpensive basic electronic theory course of inestimable help in either career or hobby. Other books are on specialised electronic subjects – they are all built on this earlier groundwork. Book 4: *Microprocessing Systems and Circuits* is described under book number BP77; Book 5: *Communication* under book number BP89; Book 6: *Audio* under book number BP111.

Book 1: 0 900162 82 1	*224 pages*	*1979*	**£3.50**
Book 2: 0 900162 83 X	*224 pages*	*1979*	**£3.50**
Book 3: 0 900162 84 8	*224 pages*	*1979*	**£3.50**

MICROPROCESSING SYSTEMS AND CIRCUITS **BP77**
(Elements of Electronics – Book 4)
F. A. Wilson, CGIA, CEng, FIEE, FIERE, FBIM

A truly comprehensive guide to the elements of microprocessing systems which really starts at the beginning. Teaches the reader the essential fundamentals that are so important for a sound understanding of the subject.

0 900162 97 X *256 pages* *1980* **£2.95**

COMMUNICATION **BP89**
(Elements of Electronics – Book 5)
F. A. Wilson, CGIA, CEng, FIEE, FIERE, FBIM

A look at the electronic fundamentals over the whole of the communication scene. This book aims to teach the important elements of each branch of the subject in a style as interesting and practical as possible. While not getting involved in the more complicated theory and mathematics, most of the modern transmission system techniques are examined including line, microwave, submarine, satellite and digital multiplex systems, radio and telegraphy. To assist in understanding these more thoroughly, chapters on signal processing, the electromagnetic wave, networks and transmission assessment are included, finally a short chapter on optical transmission. Preparing the way, the opening chapter looks at the very heart of communication, including channel capacity, information flow, cables and electro-acoustic transducers.

The book follows its predecessors in layout and aims. It is not an expert's book but neither is it for those looking for the easy way – it is of serious intent but interesting and with the objective of leaving the reader knowledgeable and with a good technical understanding of such an extensive subject.

Ideal for readers who plan to enter a technical college or university and wish to do so with some fair appreciation of the subject. Also invaluable for people with careers in mind or who need revision or updating or those who just wish to study an absorbing subject at home.

0 85934 064 3 *256 pages* *1981* **£2.95**

AUDIO **BP111**

(Elements of Electronics – Book 6)

F. A. Wilson, CGIA, CEng, FIEE, FIERE, FBIM

Analysis of the sound wave and an explanation of acoustical quantities prepare the way. These are followed by a study of the mechanism of hearing and examination of the various sounds we hear. A look at room acoustics with a subsequent chapter on microphones and loudspeakers then sets the scene for the main chapter on audio systems – amplifiers, oscillators, disc and magnetic recording and electronic music.

The book follows its predecessors in layout and aims. It is not the expert's book but neither is it for those looking for the easy way – it is of serious intent but interesting and with the objective of leaving the reader knowledgeable with a good technical understanding of such an extensive subject.

Ideal for readers who plan to enter a technical college or university and wish to do so with some fair appreciation of the subject. Also invaluable for people with careers in mind or who need revision or updating or those who just wish to study an absorbing subject at home.

0 85934 086 4 *320 pages* *1985* **£3.50**

Notes

Please note following is a list of other titles that are available in our range of Radio, Electronics and Computer books.

These should be available from all good Booksellers, Radio Components Dealers and Mail Order Companies.

However, should you experience difficulty in obtaining any title in your area, then please write directly to the Publisher enclosing payment to cover the cost of the book plus adequate postage.

If you would like a complete catalogue of our entire range of Radio, Electronics and Computer books then please send a Stamped Addressed Envelope to:–

BERNARD BABANI (publishing) LTD
THE GRAMPIANS
SHEPHERDS BUSH ROAD
LONDON W6 7NF
ENGLAND

160	Coil Design and Construction Manual	£2.50
202	Handbook of Integrated Circuits (ICs) Equivalents and Substitutes	£2.95
205	Hi-Fi Loudspeaker Enclosures	£2.95
208	Practical Stereo and Quadrophony Handbook	£0.75
214	Audio Enthusiast's Handbook	£0.85
219	Solid State Novelty Projects	£0.85
220	Build Your Own Solid State Hi-Fi and Audio Accessories	£0.85
221	28 Tested Transistor Projects	£2.95
222	Solid State Short Wave Receivers for Beginners	£1.95
223	50 Projects Using IC CA3130	£1.25
224	50 CMOS IC Projects	£2.95
225	A Practical Introduction to Digital ICs	£1.75
226	How to Build Advanced Short Wave Receivers	£2.95
227	Beginners Guide to Building Electronic Projects	£1.95
228	Essential Theory for the Electronics Hobbyist	£2.50
BP1+14	First & Second Books of Transistor Equivalents & Substitutes	£3.50
BP2	Handbook of Radio, TV, Industrial and Transmitting Tube and Valve Equivalents	£0.60
BP6	Engineer's and Machinist's Reference Tables	£1.25
BP7	Radio and Electronic Colour Codes Data Chart	£0.95
BP27	Chart of Radio, Electronic, Semiconductor and Logic Symbols	£0.95
BP28	Resistor Selection Handbook	£0.60
BP29	Major Solid State Audio Hi-Fi Construction Projects	£0.85
BP33	Electronic Calculator Users Handbook	£1.50
BP34	Practical Repair and Renovation of Colour TVs	£2.95
BP36	50 Circuits Using Germanium Silicon and Zener Diodes	£1.50
BP37	50 Projects Using Relays, SCRs and TRIACs	£1.95
BP39	50 (FET) Field Effect Transistor Projects	£1.75
BP42	50 Simple LED Circuits	£1.95
BP44	IC 555 Projects	£2.50
BP45	Projects in Opto-electronics	£1.95
BP48	Electronic Projects for Beginners	£1.95
BP49	Popular Electronic Projects	£2.50
BP53	Practical Electronics Calculations and Formulae	£2.95
BP54	Your Electronic Calculator and Your Money	£1.35
BP56	Electronic Security Devices	£2.50
BP58	50 Circuits Using 7400 Series ICs	£2.50
BP59	Second Book of CMOS IC Projects	£1.95
BP60	Practical Construction of Pre-amps, Tone Controls, Filters and Attenuators	£1.95
BP61	Beginners Guide to Digital Techniques	£1.95
BP62	The Simple Electronic Circuit & Components (Elements of Electronics – Book 1)	£3.50
BP63	Alternating Current Theory (Elements of Electronics – Book 2)	£3.50
BP64	Semiconductor Technology (Elements of Electronics – Book 3)	£3.50
BP65	Single IC Projects	£1.50
BP66	Beginners Guide to Microprocessors and Computing	£1.95
BP67	Counter Driver and Numeral Display Projects	£2.95
BP68	Choosing and Using Your Hi-Fi	£1.65
BP69	Electronic Games	£1.75
BP70	Transistor Radio Fault-finding Chart	£0.95
BP71	Electronic Household Projects	£1.75
BP72	A Microprocessor Primer	£1.75
BP73	Remote Control Projects	£2.50
BP74	Electronic Music Projects	£2.50
BP75	Electronic Test Equipment Construction	£1.75
BP76	Power Supply Projects	£2.50
BP77	Microprocessing Systems and Circuits (Elements of Electronics – Book 4)	£2.95
BP78	Practical Computer Experiments	£1.75
BP79	Radio Control for Beginners	£1.75
BP80	Popular Electronic Circuits – Book 1	£2.95
BP82	Electronic Projects Using Solar Cells	£1.95
BP83	VMOS Projects	£1.95
BP84	Digital IC Projects	£1.95
BP85	International Transistor Equivalents Guide	£3.50
BP86	An Introduction to BASIC Programming Techniques	£1.95
BP87	50 Simple LED Circuits – Book 2	£1.35
BP88	How to Use Op-Amps	£2.95
BP89	Communication (Elements of Electronics – Book 5)	£2.95
BP90	Audio Projects	£1.95
BP91	An Introduction to Radio DXing	£1.95
BP92	Electronics Simplified – Crystal Set Construction	£1.75
BP93	Electronic Timer Projects	£1.95
BP94	Electronic Projects for Cars and Boats	£1.95
BP95	Model Railway Projects	£1.95
BP97	IC Projects for Beginners	£1.95
BP98	Popular Electronic Circuits – Book 2	£2.25
BP99	Mini-matrix Board Projects	£1.95
BP101	How to Identify Unmarked ICs	£0.95
BP103	Multi-circuit Board Projects	£1.95
BP104	Electronic Science Projects	£2.25

BP105	Aerial Projects	£1.95
BP106	Modern Op-amp Projects	£1.95
BP107	30 Solderless Breadboard Projects – Book 1	£2.25
BP108	International Diode Equivalents Guide	£2.25
BP109	The Art of Programming the 1K ZX81	£1.95
BP110	How to Get Your Electronic Projects Working	£1.95
BP111	Audio (Elements of Electronics – Book 6)	£3.50
BP112	A Z-80 Workshop Manual	£3.50
BP113	30 Solderless Breadboard Projects – Book 2	£2.25
BP114	The Art of Programming the 16K ZX81	£2.50
BP115	The Pre-computer Book	£1.95
BP117	Practical Electronic Building Blocks – Book 1	£1.95
BP118	Practical Electronic Building Blocks – Book 2	£1.95
BP119	The Art of Programming the ZX Spectrum	£2.50
BP120	Audio Amplifier Fault-finding Chart	£0.95
BP121	How to Design and Make Your Own P.C.B.s	£1.95
BP122	Audio Amplifier Construction	£2.25
BP123	A Practical Introduction to Microprocessors	£1.95
BP124	Easy Add-on Projects for Spectrum, ZX81 & Ace	£2.75
BP125	25 Simple Amateur Band Aerials	£1.95
BP126	BASIC & PASCAL in Parallel	£1.50
BP127	How to Design Electronic Projects	£2.25
BP128	20 Programs for the ZX Spectrum and 16K ZX81	£1.95
BP129	An Introduction to Programming the ORIC-1	£1.95
BP130	Micro Interfacing Circuits – Book 1	£2.25
BP131	Micro Interfacing Circuits – Book 2	£2.25
BP132	25 Simple Shortwave Broadcast Band Aerials	£1.95
BP133	An Introduction to Programming the Dragon 32	£1.95
BP134	Easy Add-on Projects for Commodore 64, Vic-20, BBC Micro and Acorn Electron	£2.95
BP135	Secrets of the Commodore 64	£1.95
BP136	25 Simple Indoor and Window Aerials	£1.75
BP137	BASIC & FORTRAN in Parallel	£1.95
BP138	BASIC & FORTH in Parallel	£1.95
BP139	An Introduction to Programming the BBC Model B Micro	£1.95
BP140	Digital IC Equivalents and Pin Connections	£5.95
BP141	Linear IC Equivalents and Pin Connections	£5.95
BP142	An Introduction to Programming the Acorn Electron	£1.95
BP143	An Introduction to Programming the Atari 600/800 XL	£1.95
BP144	Further Practical Electronics Calculations and Formulae	£4.95
BP145	25 Simple Tropical and MW Band Aerials	£1.75
BP146	The Pre-BASIC Book	£2.95
BP147	An Introduction to 6502 Machine Code	£2.50
BP148	Computer Terminology Explained	£1.95
BP149	A Concise Introduction to the Language of BBC BASIC	£1.95
BP150	An Introduction to Programming the Sinclair QL	£1.95
BP152	An Introduction to Z80 Machine Code	£2.75
BP153	An Introduction to Programming the Amstrad CPC 464 and 664	£2.50
BP154	An Introduction to MSX BASIC	£2.50
BP155	International Radio Stations Guide	£2.95
BP156	An Introduction to QL Machine Code	£2.50
BP157	How to Write ZX Spectrum and Spectrum + Games Programs	£2.50
BP158	An Introduction to Programming the Commodore 16 and Plus 4	£2.50
BP159	How to Write Amstrad CPC 464 Games Programs	£2.50
BP161	Into the QL Archive	£2.50
BP162	Counting on QL Abacus	£2.50
BP163	Writing with QL Quill	£2.50
BP164	Drawing on QL Easel	£2.50
BP169	How to Get Your Computer Programs Running	£2.50
BP170	An Introduction to Computer Peripherals	£2.50
BP171	Easy Add-on Projects for Amstrad CPC 464, 664/6128 and MSX Computers	£2.95
BP173	Computer Music Projects	£2.95
BP174	More Advanced Electronic Music Projects	£2.95
BP175	How to Write Word Game Programs for the Amstrad CPC 464, 664 and 6128	£2.95
BP176	A TV-DXers Handbook	£5.95
BP177	An Introduction to Computer Communications	£2.95
BP178	An Introduction to Computers in Radio	£2.95
BP179	Electronic Circuits for the Computer Control of Robots	£2.95
BP180	Computer Projects for Model Railways	£2.95
BP181	Getting the Most from Your Printer	£2.95
BP182	MIDI Projects	£2.95
BP183	An Introduction to CP/M	£2.95
BP184	An Introduction to 68000 Assembly Language	£2.95
BP185	Electronic Synthesiser Construction	£2.95
BP186	Walkie-Talkie Projects	£2.95
BP187	A Practical Reference Guide to Word Processing on the Amstrad PCW 8256 and PCW 8512	£5.95
BP188	Getting Started with BASIC and LOGO on the Amstrad PCW 8256 and PCW 8512	£6.95
BP189	Using Your Amstrad CPC Disc Drives	£2.95
BP190	More Advanced Electronic Security Projects	£2.95
BP191	Simple Applications of the Amstrad CPCs for Writers	£2.95
BP192	More Advanced Power Supply Projects	£2.95
BP193	Starting LOGO	£2.95
BP194	Modern Opto Device Projects	£2.95
BP195	An Introduction to Communications and Direct Broadcast Satellites	£3.95
BP196	BASIC & LOGO in Parallel	£2.95

BERNARD BABANI BP144

Further Practical Electronics Calculations and Formulae

- This book is the companion volume to BP53 PRACTICAL ELECTRONICS CALCULATIONS AND FORMULAE, carrying on from where the first book leaves off. However, each of the two books stands on its own. The first book embraces components, elementary circuit analysis, networks and measurements. This book continues by covering many aspects of electronics where a knowledge and familiarity of the appropriate formulae is essential for a fuller understanding of the subject.

- Written in the same style as the first book, so as to be a practical workshop manual. Tedious and higher mathematics have been deliberately kept low key and many worked examples are included.

- A glance at the contents page will show the many areas covered by this book. Each of the fourteen chapters are as self-contained as possible, with all key formulae developed or quoted being repeated at the end for quick and easy reference.

- An essential addition to the library of all those interested in electronics be they amateur, professional or student.

£4.95

ISBN 0-85934-119-4
00495
9 780859 341196