output of the machine? Find the armature current and power factor for this condition. (*H.N.C.*)

Ans. 10·9 MW; 1,480 A; 0·737 leading.

12.9 Show that the maximum power that a synchronous generator can supply when connected to constant-voltage constant-frequency busbars increases with the excitation.

An 11 kV 3-phase star-connected turbo-alternator delivers 240 A at unity power factor when running on constant voltage and frequency busbars. If the excitation is increased so that the delivered current rises to 300 A, find the power factor at which the machine now works and the percentage increase in the induced e.m.f. assuming a constant steam supply and unchanged efficiency. The armature resistance is 0·5 Ω per phase and the synchronous reactance 10 Ω per phase. (*H.N.C.*)

Ans. 0·802 lagging; 24 per cent.

12.10 An 11 kV 300 MVA 3-phase alternator has a steady short-circuit current equal to half its rated value. Determine graphically or otherwise the maximum load the machine can deliver when connected to 11 kV constant-voltage constant-frequency busbars with its field excited to give an open-circuit voltage of 12·7 kV/phase. Find also the armature current and power factor corresponding to this load. Ignore armature resistance. (*H.N.C.*)

Ans. 300 MW; 17·4 kA; 0·895 leading.

12.11 An alternator having a synchronous impedance of $R + jX$ ohms/phase is supplying constant voltage and frequency busbars. Describe, with the aid of complexor diagrams, the changes in current and power factor when the excitation is varied over a wide range, the steam supply remaining unchanged. The complexor diagrams should show the locus of the induced e.m.f.

A star-connected alternator supplies 300 A at unity power factor to 6,600 V constant voltage and frequency busbars. If the induced e.m.f. is now reduced by 20 per cent, the steam supply remaining unchanged, determine the new values of the current and power factor. Assume the synchronous reactance is 5 Ω/phase, the resistance is negligible and the efficiency constant. (*H.N.C.*)

Ans. 350 A; 0·85 leading.

12.12 Deduce an expression for the synchronizing power of an alternator.

Calculate the synchronizing power in kilowatts per degree of mechanical displacement at full load for a 1,000 kVA 6,600 V 0·8 power-factor 50 Hz 8-pole star-connected alternator having a negligible resistance and a synchronous reactance of 60 per cent. (*L.U.*)

Ans. 158 kW per mechanical degree.

12.13 A 40 MVA 50 Hz 3,000 rev/min turbine-driven alternator has an equivalent moment of inertia of 1,310 kg-m^2, and the machine has a steady short-circuit current of four times its normal full-load current.

Deducing any formula used, estimate the frequency at which hunting may take place when the alternator is connected to an "infinite" grid system. (*H.N.C.*)

Ans. 3·14 Hz.

12.14 An 11 kV 3-phase star-connected turbo-alternator is connected to constant-voltage constant-frequency busbars. The armature resistance is negligible and the synchronous reactance is 10 Ω. The alternator is excited to deliver 300 A at unity power factor. Determine the armature current and power factor when the excitation is increased by 25 per cent, the load on the machine being unchanged. (*H.N.C.*)

Ans. 356 A; 0·843 lagging.

12.15 Derive an expression for the synchronous reactance per phase of a 3-phase cylindrical-rotor synchronous machine assuming the armature m.m.f. is sinusoidally distributed, the armature winding uniform and the air-gap uniform. Neglect armature leakage flux and all reluctance except that of the air-gap.

Using these assumptions, calculate the approximate synchronous reactance per phase of a 3-phase 200 MVA 11 kV 2-pole 50 Hz star-connected synchronous generator having a stator internal diameter of 1 m, a core length of 4·6 m and an air-gap length of 4·7 cm. A flux per pole of 5 Wb is required to give an e.m.f. equal to rated voltage.

Ans. 1·21 Ω.

12.16 If a synchronous generator operating on infinite busbars has a fast-acting field excitation system, it is possible to operate the machine beyond the steady-state limit of stability at load angles in excess of 90 electrical degrees.

A 3-phase 75 MVA 11·8 kV star-connected synchronous generator has such an excitation system. Determine the maximum leading MVAr the machine can supply to the infinite busbars to which the machine is synchronized without exceeding its rating. The load on the machine is 30 MW. What is the load angle for this condition?

The synchronous reactance is 3 Ω/phase and the resistance is negligible.

Ans. 68·5 MVAr leading; 128°.

12.17 Explain why a synchronous machine synchronized to infinite busbars tends to remain synchronized.

A 3-phase 75 MVA 11·8 kV 50 Hz 2-pole synchronous generator has a period of oscillation of 1·3 s when synchronized to infinite busbars, the excitation and steam supply to the prime mover being so adjusted that there is no current transfer under steady conditions.

Neglecting the effect of damping, calculate the moment of inertia of the rotating system. The synchronous reactance per phase may be taken as 3 Ω and the resistance per phase to be negligible.

Ans. 5×10^3 kg-m^2.

Chapter 13

INDUCTION MACHINES

It was shown in Chapter 11 that, when a polyphase stator winding is excited from a balanced polyphase supply, a stator m.m.f. distribution is set up and travels at synchronous speed given by eqn. (11.10) as

$$n_0 = \frac{f}{p} \tag{13.1}$$

Associated with the stator m.m.f. distribution is a flux density distribution which also travels at synchronous speed and is often referred to as a "rotating field".

The stator field induces voltages in the rotor phase windings so that a rotor m.m.f. distribution and an associated flux density distribution are set up. The rotor distributions travel at the same speed as the stator distribution. The axes of the stator and rotor distributions have an angular displacement, and as a result a torque acts on the rotor and causes it to accelerate in the same direction as the stator field.

The steady-state rotor speed is normally slightly less than synchronous so that the motor runs with a *per-unit slip*, s, defined as

$$s = \frac{n_0 - n_r}{n_0} \tag{13.2}$$

where n_r is the rotor speed.

At standstill, $n_r = 0$ and $s = 1$. For the rotor to reach synchronous speed ($n_r = n_0$ and $s = 0$), an external drive is necessary, since for this condition there is no rotor e.m.f. and hence no rotor current or torque. If the rotor is driven so that $n_r > n_0$, the slip becomes negative, the rotor torque opposes the external driving torque and the machine acts as an induction generator.

In all cases the slip speed is

$$n_s = n_0 - n_r \tag{13.3}$$

From eqn. (13.2),

$$n_s = sn_0 \tag{13.4}$$

and

$$n_r = (1 - s)n_0 \tag{13.5}$$

The frequency of the rotor e.m.f.s and currents is proportional to the difference in speed between the rotating field and the rotor, so that

$$f_r = (n_0 - n_r)p = sn_0p = sf$$

where p is the number of pole pairs. Hence

$$\frac{f_r}{f} = s \tag{13.6}$$

Similarly for the angular frequencies corresponding to f_r and f:

$$\frac{\omega_r}{\omega_0} = \frac{2\pi f_r}{2\pi f} = s \tag{13.7}$$

The 3-phase induction motor has a torque characteristic similar to that of the d.c. shunt motor, is robust, and is low in initial cost. Other forms of asynchronous machine are the a.c. commutator motor, which gives a wide range of speed control, and various types of single-phase motor, which are employed for fractional-horsepower drives, in individual units, and in traction.

13.1 Construction

The induction machine consists essentially of a stator, which carries a 3-phase winding, and a rotor. The stator winding is a 3-phase winding of one of the types described in Chapter 11, often being a narrow-spread mesh-connected closed winding. The winding is laid in open or half-closed slots in a laminated silicon-steel core.

The rotor winding is placed in half-closed or closed slots, the air-gap between stator and rotor being reduced to a minimum. There are two main types of rotor, the *wound rotor* and the *squirrel-cage rotor*. In the squirrel-cage rotor, solid conducting rods are inserted into closed slots, and at each end the rods are connected to a heavy short-circuiting ring. This forms a permanently short-circuited winding which is practically indestructible. In some smaller machines the conductors, end rings and fan are cast in one piece in aluminium. The cage rotor is cheap and robust, but suffers from the disadvantage of a low starting torque.

The wound rotor has a 3-phase winding with the same number of poles as the stator; the ends of the rotor winding may be brought out to three slip rings. The advantage of the wound-rotor machine is that an external starting resistance can be connected to the slip rings to give a large starting torque. This resistance is reduced to zero as the machine runs up to speed.

13.2 Equivalent Circuit of Induction Machine at Any Slip

The approximate equivalent circuit per phase of a polyphase induction machine at standstill ($s = 1$) is shown in Fig. 13.1. The equivalent circuit takes the same form as that adopted for the power transformer, since at standstill the induction machine consists of two polyphase windings linked by a common flux.

Unlike that of a power transformer, the magnetic circuit of the induction machine has an air-gap, and this makes the per-unit

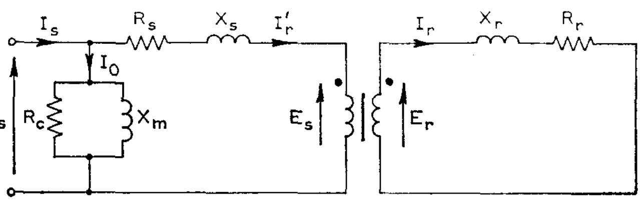


Fig. 13.1 EQUIVALENT CIRCUIT OF THE POLYPHASE INDUCTION MACHINE AT STANDSTILL

value of magnetizing current much higher than that of the power transformer. As a result the approximation of showing the shunt magnetizing branch of the equivalent circuit at the input terminals is less close than for the power transformer. The approximation is nevertheless acceptable for large machines, but not for small machines. To keep the magnetizing current as small as possible, the air-gap length of induction machines is made as short as is consistent with mechanical considerations.

A further difference between the polyphase induction machine and the power transformer is that in the former the windings are distributed, and this affects the effective turns ratio.

In this and subsequent sections it is assumed that the rotor has a 3-phase winding. A cage rotor is, in effect, a rotor with a large number of short-circuited phases. Such an arrangement may be represented by an equivalent 3-phase winding; I_r is not then the current in an actual rotor phase, but the stator current I_s is preserved as the true stator current.

The induced stator e.m.f. per phase when connected to a supply of frequency f hertz is, from eqn (11.20),

$$E_s = K_{ds}K_{ss}\frac{\omega_0\Phi N_s}{\sqrt{2}} \tag{13.8}$$

where $\omega_0 = 2\pi f = 2\pi n_0 p$.

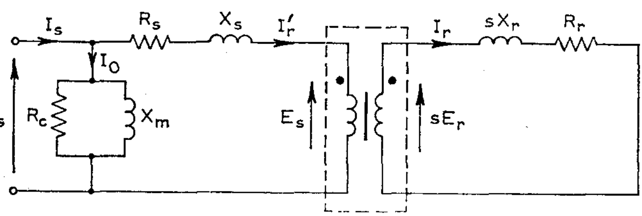


Fig. 13.2 EQUIVALENT CIRCUIT OF THE POLYPHASE INDUCTION MACHINE AT ANY SLIP, *s*

At standstill the frequency of the rotor e.m.f. per phase is the supply frequency and the rotor e.m.f. per phase is, from eqn (11.20),

$$E_r = K_{dr}K_{sr}\frac{\omega_0\Phi N_r}{\sqrt{2}} \tag{13.9}$$

When the rotor rotates the rotor e.m.f. per phase is altered both in size and frequency.

$$\left.\begin{array}{l}\text{Rotor e.m.f. per phase}\\ \quad\text{at any slip } s\end{array}\right\} = K_{dr}K_{sr}\frac{\omega_r\Phi N_r}{\sqrt{2}}$$

$$= K_{dr}K_{sr}\frac{s\omega_0\Phi N_r}{\sqrt{2}}$$

$$= sE_r$$

The rotor reactance per phase at standstill is X_r. At any slip s, therefore, the rotor reactance per phase will be sX_r, since reactance is proportional to frequency. The equivalent circuit per phase of a polyphase inductor motor at any slip s is shown in Fig. 13.2.

13.3 Slip Ratios

The element enclosed by the dotted box in Fig. 13.2 represents an ideal or lossless induction machine. This differs from the ideal transformer considered in Chapter 9 in that (*a*) the current and voltage transformation ratios differ, and (*b*) the frequencies of the voltages and currents at the input and output terminal pairs of the ideal element also differ.

From eqns. (13.8) and (13.9) the effective turns ratio, k_t, at standstill ($s = 1$) is

$$k_t = \frac{E_s}{E_r} = \frac{K_{ds}K_{ss}N_s}{K_{dr}K_{sr}N_r} \tag{13.10a}$$

At any slip s the voltage ratio is

$$\frac{E_s}{sE_r} = \frac{k_t}{s} \tag{13.10b}$$

At any slip s, m.m.f., balance must exist between the stator and rotor phase windings so that

$$I_r'\, K_{ds}K_{ss}N_s = I_r K_{dr}K_{sr}\, N_r$$

or

$$\frac{I_r'}{I_r} = \frac{K_{dr}K_{sr}N_r}{K_{ds}K_{ss}N_s} = \frac{1}{k_t} \tag{13.11}$$

Assuming there are three phases on both the stator and rotor,

$$\left.\begin{array}{c}\text{Power absorbed by ideal}\\ \text{stator winding}\end{array}\right\} P_0 = 3E_sI_r' \cos \phi_r \tag{13.12}$$

This power is obtained from the supply when the machine acts as a motor, and from the prime mover driving the rotor when it acts as a generator.

$$\left.\begin{array}{c}\text{Power dissipated in}\\ \text{the rotor circuit}\end{array}\right\} \quad P_r = 3sE_rI_r \cos \phi_r$$

$$= 3s\frac{E_s}{k_t} k_tI_r' \cos \phi_r$$

$$= 3sE_sI_r' \cos \phi_r \tag{13.13}$$

Dividing eqn. (13.12) by eqn. (13.13),

$$\frac{P_0}{P_r} = \frac{1}{s} \tag{13.14}$$

The power dissipated in the rotor circuit consists of winding loss in the rotor circuit and core loss in the rotor magnetic circuit. Since the core loss varies with frequency this implies that the equivalent circuit-element, R_r, is frequency dependent. Under normal running conditions, for plain induction motors, however, the rotor frequency and rotor core loss are low and the latter may usually be neglected. The power dissipated in the rotor is obtained from the ideal stator winding when the machine acts as a motor and from the prime mover when the machine acts as a generator.

When the machine acts as a motor the power absorbed by the ideal stator windings is greater than that dissipated in the rotor circuit except when the rotor is stationary (standstill) when they are equal. The difference in these two powers appears as gross mechanical power output:

$$\text{Mechanical power, } P_m = P_0 - P_r = P_0 - sP_0$$

or

$$P_m = P_0(1 - s) \tag{13.15}$$

Combining eqns. (13.14) and (13.15),

$$P_0 : P_r : P_m = 1 : s : (1 - s) \tag{13.16}$$

When the machine acts as a generator the net mechanical power input is the sum of the stator and rotor powers:

$$\text{Mechanical power, } P_m = P_0 + P_r$$

or

$$P_m = P_0(1 + s) \tag{13.17}$$

For generator action, therefore,

$$P_0 : P_r : P_m = 1 : s : (1 + s) \tag{13.18}$$

Eqn. (13.16) will serve for both motor and generator action if the slip s, the power absorbed by the ideal stator winding P_0 and the mechanical power P_m are taken to be negative for generator action, and it is remembered that P_m is the gross mechanical power output for the motoring mode and the net power input for the generating mode. Figs. 13.3(*a*) and (*b*) are block diagrams representing the power transfer in a plain induction machine for motor and generator action.

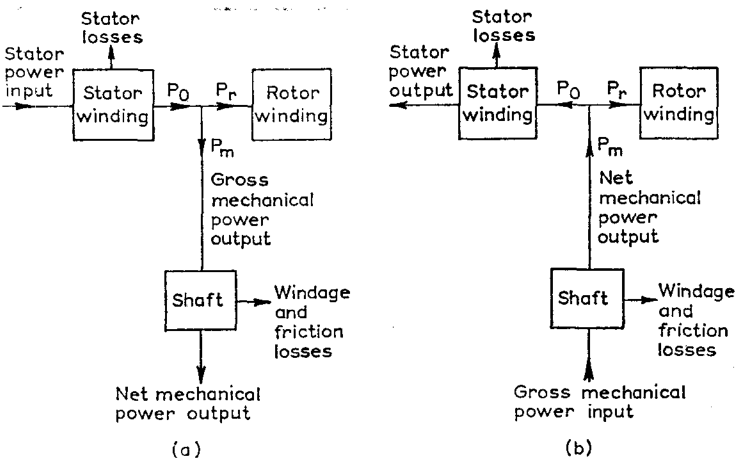


Fig. 13.3 POWER TRANSFER IN A PLAIN INDUCTION MACHINE
(*a*) Motoring mode (*b*) Generating mode

EXAMPLE 13.1 A 37·3 kW 4-pole 50 Hz induction machine has a friction and windage torque of 22 N-m. The stator losses equal the rotor circuit loss. Calculate:

(*a*) The input power to the stator when delivering full-load output at a speed of 1,440 rev/min.
(*b*) The gross input torque and stator output power when running at a speed of 1,560 rev/min. The stator losses are as in (*a*) and the windage and friction torque is unchanged.

(*a*) Synchronous speed $= \dfrac{f}{p} \times 60 = \dfrac{50 \times 60}{2} = 1{,}500\,\text{rev/min}$

$$\text{Per-unit slip, } s = \frac{n_0 - n_r}{n_r} = \frac{1{,}500 - 1{,}440}{1{,}500} = 0{\cdot}04$$

$$\text{Windage and friction loss} = 2\pi n_r T = \frac{2\pi \times 1{,}440 \times 22}{60} = 3{,}320\,\text{W}$$

$$\text{Gross mechanical power output, } P_m = 37{,}300 + 3{,}320 = 40{,}620\,\text{W}$$

$$\text{Power absorbed by ideal stator winding, } P_0 = \frac{P_m}{1-s} = \frac{40{,}620}{0{\cdot}96} = 42{,}300\,\text{W}$$

$$\text{Stator losses} = \text{Rotor loss} = sP_0 = 0{\cdot}04 \times 42{,}300 = 1{,}690\,\text{W}$$

$$\text{Stator input power} = 42{,}300 + 1{,}690 = \underline{\underline{44{,}000\,\text{W}}}$$

$$(b)\quad \text{Per-unit slip} = \frac{1{,}500 - 1{,}560}{1{,}500} = -0{\cdot}04$$

That is, the machine is now operating as an induction generator. Since the rotor circuit loss and the stator losses are equal and the latter are unchanged,

Rotor circuit loss, $P_r = 1{,}600\,\text{W}$

Net mechanical power input, $P_m = \dfrac{1+s}{s} P_r = \dfrac{1{\cdot}04}{0{\cdot}04} \times 1{,}690 = 44{,}000\text{W}$

Torque corresponding to net mechanical power input $= \dfrac{44{,}000 \times 60}{2\pi \times 1{,}560}$

$= 269\,\text{N-m}$

Gross input torque $= 269 + 22 = \underline{\underline{291\,\text{N-m}}}$

Power absorbed by ideal stator winding, $P_0 = \dfrac{P_m}{1+s} = \dfrac{44{,}000}{1{\cdot}04} = 42{,}300\,\text{W}$

Stator output power $= 42{,}300 - 1{,}690 = \underline{\underline{40{,}300\,\text{W}}}$

13.4 Transformer Equivalent Circuit of the Induction Machine

Referring to the equivalent circuit of Fig. 13.2, the rotor current per equivalent phase is

$$I_r = \frac{sE_r}{R_r + jsX_r}$$

If the numerator and denominator are divided by s this gives

$$I_r = \frac{E_r}{\dfrac{R_r}{s} + jX_r} \tag{13.19}$$

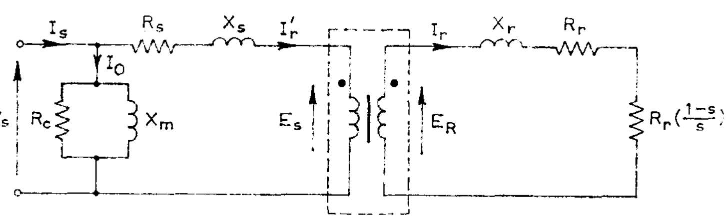


Fig. 13.4 TRANSFORMER EQUIVALENT CIRCUIT OF THE INDUCTION MACHINE

This latter expression for the rotor current per equivalent phase is consistent with the rotor equivalent circuit shown in Fig. 13.4. Although the value of I_r is unchanged this equivalent circuit is significantly different from that of Fig. 13.2 in that both the induced voltage and the reactance per equivalent rotor phase have their standstill ($s = 1$) values. Nor do the voltage and current ratios now differ as they did in the equivalent circuit of Fig. 13.1. Both are now equal to the virtual turns ratio, k_t. The element enclosed in the dotted box of Fig. 13.4 represents an ideal transformer. Therefore the power absorbed by the ideal stator winding must be equalled by the power delivered by the ideal rotor winding to R_r/s. Thus the power dissipated in the rotor equivalent circuit of Fig. 13.4 must be both the rotor loss and the gross mechanical power output. This may be shown to be so as follows.

$$\left.\begin{array}{l}\text{Added resistance per equivalent}\\ \quad\text{rotor phase}\end{array}\right\} = \frac{R_r}{s} - R_r = R_r\left(\frac{1-s}{s}\right)$$

$$\left.\begin{array}{l}\text{Power dissipated in}\\ \quad\text{added rotor resistance}\end{array}\right\} = 3I_r^2R_r\left(\frac{1-s}{s}\right) = P_r\left(\frac{1-s}{s}\right) = P_m$$

That is, the equivalent circuit has additional resistance of $R_r\left(\frac{1-s}{s}\right)$ in each phase and the power dissipated in these additional resistances is equal to the gross mechanical power output. Fig. 13.4 shows the equivalent circuit of an induction machine which is also the equivalent circuit of a transformer the power dissipated in whose secondary load is equal to the gross mechanical power output of the induction machine when operating in the motoring mode.

Since the equivalent circuit of Fig. 13.4 is that of a transformer, the secondary values may be referred to the primary by multiplying them by k_t^2, the square of the virtual turns ratio. The turns ratio is defined in eqn. (13.10*a*). Fig 13.5 shows the referred equivalent circuit in which

$$\frac{R_r'}{s} = k_t^2\frac{R_r}{s} \quad \text{and} \quad X_r' = k_t^2X_r$$

It should be noted that the equivalent circuits are valid only if the variations in speed or slip are relatively slow. Usually the moment of inertia of the rotor is sufficiently large for this condition to be realized.

13.5 Torque developed by an Induction Machine

An expression for the torque developed by an induction machine may be obtained by reference to the equivalent circuit of Fig. 13.5. Assuming that the stator winding has three phases,

$$\text{Rotor circuit loss, } P_r = 3(I_r')^2R_r' \tag{13.20}$$

$$\left.\begin{array}{l}\text{Power absorbed by ideal}\\ \quad\text{stator winding}\end{array}\right\} P_0 = \frac{P_r}{s} = 3(I_r')^2\frac{R_r'}{s} \tag{13.21}$$

$$\left.\begin{array}{r}\text{Gross mechanical}\\ \text{power output}\end{array}\right\} P_m = 3(I_r')^2 R_r' \frac{1-s}{s} \tag{13.22}$$

$$\text{Gross torque developed, } T = \frac{3}{2\pi n_r}(I_r')^2 R_r' \frac{1-s}{s} \tag{13.23}$$

where $\quad n_r = (1 - s)n_0 \qquad (13.5)$

Substituting for n_r in eqn. (13.23),

$$T = \frac{3}{2\pi n_0}(I_r')^2 \frac{R_r'}{s} \tag{13.24}$$

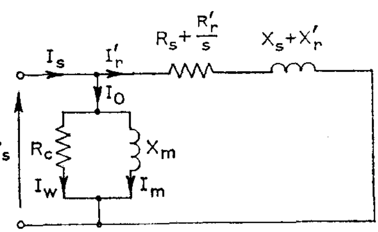


Fig. 13.5 REFERRED TRANSFORMER EQUIVALENT CIRCUIT OF THE INDUCTION MACHINE

Comparing eqn. (13.23) with eqn. (13.21) it will be seen that

$$T = \frac{P_0}{2\pi n_0} \tag{13.25}$$

Thus the torque developed is proportional to the power, P_0, absorbed by the ideal stator winding . The quantity P_0 is sometimes referred to as the torque measured in "synchronous watts", which presumably implies that, if this power is divided by the synchronous angular velocity, the torque is obtained.

Referring to the equivalent circuit of Fig. 13.5, evidently

$$I_r' = \frac{V_s}{\sqrt{\left\{\left(R_s + \frac{R_r'}{s}\right)^2 + (X_s + X_r')^2\right\}}}$$

Substituting for $(I_r')^2$ in eqn. (13.24),

$$T = \frac{3}{2\pi n_0} \frac{V_s^2}{\left(R_s + \frac{R_r'}{s}\right)^2 + (X_s + X_r')^2} \frac{R_r'}{s} \tag{13.26}$$

Multiplying numerator and denominator by s gives

$$T = \frac{3V_s^2}{2\pi n_0} \frac{sR_r'}{(sR_s + R_r')^2 + s^2(X_s + X_r')^2} \tag{13.27}$$

If the stator impedance is neglected, this equation reduces to

$$T = \frac{3V_s^2}{2\pi n_0} \frac{sR_r'}{(R_r')^2 + s^2(X_r')^2} \tag{13.28}$$

Eqns. (13.27) and (13.28) have been obtained by considering motor action, but they apply equally to generator action. If the torque is taken to be positive when the machine acts as a motor it will be negative for generator action since the slip then becomes negative. For motor action eqn. (13.28) gives the gross torque developed, but for generator action it gives the net input torque, as will be evident from a consideration of Figs. 13.2(*a*) and (*b*), the torque in each case being $P_m/2\pi n_r$.

13.6 Slip/Torque Characteristics of the Induction Machine

Referring to the expression for torque given by eqn. (13.27), since the slip s is positive in both the numerator and the denominator, the

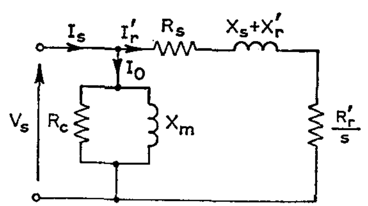


Fig. 13.6 PERTAINING TO MAXIMUM DEVELOPED TORQUE

torque will be zero both at $s = 0$ and $s = \infty$. Therefore, the torque will be a maximum at some intermediate value of slip.

From eqn. (13.24),

$$T = \frac{3}{2\pi n_0} (I_r')^2 \frac{R_r'}{s} \tag{13.24}$$

Therefore the torque will be a maximum when there is maximum power transfer into the load R_r'/s. As shown in Fig. 13.6, the source impedance is $R_s + j(X_s + X_r')$ assuming the supply itself to have zero internal impedance. According to Section 2.4 the maximum power is transferred to R_r'/s when

$$\frac{R_r'}{s} = \pm\sqrt{\{R_s^2 + (X_s + X_r')^2\}}$$

or

$$s = \pm\frac{R_r'}{\sqrt{\{R_s^2 + (X_s + X_r')^2\}}} \tag{13.29}$$

The plus sign refers to motor action, the minus sign to generator action.

An expression for maximum torque may be obtained by substituting in eqn. (13.27) the value of s given in eqn. (13.29). The algebra is a little simplified if in the first instance a substitution for $s^2(X_s + X_r')^2$ is made. From eqn. (13.29),

$$s^2(X_s + X_r')^2 = (R_r')^2 - s^2R_s^2$$

In eqn. (13.27),

$$T_{max} = \frac{3V_s^2}{2\pi n_0}\frac{sR_r'}{(sR_s + R_r')^2 + (R_r')^2 - s^2R_s^2}$$

$$= \frac{3V_s^2}{2\pi n_0}\frac{1}{2}\frac{s}{(sR_s + R_r')}$$

Substituting for s and simplifying further,

$$T_{max} = \pm\frac{3V_s^2}{2\pi n_0}\frac{1}{2}\left[\frac{1}{\sqrt{\{R_s^2 + (X_s + X_r')^2\}} \pm R_s}\right] \tag{13.30}$$

Here again the plus signs refer to motor action, the minus signs to generator action. It is to be noted that the maximum net input torque for generator action is greater than the maximum gross output torque for motor action. This inequality would disappear if the stator resistance were negligible. Although, as eqn. (13.29) shows, the slip at which the maximum torque occurs is proportional to the referred value of rotor resistance per stator phase, R_r', the actual value of the maximum torque in independent of R_r'. Therefore variation of R_r' changes the slip at which maximum torque occurs without affecting its value. The maximum torque is sometimes called the *pull-out torque*.

If stator impedance is neglected (i.e. $R_s = 0$, $X_s = 0$) eqn. (13.29) becomes

$$s = \pm\frac{R_r'}{X_r'} \tag{13.31}$$

and eqn. (13.30) becomes

$$T_{max} = \pm \frac{3}{2\pi n_0} \frac{1}{2X_r'} \tag{13.32}$$

It will be appreciated that, to obtain a high starting torque and a high maximum torque, the combined rotor and stator leakage reactance must be small. The shorter the air-gap is made the more the leakage flux is reduced. This is an additional reason for minimizing the air-gap.

Fig. 13.7 is a typical slip/torque characteristic of an induction machine. The hatched areas in the region of $s = 0$ show the normal operating range of the machine for motor and generator action. Referring to motor action, AB represents the torque at standstill or starting torque. Provided the load torque is less than this the motor will accelerate until the developed motor torque and load torque come into equality at a speed close to but less than synchronous speed. The machine operates stably as a motor over the range indicated. If the machine is operating in this region and a load torque in excess of the maximum motoring torque, CD, is imposed on the machine, it will decelerate to standstill or stall. The range DB represents unstable motor action.

Fig. 13.7 shows positive values of slip greater than unity. To achieve such values the rotor must be coupled to a prime mover and driven in the opposite direction to that of the stator rotating field, the stator still being connected to the 3-phase supply. In such conditions of operation the machine acts as neither a motor nor a generator as it receives both electrical and mechanical input power, all the power input being dissipated as loss. This mode of operation is referred to as *brake action.*

For the machine to operate as an induction generator a prime mover must drive the rotor in the same direction as the stator rotating field but at a higher speed with the stator connected to a pre-existing supply. In Fig. 13.7 OF represents the range of stable generator action. If the input torque to the generator exceeds the maximum generating torque, EF, the machine accelerates and passes into the region of unstable generator action, FG. Unless the input torque is removed, the speed may rise dangerously.

At some negative value of slip the machine will pass from unstable generator action to brake action. This will occur when the stator I^2R loss exceeds the power absorbed by the ideal stator winding, since the machine must then draw power from the electrical supply to meet completely the stator loss as well as having mechanical power input.

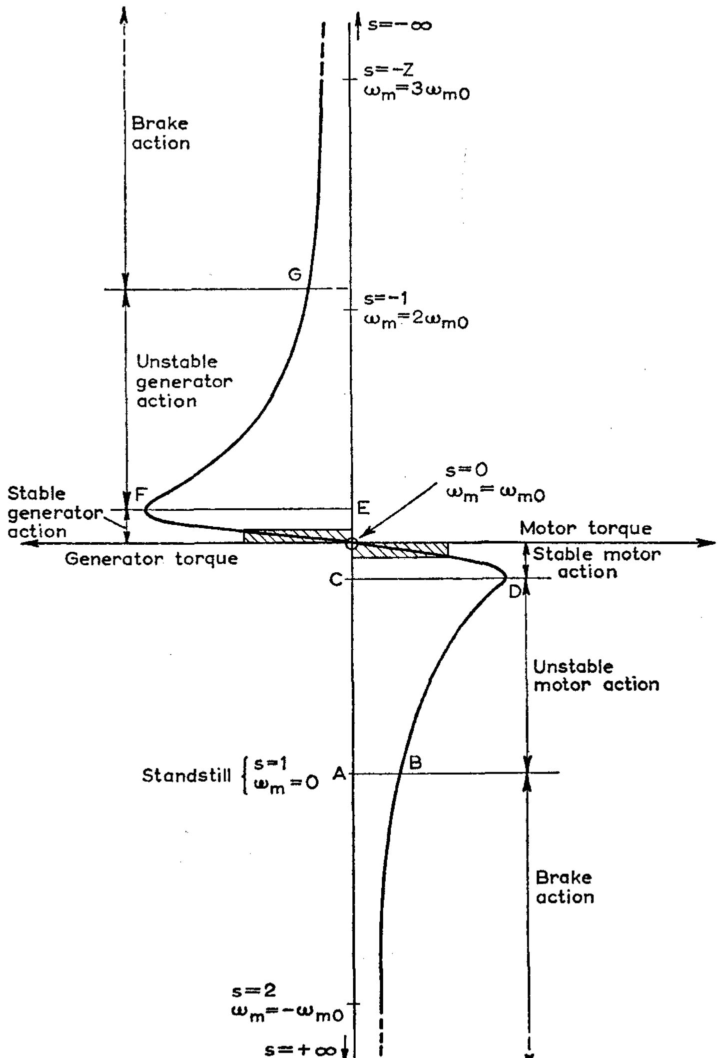


Fig. 13.7 SLIP/TORQUE CHARACTERISTIC OF THE INDUCTION MACHINE

$$\text{Stator } I^2R \text{ loss} = 3(I_r')^2R_s \tag{13.33}$$

$$\text{Power absorbed by ideal stator, } P_0 = -3(I_r')^2\frac{R_r'}{s} \tag{13.21}$$

P_0 is taken to be negative when the input to the ideal stator winding is derived from a mechanical source. Equating eqns. (13.33) and (13.21),

$$s = -\frac{R_r'}{R_s} \tag{13.34}$$

At this value of slip the machine changes from unstable generator to brake action. Since R_r' and R_s will be of the same order this value of slip will be approximately -1.

The slip/torque characteristic shown in Fig. 13.7 has its maximum value at a relatively small value of slip. This is typical of induction machines and is desirable, since when the machine operates as a motor the speed regulation with load is small, giving the machine a speed/torque characteristic over its operating region similar to that of the d.c. shunt motor. This matter is discussed further in Section 13.7. Further, since from eqn. (13.14) the slip is equal to $s = P_r/P_0$, an induction machine operating with a large value of slip would have a large rotor loss and consequently a low efficiency. A disadvantage of having the maximum torque occur at a low value of slip is that, as Fig. 13.7 shows, this arrangement makes the starting torque low.

EXAMPLE 13.2 A 440 V 4-pole 3-phase 50 Hz slip-ring induction motor has its stator winding mesh connected and its rotor winding star connected. The standstill voltage measured between slip rings with the rotor open-circuited is 218 V. The stator resistance per phase is 0·6 Ω and the stator reactance per phase is 3 Ω. The rotor resistance per phase is 0·05 Ω and the rotor reactance per phase is 0·25 Ω. Calculate the maximum torque and the slip at which it occurs. If the ratio of full-load to maximum torque is 1 : 2·5 find the full-load slip and the power output.

All values must be referred to either the stator or the rotor. It is usual to refer to the stator. Since the rotor is star connected:

$$\text{Induced standstill rotor voltage, } E_r = \frac{218}{\sqrt{3}} = 126\text{ V}$$

From eqn. (13.10*a*) the standstill turns ratio is

$$k_t = \frac{E_s}{E_r} = \frac{440}{126} = 3{\cdot}49$$

Rotor resistance/phase referred to stator, $R_r' = 0{\cdot}05 \times 3{\cdot}49^2 = 0{\cdot}61\ \Omega$
Rotor reactance/phase referred to stator, $X_r' = 0{\cdot}25 \times 3{\cdot}49^2 = 3{\cdot}05\ \Omega$

From eqn. (13.29) the slip for maximum torque is

$$s = \frac{R_r'}{\sqrt{\{(R_s')^2 + (X_s \times X_r')^2\}}} = \frac{0{\cdot}61}{\sqrt{(0{\cdot}6^2 + 6{\cdot}05^2)}} = \underline{\underline{0{\cdot}1}}$$

$$\text{Synchronous speed, } n_0 = \frac{f}{p} = \frac{50}{2} = 25\,\text{rev/s}$$

From eqn. (13.30) the maximum torque is

$$T_{max} = \frac{3V_s^2}{2\pi n_0}\frac{1}{2}\left[\frac{1}{\sqrt{\{R_s^2 + (X_s + X_r')^2\}} + R_s}\right]$$

$$= \frac{3 \times 440^2}{2\pi 25}\frac{1}{2}\frac{1}{6{\cdot}06 + 0{\cdot}6} = \underline{\underline{278\,\text{N-m}}}$$

$$\text{Full-load torque} = \frac{278}{2{\cdot}5} = 111\,\text{N-m}$$

The slip for full-load torque may be obtained from eqn. (13.27):

$$T = \frac{3V_s^2}{2\pi n_0}\frac{sR_r'}{(sR_s + R_r')^2 + s^2(X_s + X_r')^2} \tag{13.27}$$

i.e.

$$111 = \frac{3 \times 440^2}{2\pi \times 25}\frac{0{\cdot}61s}{(0{\cdot}6s + 0{\cdot}61)^2 + 6{\cdot}05^2 s^2}$$

This gives

$$s^2 - 0{\cdot}53s + 0{\cdot}01 = 0$$

so that

$$s = \frac{0{\cdot}53 \pm \sqrt{(0{\cdot}53^2 - 4 \times 0{\cdot}01)}}{2} = \underline{\underline{0{\cdot}02 \text{ or } 0{\cdot}51}}$$

Since the slip for maximum torque is $0{\cdot}1$, $s = 0{\cdot}02$ is on the stable part of the slip/torque characteristic and $s = 0{\cdot}51$ is on the unstable part. Selecting the value of s giving stable operation,

$$\text{Power output, } P_m = 2\pi n_r T = 2\pi \times 25(1 - 0{\cdot}02) \times 111 = \underline{\underline{17{\cdot}1\ \text{kW}}}$$

13.7 Starting

SLIP-RING MACHINES

To obtain a satisfactory operating characteristic giving a reasonable efficiency and a small speed regulation with load, the slip for the maximum torque developed by an induction motor must have a value in the range from $0{\cdot}1$ to $0{\cdot}2$, as explained at the end of Section 13.6. From eqn. (13.29) the slip for maximum torque is

$$s = \frac{R_r'}{\sqrt{\{R_s^2 + (X_s + X'_r)^2\}}} = \text{from } 0{\cdot}1 \text{ to } 0{\cdot}2$$

At starting $s = 1$, and to obtain maximum torque on starting,

$$\frac{R_r'}{\sqrt{\{R_s^2 + (X_s + X'_r)^2\}}} = 1$$

In the plain-cage-rotor induction motor these conflicting requirements cannot well be met, though specially shaped rotor slots or a double-cage rotor may make this possible. In slip-ring machines,

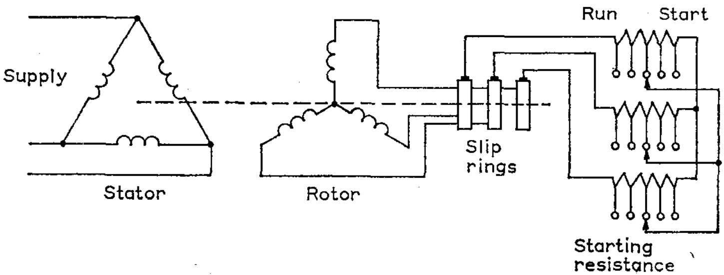


Fig. 13.8 STARTING OF WOUND-ROTOR MACHINES

however, the rotor resistance per phase is such as to give a satisfactory operating characteristic.

Slip-ring machines are invariably started by means of external resistances connected through the slip rings to the rotor circuit (Fig. 13.8). The machine is started with all the resistances in, giving a high starting torque. As the machine runs up to speed the

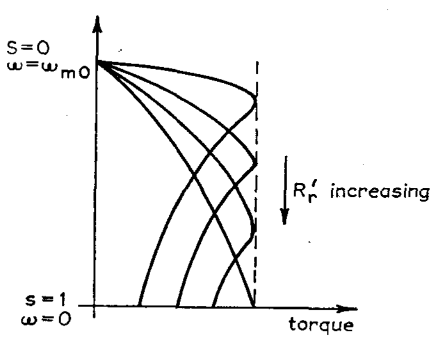


Fig. 13.9 SLIP/TORQUE CURVES FOR ROTOR-RESISTANCE STARTING

external resistance is reduced until the machine attains full speed with no external resistance.

Fig. 13.9 shows the slip/torque curves of a slip-ring induction motor corresponding to various positions of the starting resistance.

NON-SLIP-RING MACHINES

Stator starting must be used for cage-rotor machines, since no connexion can be made to the rotor, and direct switching of large machines would cause huge starting currents, which must be avoided. The cage-rotor machine suffers from the disadvantage that the starting torque is low if the resistance is low, while the efficiency is reduced if the rotor resistance is high. Various methods of starting will be examined.

Direct-on-line starting. In small workshops, direct-on-line starting may be restricted to motors of 2 kW or less, but in large industrial premises the tendency is to "direct switch" whenever possible, direct-on starting of motors of 40 kW at medium voltages being common.

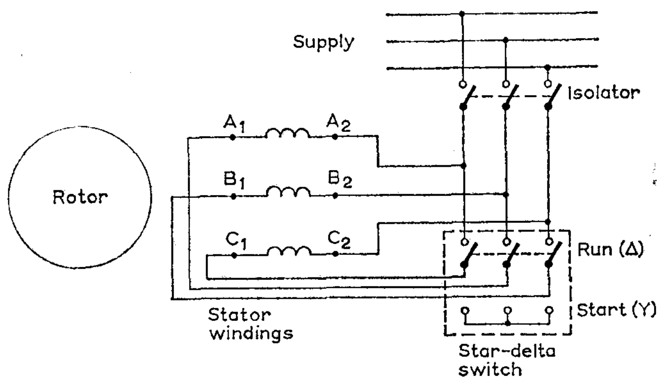


Fig. 13.10 STAR-DELTA STARTING

Reduced-voltage starters, described below, may be used to limit the initial starting torque and thus to reduce the shock to the driven machine.

Star-delta starting. In this type of starter the stator is star-connected for running up to speed, and is then delta connected. The applied voltage per phase in star is only $1/\sqrt{3}$ of the value which would be applied if the windings were connected in delta; hence, from eqn. (13.27), the starting torque is reduced to 1/3. The phase current in star is $1/\sqrt{3}$ of its value in delta, so that the line current for star connexion is 1/3 of the value for delta. Fig. 13.10 shows a connexion diagram for a star-delta starter.

Auto-transformer starting. In auto-transformer starting the transformer has at least three tappings giving open-circuit voltages of not less than 40, 60 and 75 per cent of line voltage for starting, and the stator is switched directly to the mains when the motor has run up to speed (Fig. 13.11). If the fractional tapping is x, then the applied voltage per phase on starting is xV_1 (where V_1 is the mains voltage), and the starting torque is reduced by x^2. The starting current from the mains will also be reduced by approximately x^2.

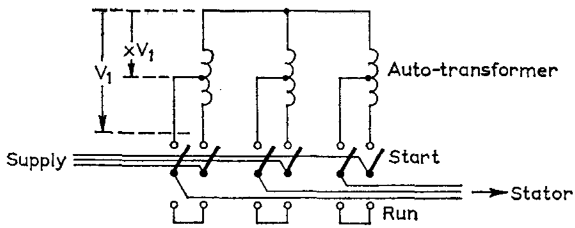


Fig. 13.11 AUTO-TRANSFORMER STARTING

EXAMPLE 13.3 A 3-phase squirrel-cage induction motor has a stator resistance per phase of 0·5 Ω and a rotor resistance per phase referred to the stator of 0·5 Ω. The total standstill reactance per phase referred to the stator is 4·92 Ω. If the ratio of maximum torque to full-load torque is 2:1, find the ratio of actual starting to full-load torque for (*a*) direct starting, (*b*) star-delta starting and (*c*) auto-transformer starting with a tapping of 75 per cent.

The maximum torque is given by eqn. (13.30):

$$T_{max} = \pm \frac{3V_s^2}{2\pi n_0}\frac{1}{2}\left[\frac{1}{\sqrt{\{R_s^2 + (X_s + X_r')^2\}} \pm R_s}\right]$$

For motor action

$$T_{max} = k\frac{V_s^2}{2}\frac{1}{\sqrt{(0{\cdot}5^2 + 4{\cdot}92^2)} + 0{\cdot}5} = \frac{kV_s^2}{10}$$

Full-load torque, $T_{FL} = \frac{1}{2}T_{max} = \dfrac{kV_s^2}{20}$

(*a*) *Direct-on-line starting.* The starting torque is obtained by substituting $s = 1$ in eqn. (13.27), which is

$$T = \frac{3V_s^2}{2\pi n_0}\frac{sR_r'}{(sR_s + R_r')^2 + s^2(X_s + X_r')^2} \qquad (13.27)$$

The starting torque is

$$T_0 = kV_s^2\frac{0{\cdot}5}{(0{\cdot}5 + 0{\cdot}5)^2 + 4{\cdot}92^2} = \frac{kV_s^2}{2 \times 24{\cdot}2}$$

Therefore

$$\frac{\text{Starting torque}}{\text{Full-load torque}} = \frac{T_0}{T_{FL}} = \frac{kV_s^2}{2 \times 24{\cdot}2} \times \frac{20}{kV_s^2} = \underline{\underline{0{\cdot}413}}$$

(*b*) *Star-delta starting.* The effective phase voltage is reduced to $1/\sqrt{3}$ of its original value. Therefore

$$T_0 = \left(\frac{1}{\sqrt{3}}\right)^2 \text{ of } T_0 \text{ for direct starting, and}$$

$$\frac{\text{Starting torque}}{\text{Full-load torque}} = \frac{0{\cdot}393}{3} = \underline{\underline{0{\cdot}131}}$$

(*c*) *Auto-transformer starting.* The effective phase voltage is reduced to 0·75 of its original value. Thus

$$T_0 = 0{\cdot}75^2 \text{ of } T_0 \text{ for direct starting}$$

and

$$\frac{\text{Starting torque}}{\text{Full-load torque}} = 0{\cdot}393 \times 0{\cdot}75^2 = \underline{\underline{0{\cdot}221}}$$

Where a high starting torque is required from a squirrel-cage motor, it may be achieved by a double-cage arrangement of the rotor conductors, as shown in Fig. 13.12(*a*). The equivalent electrical

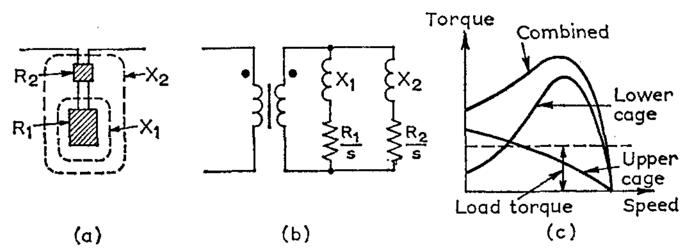


Fig. 13.12 IMPROVEMENT IN STARTING TORQUE OF CAGE ROTORS

(*a*) Double-cage rotor
(*b*) Equivalent circuit
(*c*) Combined torque/speed characteristics

rotor circuit is shown at (*b*), where X_1 and X_2 are leakage reactances. This equivalent circuit neglects mutual inductance between the cages. For the upper cage the resistance is made intentionally high, giving a high starting torque, while for the lower cage the resistance is low, and the leakage reactance is high, giving a low starting torque but high efficiency on load. The resultant characteristic will be roughly the sum of these two as shown in Fig. 13.12(*c*).

If a 3-phase induction motor starts in the wrong direction, this can be remedied by interchanging any two of the three supply leads to the stator.

13.8 Stability and Crawling

Curve (*a*) in Fig. 13.13 is the torque/speed curve for a typical induction motor. Consider that this motor is required to drive a constant-torque load having the torque/speed characteristic illustrated by curve (*b*). T_0, the starting torque with direct-on switching, is greater than the load torque T_b; thus there will be an excess starting torque (T_0–T_b) which will accelerate the motor and the load. The acceleration at any speed will be proportional to the torque difference (T–T_b)

so that the acceleration will be a maximum when the driving torque, T, is a maximum; and the acceleration will be zero, i.e. a steady speed will be obtained, when the speed corresponds to the operating point A. This is a stable operating point, since if the speed rose by a small amount Δn_r from its value at A, the load torque would exceed the driving torque and there would be a deceleration back to the

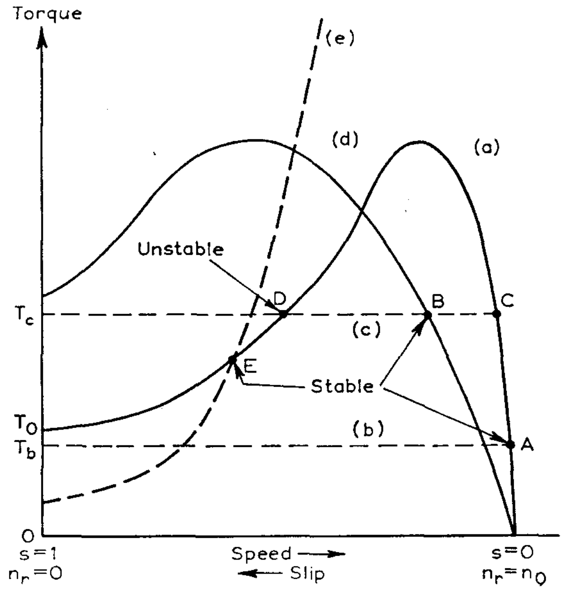


Fig. 13.13 PERTAINING TO STABILTY

--- Load characteristics
—— Motor characteristics

speed at A, and vice versa for a decrease in speed. The conclusions from this argument are:

(*a*) The operating point must be at the intersection of the two torque/speed characteristics.

(*b*) The slope of the load torque/speed curve must be greater than that of the driving-motor torque/speed curve for the operating point to be stable; i.e.

$$\frac{dT}{dn_r} \text{ for the load} > \frac{dT}{dn_r} \text{ for the drive}$$

Curve (*c*) in Fig. 13.13 represents a second load. In this case the load torque T_c is greater than the starting torque T_0 of the motor.

and with direct-on switching the motor would fail to start. The motor could be started by the use of additional rotor resistance sufficient to give the motor the characteristic of curve (*d*). The motor would drive the load at the speed corresponding to point C when the additional rotor resistance is short-circuited. Operation at point C may be unsatisfactory, since C is relatively near the maximum torque point and fluctuations in the load might too easily stall the motor. If the fluctuation does not cause the speed to fall

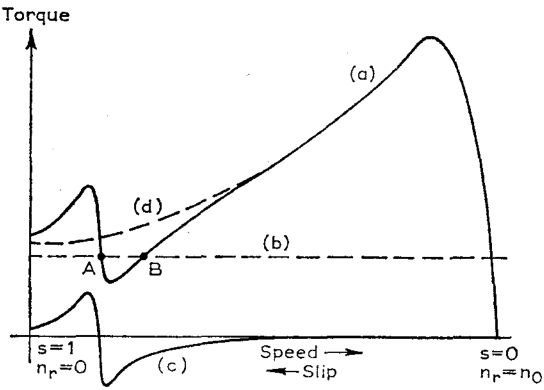


Fig. 13.14 TORQUE/SPEED CHARACTERISTICS TO ILLUSTRATE CRAWLING

(*a*) Resultant of (*c*) and (*d*)
(*b*) Constant-torque characteristic
(*c*) Characteristic due to 7th harmonic flux density
(*d*) Characteristic due to fundamental flux density

below that at D the motor should accelerate the load back to its speed at C when the load torque returns to normal. Since dT/dn_r for the load is not greater than dT/dn_r for the drive at D, this is an unstable operating point; i.e. if some random cause makes the speed fall slightly the load torque will exceed the driving torque and cause a further reduction in speed, or vice versa. The portion of the normal characteristic curve (*a*) which lies to the left of the maximum-torque point is called the *nominally unstable* portion. Though it is not normally possible for a motor to operate at a point on the nominally unstable portion of its characteristic, this may be arranged if a load, such as a fan, with a steep rising characteristic is chosen. A particular case is represented by curve (*e*); the motor would drive this load at a speed corresponding to the point E.

An induction motor may sometimes run in a stable manner at a low speed on a constant-torque load. This can be the result of a kink in the normal torque/speed characteristic. In Fig. 13.14

curve (*a*) shows a torque/speed curve with such a kink, and curve (*b*) represents a constant-torque load. The intersection A of curves (*a*) and (*b*) represents a stable operating point, so that the machine would not run up to full speed but merely drive the load at the speed corresponding to point A. This is termed *crawling*. To make the motor run up to full speed the load would have to be reduced to a value less than that of the minimum occurring between A and B. The kinks are due to irregularities (such as teeth) in the machine which accentuate the effect of space harmonics in the flux density distribution. Curve (*c*) in Fig. 13.14 shows the slip/torque characteristic due to the 7th space harmonic, which rotates at a speed of $n_0/7$ in the same direction as the fundamental. When the rotor speed n_r is less than $n_0/7$ this space harmonic produces a motoring torque, but when n_r is greater than $n_0/7$ it produces a generating torque. When this torque/slip characteristic is added to the torque/slip characteristic due to the fundamental (curve (*d*)) a kink in the resultant torque/slip characteristic occurs.

13.9 Stator Current Locus of the Induction Machine

Fig. 13.15(*a*) shows the approximate referred equivalent circuit of the induction machine. The impedance $\boldsymbol{Z_s}$ is

$$\boldsymbol{Z_s} = R_s + \frac{R_r'}{s} + j(X_s + X_r') \tag{13.35}$$

As the slip s varies the locus of $\boldsymbol{Z_s}$ is that of an impedance of fixed reactance $X_s + X_r'$ and of variable resistance. In Fig. 13.15(*b*) W′AW represents the locus of $\boldsymbol{Z_s}$. For any positive slip s, $\mathrm{OP_s}$ represents the impedance $\boldsymbol{Z_s}$ in modulus and phase, OA represents the fixed reactance $X_s + X_r'$, and $\mathrm{AP_s}$ represents the particular resistance $R_s + R_r'/s$. For negative values of slip points such as $\mathrm{P_s}'$ fall to the left of A and represent induction generator action.

As shown in Section 1.7, if the complexor representing the current $\boldsymbol{I_r}'$ in the impedance $\boldsymbol{Z_s}$, is drawn from the origin O and $\boldsymbol{V_s}$ is taken as reference complexor, the extremity of $\boldsymbol{I_r}'$ must lie on a circle of diameter $V_s/X_s + X_r')$. This locus is shown in Fig. 13.15(*b*). OB is the current $V_s/(X_s + X_r')$ when the slip takes up a value such as to make $R_s + R_r'/s = 0$.

For any value of $\boldsymbol{Z_s}$ (such as $\mathrm{OP_s}$ making an angle θ with the first reference axis), the corresponding value of $\boldsymbol{I_r}'$ is found by drawing $\mathrm{OQ_s}$ making an angle $-\theta$ with the first reference axis, as shown in

Fig. 13.15(*b*). OQ$_s$′ is the current complexor corresponding to the impedance OP$_s$′ drawn for a negative slip.

To obtain the total stator current I_s the fixed current I_0 must be added to I_r', whose value varies as s varies. This may be done most

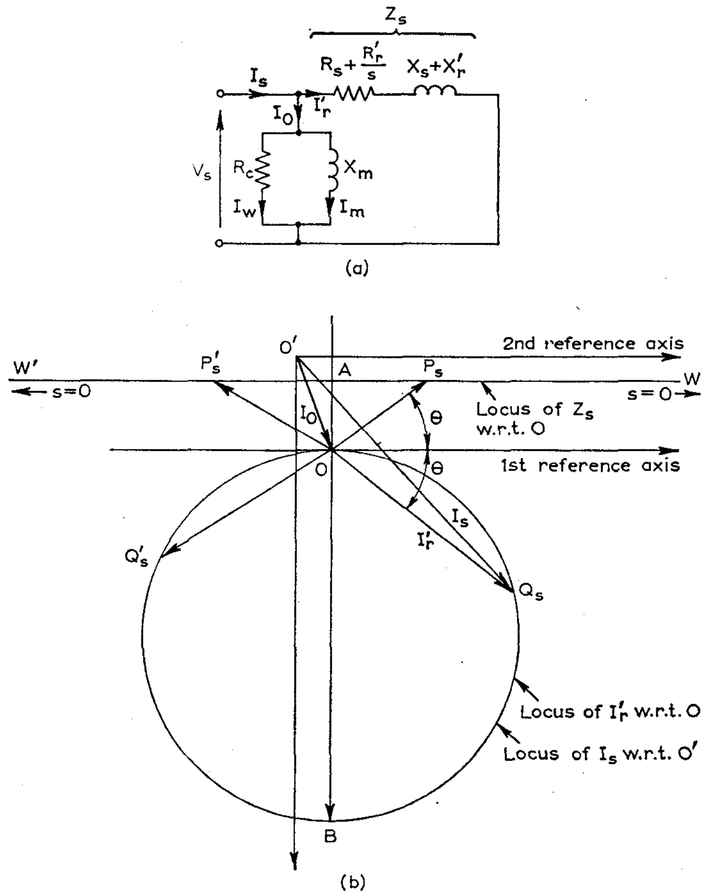


Fig. 13.15 STATOR CURRENT LOCUS OF THE INDUCTION MACHINE

(*a*) Equivalent circuit at any slip, *s*
(*b*) Impedance and current loci

conveniently by shifting the origin of measurement for the current by an amount equal to I_0 (to scale) from O to O′ as shown in Fig. 13.15(*b*). The locus of I_s is then the same circle as the locus of I_r', but the origin of measurement for I_s is O′, while that for I_r' is O.

13.10 Torque and Mechanical Power Lines of the Induction Machine Circle Diagram

Fig. 13.16 again shows the locus of the impedance

$$Z_s = R_s + \frac{R_r'}{s} + j(X_s + X_r')$$

and of the current, I_r', flowing through this impedance. For the

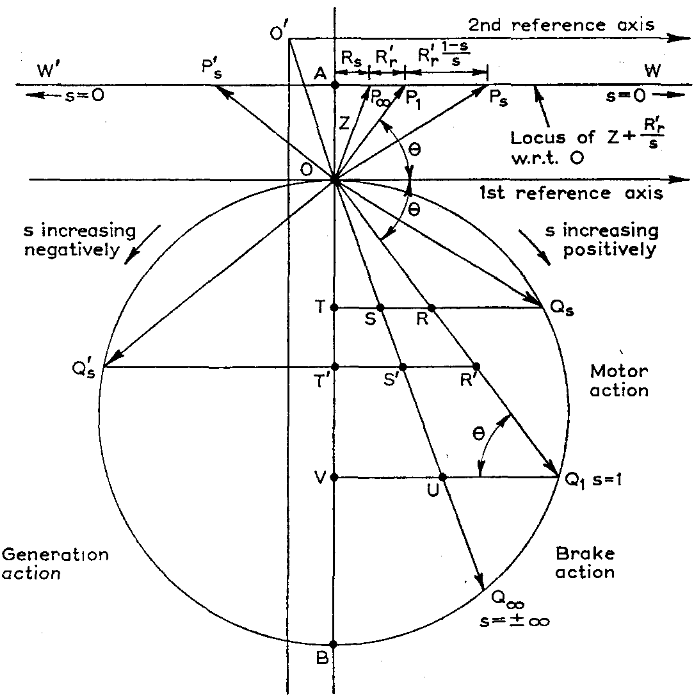


Fig. 13.16 TORQUE AND MECHANICAL-POWER LINES ON INDUCTION-MACHINE CIRCLE DIAGRAM

sake of clarity the size of the complexor O′O representing the no-load current I_0 has been exaggerated.

OP_∞ represents the fixed impedance $Z = R_s + j(X_s + X_r')$ so that $AP_\infty = R_s$ and $OA = X_s + X_r'$. The locus of the impedance $Z + R_r'/s$ as the slip s varies from $+\infty$ to $-\infty$ is the straight line $W'P_\infty W$ parallel to the reference axis. P_s, any point on the locus, lies to the right of P_∞ for positive values of slip (motor action) and to the left of P_∞ for negative values (generator action).

When $s = \pm\infty$, $R_r'/s = 0$, so that any point P_s on the impedance locus is at P_∞.

When $s = 1$, $R_r'/s = R_r' \equiv P_\infty P_1$, and P_s is at P_1.

For any positive slip s,

$$\frac{R_r'}{s} = P_\infty P_s$$

$$P_1P_s = P_\infty P_s - P_\infty P_1 = \frac{R_r'}{s} - R_r' = R_r'\left(\frac{1-s}{s}\right)$$

$$AP_\infty : P_\infty P_1 : P_1P_s = R_s : R_r' : R_r'\left(\frac{1-s}{s}\right) \tag{13.36}$$

$OA = X_s + X_r'$ is the minimum value of the impedance $Z + R_r'/s$ and OB is the corresponding maximum value of current $V_s/(X_s + X_r')$. The circle on OB as diameter is the locus of $V_s/(Z + R_r'/s)$ with respect to O when V_s is taken as reference complexor. Points Q_∞, Q_1, Q_s and Q_s' on this locus corresponding to points P_∞, P_1, P_s and P_s' respectively. The subscripts refer to the values of slip to which the points correspond.

When $s = +\infty$, a condition not practically attainable since it implies that the rotor is driven at infinite speed, any point P_s on the impedance locus is at P_∞, and any point Q_s on the admittance locus is correspondingly at Q_∞. As the value of s decreases, P_s moves to the right to P_1, corresponding to $s = 1$ (standstill), and Q_s moves to Q_1. In the range between P_∞ and P_1, the mode of operation is brake action as explained in Section 13.6. As the slip decreases from unity, P_s moves to the right of P_1 and correspondingly Q_s moves round the circular locus towards O. If the slip were to become zero, P_s would be an infinite distance to the right of P_1 and Q_s would move to O. For the range of slip from +1 to almost zero the mode of operation is motor action.

When $s = -\infty$, a condition also not practically realizable, P_s is at P_∞ and Q_s at Q_∞. As s takes up smaller and smaller negative values, P_s moves to the left of P_∞ and Q_s moves anticlockwise round the circular locus towards O, which it would reach when P_s was an infinite distance to the left of P_∞.

MECHANICAL POWER LINE

In Fig. 13.16 Q_sT is drawn parallel to the reference axis cutting OQ_1 in R, OQ_∞ in S and meeting OB in T. Q_1V is also drawn parallel to the reference axis, cutting OQ_∞ in U and meeting OB in V. Evidently,

$$VU : UQ_1 = AP_\infty : P_\infty P_1 = R_s : R_r' \tag{13.37}$$

Also

$$\cot\,\theta = \frac{\text{VQ}_1}{\text{OV}} = \frac{\text{AP}_1}{\text{OA}} = \frac{R_s + R_r'}{X_s + X_r'} \tag{13.38}$$

$$I_r' = \text{OQ}_s = \frac{V_s}{R_s + \dfrac{R_r'}{s} + j(X_s + X_r')}$$

V_s is the reference complexor, i.e. $V_s = V\underline{/0^\circ}$.

$$\text{OQ}_s = \frac{V_s}{Z_s^2}(R_s + R_r'/s) - j\frac{V_s}{Z_s^2}(X_s + X_r')$$

or

$$\text{OQ}_s = \text{TQ}_s - j\text{OT}$$

$$\text{TR} = \text{OT}\frac{\text{VQ}_1}{\text{OV}} = \frac{V_s}{Z_s^2}(X_s + X_r')\frac{R_s + R_r'}{X_s + X_r'} = \frac{V_s}{Z_s^2}(R_s + R_r')$$

$$\text{RQ}_s = \text{TQ}_s - \text{TR} = \frac{V_s}{Z_s^2}(R_s + R_r'/s) - \frac{V_s}{Z_s^2}(R_s + R_r')$$

$$= \frac{V_s}{Z_s^2}R_r'\left(\frac{1-s}{s}\right)$$

The mechanical power is

$$P_m = 3(I_2')^2R_r'\left(\frac{1-s}{s}\right) \tag{13.22}$$

$$= 3\frac{V_s^2}{Z_s^2}R_r'\left(\frac{1-s}{s}\right)$$

$$= 3V_s\,.\,\text{RQ}_s \tag{13.39}$$

Thus for any point Q_s on the current locus the distance RQ_s parallel to the reference axis is proportional to the mechanical power, so that OQ_1 is called the *mechanical power line*. For generator action $\text{R}'\text{Q}_s'$ represents the mechanical power input.

TORQUE LINE

$$\text{SR} = \text{RQ}_s \times \frac{\text{P}_\infty\text{P}_1}{\text{P}_1\text{P}_s} = \frac{V_s}{Z_s^2}R_r'\left(\frac{1-s}{s}\right) \times \frac{R_r'}{R_r'\left(\dfrac{1-s}{s}\right)} = \frac{V_s}{Z_s^2}R_r'$$

$$\text{Rotor loss, } P_r = 3(I_r')^2R_r' \tag{13.20}$$

$$= 3\frac{V_s^2}{Z_s^2}R_r'$$

i.e.

$$P_r = 3V_s \,.\, \mathrm{SR} \tag{13.40}$$

$$\text{Ideal stator power transfer, } P_0 = 3(I_r')^2 \frac{R_r'}{s} \tag{13.21}$$

$$= 3\frac{V_s^2}{Z_s^2}\left\{R_r' + R_r'\left(\frac{1-s}{s}\right)\right\} = 3V_s(\mathrm{SR} + \mathrm{RQ}_s)$$

i.e.

$$P_0 = 3V_s \,.\, \mathrm{SQ_s} \tag{13.41}$$

$$\text{Torque developed, } T = \frac{3}{2\pi n_0} V_s \,.\, \mathrm{SQ_s} \tag{13.42}$$

Thus for any point $\mathrm{Q_s}$ on the current locus the distance $\mathrm{SQ_s}$ parallel to the reference axis is proportional to the developed torque, so that $\mathrm{OQ_\infty}$ is called the *torque line*. For generator action the torque is $\mathrm{S'Q_s'}$.

13.11 Determination of Equivalent Circuit and Locus Diagram

The current locus may be directly determined by test, and this is usually achieved by means of a no-load test and a short-circuit test.

NO-LOAD TEST

In Fig. 13.16, O′O represents the current I_0, which may be determined by operating the machine in such a way that $I_r' = 0$. Referring to Fig. 13.15(*a*) it will be seen that this condition could be achieved exactly by coupling the machine to a prime mover and rotating the rotor at exactly synchronous speed so that the slip was zero. In these circumstances R_r'/s is infinite and I_r' is zero. In practice it is usually sufficiently accurate to take I_0 as the input current with the machine running unloaded, when the slip will be very small and I_r' negligible compared with I_0. The test is usually performed with full voltage applied to the stator.

Let V_0 = Input voltage per phase on no-load
I_0 = Input current per phase on no-load
P_0 = Input power per phase on no-load

In the approximate conditions of the test it should be noted that the power P_0 is made up of stator core loss, stator I^2R loss, rotor loss, and windage and friction loss. The stator I^2R loss and the rotor loss are usually negligible, but the windage and friction loss may well be comparable to the stator core loss.

$$\cos \phi_0 = \frac{P_0}{V_0 I_0} \tag{13.43}$$

Hence the complexor I_0 of Fig. 13.17 can be drawn.

The quantities in the shunt arm of the equivalent circuit of Fig. 13.15(*a*) are

$$I_w = I_0 \cos \phi_0 \qquad I_m = I_0 \sin \phi_0$$

and

$$R_c = \frac{V_0}{I_w} = \frac{V_0}{I_0 \cos \phi_0} \tag{13.44}$$

$$X_m = \frac{V_0}{I_m} = \frac{V_0}{I_0 \sin \phi_0} \tag{13.45}$$

It should be noted that the power dissipated in R_c, namely $I_w^2 R_c$, represents stator core loss together with windage and friction loss for this method of testing.

LOCKED-ROTOR TEST

In Fig. 13.16, O′Q$_1$ represents the total input current I_s at $s = 1$, and this can be measured with the rotor at standstill. Since the machine produces a torque at standstill it is necessary to lock the rotor to prevent its accelerating. The test is usually performed with reduced voltage.

Let V_{sc} = Input voltage per phase at standstill
I_{sc} = Input current per phase at standstill
P_{sc} = Input power per phase at standstill

The power input P_{sc} includes stator I^2R loss, stator core loss and rotor loss. The stator core loss will be negligible, however, since V_{sc} will be a small fraction of the rated value, whereas I_{sc} will be of the same order as the rated current.

$$\cos \phi_{sc} = \frac{P_{sc}}{V_{sc} I_{sc}} \tag{13.46}$$

The standstill stator current if full voltage were applied would be

$$I_s = I_{sc} \frac{V_s}{V_{sc}} \tag{13.47}$$

Hence the complexor I_s (O′Q$_1$) of Fig. 13.17 may be drawn.

The input current I_s is $I_0 + I_r'$. Since I_0 is normally negligible compared with I_r' at standstill, the input impedance at standstill may be taken as

$$Z_1 = R_s + R_r' + j(X_s + X_r')$$

$$= \frac{V_{sc}}{I_{sc}} \cos \phi_{sc} + j \frac{V_{sc}}{I_{sc}} \sin \phi_{sc}$$

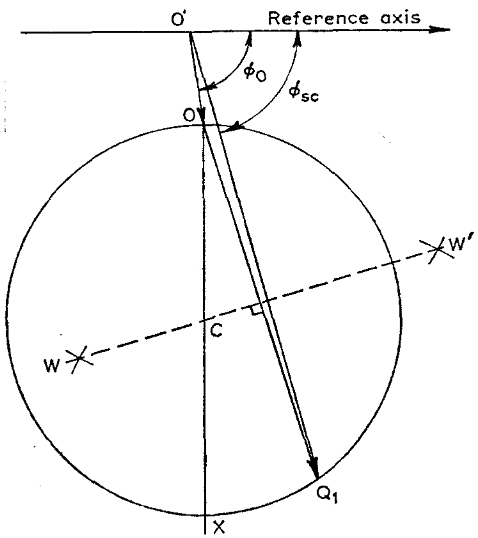


Fig. 13.17 CONSTRUCTION OF THE CURRENT LOCUS DIAGRAM FROM TEST RESULTS

Therefore

$$R_s + R_r' = \frac{V_{sc}}{I_{sc}} \cos \phi_{sc} \tag{13.48}$$

$$X_s + X_r' = \frac{V_{sc}}{I_{sc}} \sin \phi_{sc} \tag{13.49}$$

R_s may be separated from R_r' by performing a d.c. resistance test on the stator winding.

In Fig. 13.17 O′O represents I_0 and O′Q₁ represents the value of I_s at standstill with normal voltage applied. Both O and Q_1 lie on the circular locus, so OQ_1 is joined and represents a chord of the circle. WW′ is the perpendicular bisector of the chord OQ_1. The centre, C, of the circular locus lies along WW′ where it intersects OX, the line drawn through O perpendicular to the reference axis.

EXAMPLE 13.4 A 2·25 kW 440 V 4-pole 50 Hz 3-phase induction motor gave the following test results at the stator terminals.

(i) Standstill test (rotor locked): 142 V, 5 A, 205 W.
(ii) No-load test: 440 V, 2·21 A, 122 W.

The ratio of stator: rotor I^2R loss at standstill is 1·2:1.

From these data construct the circle diagram for the machine, and hence determine

(*a*) Full-load stator current, power factor, slip, speed and efficiency.
(*b*) Pull-out torque, and the slip at which it occurs for both motor and generator action.
(*c*) The maximum mechanical power for both motor and generator action.
(*d*) The starting torque for direct-on-line starting and star/delta starting.

From the standstill test the standstill current is

$$I_s = I_{sc} \times \frac{V_s}{V_{sc}} = 5 \times \frac{440}{142} = 15{\cdot}5\,\text{A}$$

and the power factor is

$$\cos\phi_{sc} = \frac{P_{sc}}{\sqrt{3}V_{sc}I_{sc}} = \frac{205}{\sqrt{3} \times 142 \times 5} = 0{\cdot}167 \text{ lagging}$$

whence

$$\phi_{sc} = 80^\circ\ 30' \text{ lagging}$$

$$I_s = 15{\cdot}5\underline{/-80^\circ\ 30'}\,\text{A}$$

From the no-load test,

$$\cos\phi_0 = \frac{122}{\sqrt{3} \times 440 \times 2{\cdot}21} = 0{\cdot}0725 \text{ lagging}$$

whence

$$\phi_0 = 85^\circ\ 50' \text{ lagging}$$

Open-circuit line current, $I_0 = 2{\cdot}21\underline{/-85^\circ\ 50'}\,\text{A}$

The circle diagram of the machine is shown in Fig. 13.18. The diagram is drawn for line values of current, and the circular locus is obtained by the method explained in Section 13.11.

I_s and I_0 are drawn as the complexors O′O and O′Q₁. The perpendicular bisector of the chord OQ_1 is drawn to cut OX in C, the centre of the circular locus. OX is the line through O perpendicular to the reference axis. A point U is chosen in VQ_1 according to eqn. (13.37) such that

$$\frac{VU}{UQ_1} = \frac{\text{Stator winding loss}}{\text{Rotor winding loss}} = \frac{R_s}{R_r'} = 1{\cdot}2$$

OQ_1 is the mechanical power line, and OU is the torque line. OZ is an axis drawn through O parallel to the reference axis.

(*a*) Full-load power output = 2·25 kW

Active component of current corresponding to full load output

$$= \frac{2{,}250}{\sqrt{3} \times 440} = 2{\cdot}93\,\text{A}$$

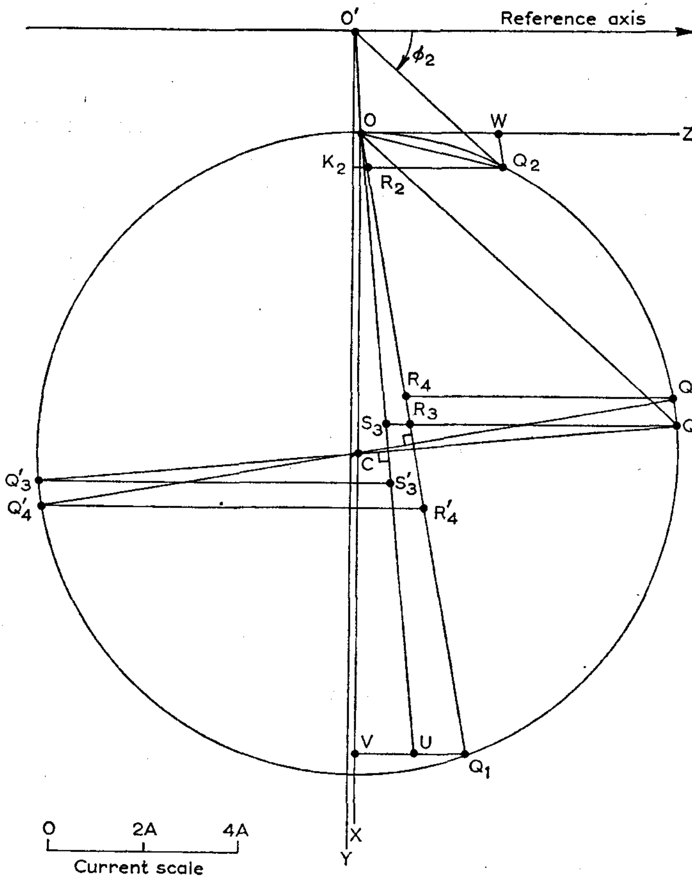


Fig. 13.18

In Fig. 13.18, OW is drawn equal to 2·93 A to scale. WQ_2 is drawn parallel to the mechanical power line to cut the circle in Q_2. OQ_2 is joined Q_2K_2 is drawn parallel to the reference axis to meet O'Y in K_2 and to cut the mechanical power line in R_2.

Full-load stator current $= O'Q_2 \times 2 = 2{\cdot}15 \times 2 = \underline{\underline{4{\cdot}30\,\text{A}}}$

Input power factor $= \cos\phi_2 = \cos 41{\cdot}5 = \underline{\underline{0{\cdot}748 \text{ lagging}}}$

From eqn. (13.14) the slip is

$$s = \frac{P_r}{P_0}$$

Fig. 13.19 shows an enlargement of the area at the origin of Fig. 13.18. In that diagram

$$s = \frac{S_2R_2}{S_2Q_2} = \frac{0{\cdot}12}{5{\cdot}94} = \underline{\underline{0{\cdot}02\,\text{p.u.}}}$$

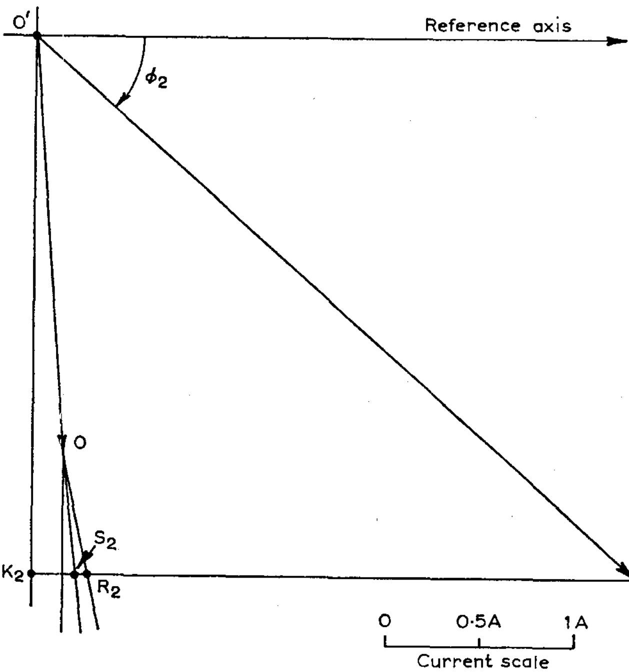


Fig. 13.19 ENLARGEMENT OF AREA AT ORIGIN OF FIG. 13.18

The speed is

$$n_r = (1 - s) = (1 - 0{\cdot}02) \times \tfrac{50}{2} = \underline{\underline{24{\cdot}5\,\text{rev/s}}} \text{ or } \underline{\underline{1{,}470\,\text{rev/min}}}$$

It should be realized that Q_2K_2 is the reference or active component of the input current $O'Q_2$. Of the current Q_2K_2, the component R_2Q_2 is equal to OW, which corresponds to the output. Hence the component R_2K_2 corresponds to losses, and Q_2K_2 corresponds to the electrical power input. Therefore

$$\text{Efficiency} = \frac{Q_2R_2}{Q_2K_2} = \frac{1{\cdot}47}{1{\cdot}60} = \underline{\underline{0{\cdot}92\,\text{p.u.}}}$$

(*b*) The pull-out torque is the maximum torque. Hence two points, Q_3 for motor action and Q_3' for generator action, must be found on the circular locus which make the horizontal distance from the locus to the torque line a maximum. This is most easily accomplished by drawing a line from the centre of the circle, perpendicular to the torque line, to meet the circle at Q_3 and Q_3'. Q_3S_3 and $Q_3'S_3'$ are then drawn in the reference direction to meet the torque line in S_3 and S_3'.

$$\text{Pull-out torque, motoring} = \frac{3V_s \,.\, S_3Q_3}{2\pi n_0} \qquad (13.42)$$

or, when line values of voltage and current are used,

$$\text{Pull-out torque, motoring} = \frac{\sqrt{3} \times 440 \times 3{\cdot}09 \times 2}{2\pi \times 50/2} = \underline{\underline{30\,\text{N-m}}}$$

$$\text{Slip for pull-out torque} = \frac{S_3R_3}{S_3Q_3} = \frac{0{\cdot}24}{3{\cdot}09} = \underline{\underline{0{\cdot}078}}$$

Since S_3Q_3 is now known to represent a torque of 30 N-m, all other torques may be found by simple proportion.

$$\text{Pull-out torque, generating} = 30 \times \frac{S_3'Q_3'}{S_3Q_3} = 30 \times \frac{3{\cdot}73}{3{\cdot}09} = \underline{\underline{36{\cdot}2\,\text{N-m}}}$$

The slip for maximum torque, generating, will be numerically the same as that for maximum torque, motoring, but will be negative.

(*c*) To find the maximum mechanical power, the points Q_4 and Q_4' must be found on the circular locus which make the horizontal distance from the locus to the mechanical power line a maximum. A similar method is used as was used to obtain the maximum torque. A line is drawn from the centre of the circle at right angles to the mechanical power line to meet the circular locus in Q_4 and Q_4'. Q_4R_4 and $Q_4'R_4'$ are then drawn in the reference direction to meet the mechanical power line in R_4 and R_4'.

$$\text{Maximum mechanical power, motoring} = 3V_s \,.\, R_4Q_4 \qquad (13.39)$$

or, when line values of voltages and current are used,

$$\begin{aligned}\text{Maximum mechanical power motoring} &= \sqrt{3} \times 440 \times 2{\cdot}84 \times 2\\ &= \underline{\underline{4{,}320\,\text{W}}}\end{aligned}$$

Otherwise, since it is already known that Q_2R_2 represents 2,240 W (full load), all other mechanical powers may be found by ratio.

$$\begin{aligned}\text{Maximum mechanical power, motoring} &= 2{,}240 \times \frac{Q_4R_4}{Q_2R_2}\\ &= 2{,}240 \times \frac{2{\cdot}84}{1{\cdot}47} = \underline{\underline{4{,}320\,\text{W}}}\end{aligned}$$

$$\begin{aligned}\text{Maximum mechanical power generating} &= 2{,}240 \times \frac{Q_4'R_4'}{Q_2R_2}\\ &= 2{,}240 \times \frac{4{\cdot}03}{1{\cdot}47} = \underline{\underline{6{,}140\,\text{W}}}\end{aligned}$$

(*d*) The starting torque is the torque when $s = 1$ and is therefore represented by UQ_1 on the diagram.

$$\text{Starting torque for direct-on starting} = 30 \times \frac{UQ_1}{S_3Q_3}$$

$$= 30 \times \frac{0{\cdot}55}{3{\cdot}09} = \underline{\underline{5{\cdot}34\,\text{N-m}}}$$

When a star-delta starter is used, the voltage applied to each stator phase is $1/\sqrt{3}$ of the normal running value. Eqn. (13.27) shows that the torque is proportional to the square of the voltage applied to each stator phase. Therefore,

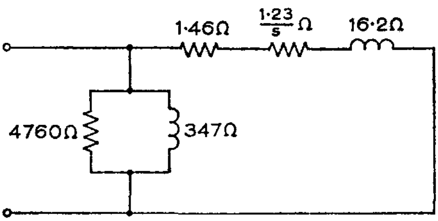


Fig. 13.20 (EXAMPLE 13.4)

when a star-delta starter is used the starting torque will be $(1/\sqrt{3})^2$ or 1/3 of that for direct-on-line starting.

$$\text{Starting torque for star-delta starting} = \frac{5{\cdot}34}{3} = \underline{\underline{1{\cdot}78\,\text{N-m}}}$$

The equivalent circuit calculated from the test results using the method given in Section 13.11 is shown in Fig. 13.20.

EXAMPLE 13.5 In the approximate equivalent circuit of one phase of a 3-phase mesh-connected induction motor shown in Fig. 13.21, $V_s = 415\text{V}$, $R_c = 250\Omega$,

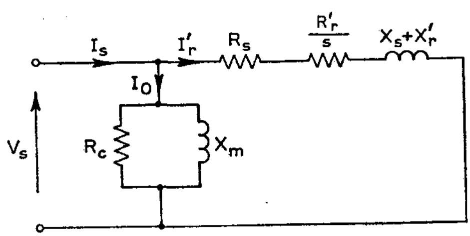


Fig. 13.21 (EXAMPLE 13.5)

$R_s = 0{\cdot}1\Omega$, $R_r' = 0{\cdot}2\Omega$, $X_m = 25\Omega$, and $X_s + X_r' = 1{\cdot}0\Omega$. Determine the input current, power factor, output power and efficiency if the full-load slip is 0·03 when the machine is connected to a 3-phase 415V 50Hz supply.

This problem could be solved by finding the input current on no load (when $s \approx 0$) and on short-circuit (when $s = 1$) and then constructing the circle diagram as in Example 13.4. An alternative solution by calculation is given below.

$$Z_s = R_s + \frac{R_r'}{s} + j(X_s + X_r') = 0{\cdot}1 + \frac{0{\cdot}2}{+0{\cdot}03} + j1{\cdot}0 = 6{\cdot}85\underline{/8{\cdot}4^\circ}$$

$$Y_s = \frac{1}{Z_s} = 146 \times 10^{-3}\underline{/-8{\cdot}4^\circ} = 144 \times 10^{-3} - j21{\cdot}3 \times 10^{-3}$$

$$Y_0 = \frac{1}{R_c} - j\frac{1}{X_m} = \frac{1}{250} - j\frac{1}{25} = 5 \times 10^{-3} - j40 \times 10^{-3}$$

Total input admittance, $Y_T = Y_0 + Y_s = 149 \times 10^{-3} - j61{\cdot}3 \times 10^{-3}$

$$= 161 \times 10^{-3}\underline{/-22{\cdot}4^\circ}$$

Input current per phase $= VY_T = 415 \times 161 \times 10^{-3} = \underline{\underline{66{\cdot}8\,\text{A}}}$

Input power factor $= \cos 22{\cdot}4^\circ = \underline{\underline{0{\cdot}925 \text{ lagging}}}$

Input power $= 3 \times 415 \times 66{\cdot}8 \times 0{\cdot}925 = 77{,}100\,\text{W}$

$I_r' = VY_s = 415 \times 146 \times 10^{-3} = 60{\cdot}5\,\text{A}$

Rotor loss, $P_r = 3(I_r')^2R_r'$

Mechanical power output, $P_m = P_r\left(\frac{1-s}{s}\right)$

$$= 3 \times 60{\cdot}5^2 \times 0{\cdot}2 \times \frac{1 - 0{\cdot}03}{0{\cdot}03}$$

$$= \underline{\underline{71{,}100\,\text{W}}}$$

Efficiency $= \frac{71{,}100}{77{,}100} = \underline{\underline{0{\cdot}923\,\text{p.u.}}}$

13.12 Speed Control

In the majority of applications the speed of a driving motor is required to be almost constant, and hence the plain induction motor is very suitable. Sometimes, however, it is desirable to be able to control the speed of the motor, and this may be achieved in three main ways. From eqn. 13.5,

$$n_r = (1 - s)n_0 = (1 - s)\frac{f}{p} \tag{13.50}$$

It follows that the speed n_0, may be varied by varying the number of stator poles or varying the supply frequency. It is rarely that variation of the supply frequency is used.

In wound-rotor machines the speed for a given load torque may be varied by varying the rotor resistance. This (as shown in Fig. 13.22) gives a range of speeds near full speed, four speeds being possible for the same load torque, with a four-position resistance. The disadvantages of this method are the heat lost in the regulating resistor, and the dependence of the speed variation on the load torque.

If two or three different operating speeds are required near the synchronous speeds of 3,000, 1,500, 1,000, 750, 500, etc., rev/min. these may be achieved by having two or more stator windings each having a different number of pole-pairs. The required speed is obtained by switching on the appropriate winding. This is only possible with cage-rotor machines, since wound-rotor machines must have a fixed number of rotor poles. To avoid the added cost of separate windings with different numbers of pole-pairs, pole-changing windings were introduced. In these, by a series or parallel grouping of the coils (achieved by switching), the number

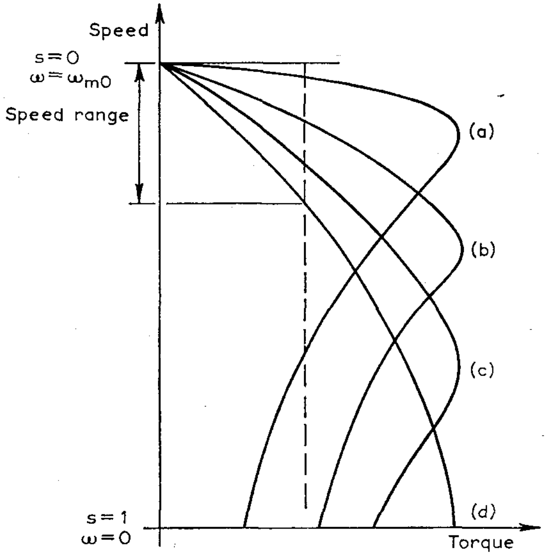


Fig. 13.22 SPEED CONTROL BY VARYING ROTOR RESISTANCE

(*a*) $R_r' = 0{\cdot}15\sqrt{[R_s^2 + (X_s + X_r')^2]}$
(*b*) $R_r' = 0{\cdot}4\sqrt{[R_s^2 + (X_s + X_r')^2]}$
(*c*) $R_r' = 0{\cdot}7\sqrt{[R_s^2 + (X_s + X_r')^2]}$
(*d*) $R_r' = \sqrt{[R_s^2 + (X_s + X_r')^2]}$

of stator poles may be altered. Speed ratios of 2:1 and 1·5:1 may be achieved in this manner.

EXAMPLE 13.6 An 8-pole 50 Hz 3-phase slip-ring induction motor has a total leakage impedance of $(2{\cdot}0 + j5{\cdot}4)\,\Omega$ per phase referred to the stator. The stator resistance per phase is $1{\cdot}1\,\Omega$. When 415 V is applied to the mesh-connected stator winding the voltage between any pair of open-circuited slip rings to which the star-connected rotor winding is connected is 239 V. The motor develops full-load torque with a slip of 0·04 with the slip rings short-circuited. Calculate the approximate speed of the machine if a 3-phase non-inductive resistor of $0{\cdot}5\,\Omega$ per phase is connected in series with the slip rings when full-load torque is applied.

Since the rotor winding is star connected,

$$\text{Induced rotor voltage/phase on open-circuit, } E_r = \frac{239}{\sqrt{3}}\,\text{V}$$

Since the stator winding is mesh connected,

Induced stator voltage/phase on open-circuit, $E = 415\,\text{V}$

Voltage drop due to any magnetizing current is neglected.

$$\text{Effective phase turns ratio} = \frac{415}{239/\sqrt{3}} = 3$$

Rotor resistance/phase referred to stator, $R_{r1}' = 2{\cdot}0 - 1{\cdot}1 = 0{\cdot}9\,\Omega$

Added rotor resistance/phase referred to stator $= 0{\cdot}5 \times 3^2 = 4{\cdot}5\,\Omega$

Total rotor resistance/phase referred to stator, $R_{r2}' = 0{\cdot}9 + 4{\cdot}5 = 5{\cdot}4\,\Omega$

$$\text{Torque developed, } T = \frac{3V_s^2}{2\pi n_0} \frac{s_1 R_{r1}'}{(s_1 R_s + R_{r1}')^2 + s_1^2(X_s + X_r')^2} \tag{13.27}$$

$$= \frac{3V_s^2}{2\pi n_0} \frac{s_2 R_{r2}'}{(s_2 R_s + R_{r2}')^2 + s_2^2(X_s + X_r')^2}$$

where s_1 = slip for full-load torque with slip rings short-circuited, and s_2 = slip for full-load torque with added rotor resistance.

$$\frac{0{\cdot}04 \times 0{\cdot}9}{(0{\cdot}04 \times 1{\cdot}1 + 0{\cdot}9)^2 + 0{\cdot}04^2 \times 5{\cdot}4^2} = \frac{5{\cdot}4 s_2}{(1{\cdot}1 s_2 + 5{\cdot}4)^2 + 5{\cdot}4^2 s_2^2}$$

Therefore

$$s_2^2 - 4{\cdot}22 s_2 + 0{\cdot}96 = 0$$

whence

$$s_2 = 0{\cdot}235 \quad \text{or} \quad 3{\cdot}92$$

The second result refers to brake action and may be neglected here.

$$\text{Speed} = (1 - s) = (1 - 0{\cdot}235) \times \tfrac{50}{4} \times 60 = \underline{\underline{573\,\text{rev/min}}}$$

13.13 Single-phase Induction Machines

The construction of the single-phase induction motor is similar to that of the 3-phase type: a single-phase winding replaces the 3-phase winding. For the same output, the size of the single-phase machine is about 1·5 times that of the corresponding 3-phase machine.

A single-phase current in a single-phase winding produces a pulsating, not a rotating, magnetic field. However, the theory of single-phase motors may be placed on the same basis as that of 3-phase motors by representing a single-phase pulsating m.m.f. by two fields of constant amplitude rotating in opposite directions. This representation is valid so long as the original pulsating field has a sinusoidal distribution round the armature and varies sinusoidally with time.

If the current in the single-phase stator winding is

$$i_s = I_{sm} \sin \omega t$$

then the fundamental stator m.m.f. at any position θ is

$$F'_s = \frac{I_{sm}N_s}{2p}\sin \omega t \sin \theta$$

$$= \frac{I_{sm}N_s}{4p}\{\cos(\omega t - \theta) - \cos(\omega t + \theta)\} \quad (13.51)$$

Thus the pulsating stator m.m.f. may be represented by two oppositely rotating cosinusoidally distributed m.m.f.s each travelling at synchronous speed. Each of these component stator m.m.f. distributions gives rise to corresponding rotor m.m.f. distributions. The resulting instantaneous torque developed has four clearly distinguishable components, which are:

1. A torque due to the interaction of the forward-travelling stator and rotor m.m.f. distributions.
2. A torque due to the interaction of the backward-travelling stator and rotor m.m.f. distributions.
3. A torque due to the forward-travelling stator m.m.f. distribution and the backward-travelling rotor m.m.f. distribution.
4. A torque due to the backward-travelling stator m.m.f. distribution and the forward-travelling rotor m.m.f. distribution.

The first component gives a steady non-pulsating torque acting on the rotor in the forward direction and gives rise to a component slip/torque characteristic of the form obtained from a polyphase induction machine as shown in Fig. 13.23. The second component gives rise to a similar slip/torque characteristic, the torque acting in the opposite, backward direction.

The third and fourth components give rise to torques which pulsate at twice supply frequency and do not contribute to the mean torque. Thus, unlike the polyphase induction machine, the single-phase induction machine has a pulsating component of developed torque. Examination of the component slip/torque curves of Fig. 13.23 reveals that at starting the forward- and backward-directed mean torques are equal, so that the single-phase machine is not self-starting. If the machine is started in either direction it will run up to speed and run stably in that direction.

Following eqn. (13.2), the slip when the rotor runs in the forward direction is

$$s_f = \frac{n_0 - n_r}{n_0} = 1 - \frac{n_r}{n_0} \quad (13.52)$$

The corresponding value of slip for the rotor measured with reference to the backward-travelling m.m.f. is

$$s_b = \frac{-n_0 - n_r}{-n_0} = 1 + \frac{n_r}{n_0} \tag{13.53}$$

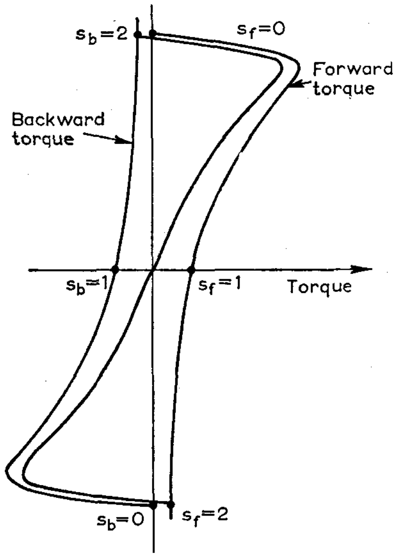


Fig. 13.23 SLIP/TORQUE CHARACTERISTIC OF A SINGLE-PHASE INDUCTION MOTOR

Adding eqns. (13.52) and (13.53),

$$s_f + s_b = 2$$

or

$$s_b = 2 - s_f \tag{13.54}$$

13.14 Equivalent Circuit of Single-phase Induction Machine

An equivalent circuit for the single-phase induction machine may be obtained by considering it as a 3-phase machine with one supply line disconnected as shown in Fig. 13.24. Evidently, from Fig. 13.24,

$$I_R = I \qquad I_Y = -I \qquad I_B = 0$$

$$V_{RS} = \tfrac{1}{2}V \qquad V_{YS} = -\tfrac{1}{2}V$$

Also, $V_{BS} = 0$ since the voltages induced in the blue phase due to its couplings with the red and yellow phases cancel out.

Considering first the symmetrical components of the voltage applied, the zero-sequence voltage is

$$V_{R0} = \tfrac{1}{3}(V_{RS} + V_{YS} + V_{BS}) = 0$$

so that

$$V_{R0} = V_{Y0} = V_{B0} = 0$$

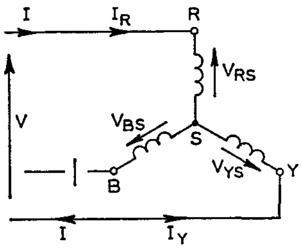


Fig. 13.24 SINGLE-PHASE OPERATION OF A 3-PHASE INDUCTION MACHINE

The applied voltage is

$$\begin{aligned} V &= V_{RS} - V_{YS} \\ &= (V_{R+} + V_{R-} + V_{R0}) - (V_{Y+} + V_{Y-} + V_{Y0}) \\ &= (1 - a^2)V_{R+} + (1 - a)V_{R-} \end{aligned} \tag{13.55}$$

Considering now the symmetrical components of the current, the zero-sequence current is

$$I_{R0} = \tfrac{1}{3}(I_R + I_Y + I_B) = 0$$

The positive phase-sequence current is

$$\begin{aligned} I_{R+} &= \tfrac{1}{3}(I_R + aI_Y + a^2 I_B) \\ &= \tfrac{1}{3}I(1 - a) \end{aligned} \tag{13.56}$$

$$\begin{aligned} I_{R-} &= \tfrac{1}{3}(I_R + a^2 I_Y + aI_B) \\ &= \tfrac{1}{3}I(1 - a^2) \end{aligned} \tag{13.57}$$

Thus from eqns. (13.56) and (13.57),

$$I = \frac{3I_{R+}}{1 - a} = \frac{3I_{R-}}{1 - a^2} \tag{13.58}$$

The total input impedance is

$$Z = \frac{V}{I} = \frac{(1-a^2)V_{R+}}{\dfrac{3I_{R+}}{1-a}} + \frac{(1-a)V_{R-}}{\dfrac{3I_{R-}}{1-a^2}}$$

$$= \frac{V_{R+}}{I_{R+}}\frac{(1-a^2)(1-a)}{3} + \frac{V_{R-}}{I_{R-}}\frac{(1-a)(1-a^2)}{3}$$

$$= \frac{V_{R+}}{I_{R+}} + \frac{V_{R-}}{I_{R-}}$$

i.e.

$$Z = Z_+ + Z_- \tag{13.59}$$

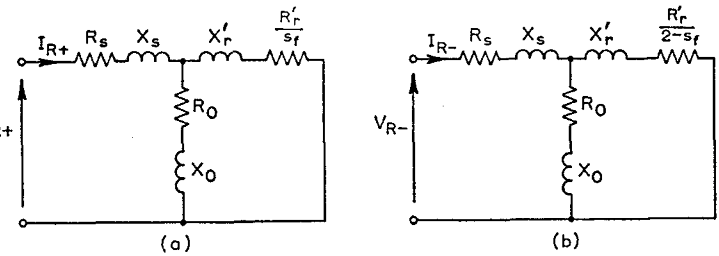


Fig. 13.25 PHASE-SEQUENCE NETWORKS OF THE 3-PHASE INDUCTION MACHINE
(*a*) Positive (*b*) Negative

where Z_+ and Z_- are the positive phase-sequence and negative phase-sequence impedance operators of the 3-phase induction machine respectively.

The positive phase-sequence equivalent circuit of the induction machine, as shown in Fig. 13.25(*a*), is evidently of the same form as the normal equivalent circuit per phase, except that the parallel magnetizing branch has been replaced by its equivalent series circuit. No simplification is obtained by the approximation of showing the magnetizing branch at the input terminals. It should be noted that the parameters in the magnetizing branch will have different values for single-phase excitation than for 3-phase excitation owing to different mutual effects.

The negative phase-sequence equivalent circuit, as shown in Fig. 13.25(*b*), is the equivalent circuit for the machine when a negative phase-sequence 3-phase supply is applied to the stator and the rotor is driven in the opposite direction to the stator field. For any given rotor speed, if s_f is the slip for the positive phase-sequence network, the corresponding slip for the negative phase-sequence network is, according to eqn. (13.54),

$$s_b = 2 - s_f \tag{13.54}$$

Evidently the positive phase-sequence currents give rise to the forward-travelling component wave and the negative phase-sequence currents to the backward-travelling wave, so that the positive phase-sequence network may be used to calculate the forward torque,

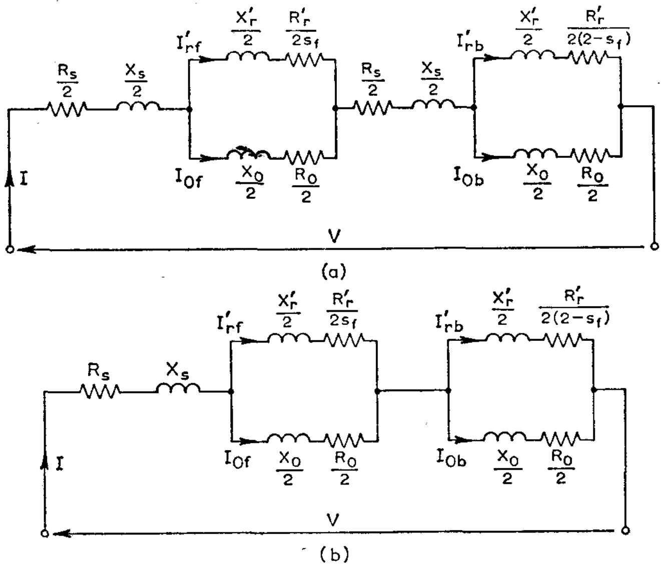


Fig. 13.26 EQUIVALENT CIRCUITS OF THE SINGLE-PHASE INDUCTION MACHINE

and the negative phase-sequence network to predict the backward torque.

Eqn. (13.59) shows that, to obtain the equivalent circuit of the single-phase induction machine, the positive and negative phase-sequence networks must be joined in series. Since one phase of the 3-phase machine corresponds to half of the single-phase winding, the equivalent circuits of the single-phase machine are as shown in Figs. 13.26(*a*) and (*b*).

Following eqn. (13.24) for the 3-phase machine, the forward torque component is

$$T_f = \frac{(I_{rf}')^2 R_r'/2s_f}{2\pi n_0} \tag{13.60}$$

and the backward torque component is

$$T_b = \frac{(I_{rb}')^2 R_r'/2(2 - s_f)}{2\pi n_0} \tag{13.61}$$

EXAMPLE 13.7 A single-phase 230 V 4-pole 50 Hz 0·5 kW induction motor gave the following test results:

Locked rotor test	60 V	1·5 A	Power factor, 0·6 lagging
No load test	230 V	0·535 A	Power factor, 0·174 lagging

Determine the approximate equivalent circuit for the machine. Find also the torque developed, the power output, the input current and power factor when the machine runs with a fractional slip of 0·05.

Assume that the stator and rotor I^2R losses at standstill are equal, and that the rotor leakage reactance referred to the stator and the stator leakage reactance are equal.

For the locked-rotor test, $s_f = 1$. The referred value of rotor impedance, $R_r' + jX_r'$, will be much smaller than the magnetizing impedance, $R_0 + jX_0$; therefore, to a good approximation the equivalent circuit for $s = 1$ is as shown in Fig. 13.27(*a*).

Input impedance with rotor locked, $Z_{sc} = R_s + R_r' + j(X_s + X_r')$

$$= \frac{60\underline{/0^\circ}}{1{\cdot}5\underline{/-\cos^{-1} 0{\cdot}6}} = (24 + j32)\,\Omega$$

$$R_s + R_r' = 24\,\Omega \quad \text{so that} \quad R_s = R_r' = 12\,\Omega$$

$$X_s + X_r' = 32\,\Omega \quad \text{so that} \quad X_s = X_r' = 16\,\Omega$$

For no-load conditions $s_f \to 0$, in which case the magnetizing impedance, $\dfrac{R_0}{2} + j\dfrac{X_0}{2}$, will be much smaller than $\dfrac{R_r'}{2s_f} + j\dfrac{X_r'}{2}$ which tends to infinity when $s_f \to 0$. On the other hand, $\dfrac{R_r'}{2(2 - s_f)} + j\dfrac{X_r'}{2}$, which tends to $\dfrac{R_r'}{4} + \dfrac{X_r'}{2}$ when $s_f \to 0$, will be much smaller than $\dfrac{R_0}{2} + j\dfrac{X_0}{2}$. A good approximation for the equivalent circuit of the single-phase machine on no-load is therefore as shown in Fig. 13.27(*b*).

Input impedance on no-load, $Z_{nl} = R_s + \dfrac{R_0}{2} + \dfrac{R_r'}{4} + j\left(X_s + \dfrac{X_0}{2} + \dfrac{X_r'}{2}\right)$

$$= \frac{230\underline{/0^\circ}}{0{\cdot}535\underline{/-\cos^{-1} 0{\cdot}174}} = (75 + j424)\,\Omega$$

Therefore

$$\frac{R_0}{2} = 75 - R_s - \frac{R_r'}{4} = 75 - 12 - 3 = 60\,\Omega$$

$$\frac{X_0}{2} = 424 - X_s - \frac{X_r'}{2} = 424 - 16 - 8 = 400\,\Omega$$

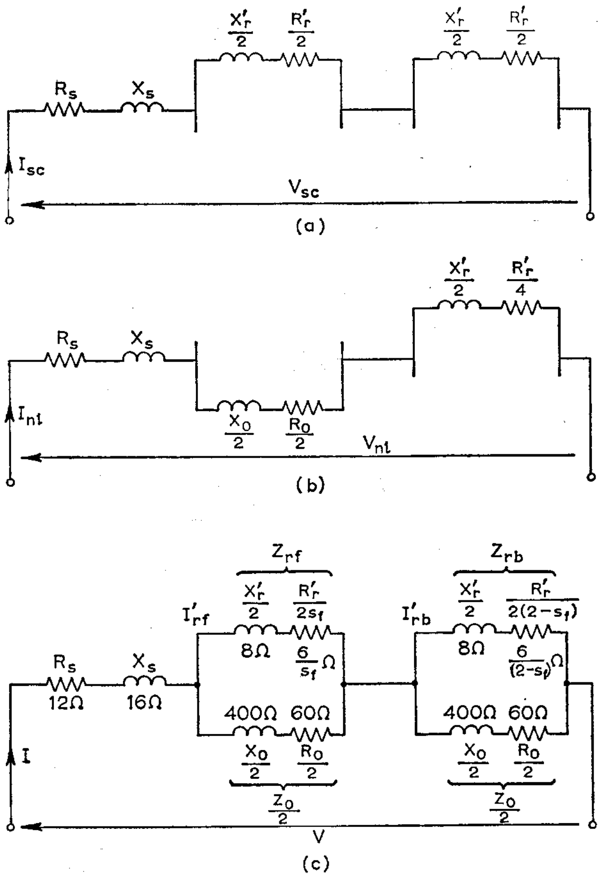


Fig. 13.27

The full equivalent circuit is shown in Fig. 13.27(c). When $s_f = 0{\cdot}05$,

$$\boldsymbol{Z}_{rf} = \frac{6}{\times\ 0{\cdot}05} + j8 = 120 + j8 = 120\underline{/3{\cdot}8^\circ}\,\Omega$$

$$\boldsymbol{Z}_{rb} = \frac{6}{(2 - 0{\cdot}05)} + j8 = 3{\cdot}08 + j8 = 8{\cdot}55\underline{/69^\circ}\,\Omega$$

$$\frac{\boldsymbol{Z}_0}{2} = 60 + j400 = 404\underline{/81{\cdot}5}\,\Omega$$

$$\boldsymbol{Z}_{rf} + \frac{\boldsymbol{Z}_0}{2} = 180 + j408 = 446\underline{/66{\cdot}2^\circ}\,\Omega$$

$$\boldsymbol{Z}_{rb} + \frac{\boldsymbol{Z}_0}{2} = 63{\cdot}1 + j408 = 413\underline{/81{\cdot}3^\circ}\,\Omega$$

The total impedance is

$$Z_T = 12 + j16 + \frac{120\underline{/3\cdot8^\circ} \times 404\underline{/81\cdot5^\circ}}{446\underline{/66\cdot2^\circ}} + \frac{8\cdot55\underline{/69^\circ} \times 404\underline{/81\cdot5^\circ}}{413\underline{/81\cdot3^\circ}}$$

$$= (118 + j59\cdot6)\Omega = 132\underline{/26\cdot9}\Omega$$

$$\text{Input current, } I = \frac{230\underline{/0^\circ}}{132\underline{/26\cdot9^\circ}} = 1\cdot74\underline{/-26\cdot9^\circ}\,\text{A}$$

Input power factor = cos 26·9° = $\underline{\underline{0\cdot892 \text{ lagging}}}$

$$\text{Synchronous speed, } n_0 = \frac{f}{p} = \frac{50}{2} = 25\,\text{rev/s}$$

$$I_{rf}' = I\frac{Z_0/2}{Z_{rf} + Z_0/2} = 1\cdot74 \times \frac{404}{446} = 1\cdot57\,\text{A}$$

$$I_{rb}' = I\frac{Z_0/2}{Z_{rb} + Z_0/2} = 1\cdot74 \times \frac{404}{413} = 1\cdot7\,\text{A}$$

From eqns. (13.60) and (13.61),

$$T_f = \frac{(I_{rf}')^2 R_r'/2s_f}{2\pi n_0} = \frac{1\cdot7^2 \times 120}{2\pi 25} = 1\cdot88\,\text{N-m}$$

$$T_b = \frac{(I_{rb}')^2 R_r'/2(2 - s_f)}{2\pi n_0} = \frac{1\cdot7^2 \times 3\cdot08}{2\pi 25} = 0\cdot0566\,\text{N-m}$$

Net forward torque = 1·88 − 0·0566 = $\underline{\underline{1\cdot82\,\text{N-m}}}$

Power output = $2\pi n_r T$ = 157(1 − 0·05) × 1·82 = $\underline{\underline{0\cdot27\,\text{kW}}}$

13.15 Starting

One of the chief disadvantages of the single-phase induction motor is the fact that special arrangements must be made for starting, so increasing the cost and limiting the motor size which is practical. A second starting winding of short time rating is used. This winding is connected to the single-phase supply through a capacitor or an inductor, producing a phase shift which causes the machine to start as a 2-phase induction motor. A centrifugal switch disconnects the starting winding from the supply when the machine runs up to speed. Fig. 13.28 shows connexion diagrams and a torque/speed curve for this type of motor. It will be observed that the starting direction is reversed with inductance starting.

13.16 Synchronous Induction Motor

A disadvantage of the induction motor, as has been seen, is the fact that it operates at a lagging power factor. It is, however, self-starting. On the other hand, the synchronous motor, which can be

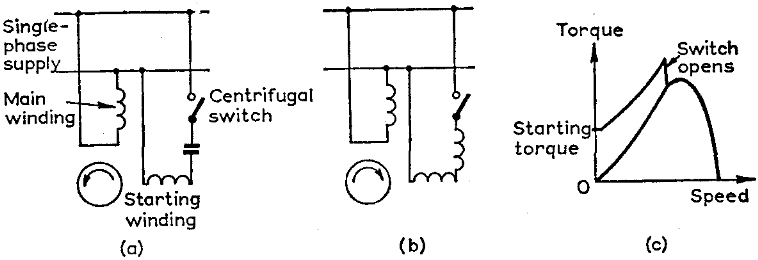


Fig. 13.28 STARTING OF SINGLE-PHASE INDUCTION MOTORS

(*a*) Capacitor start
(*b*) Inductor start
(*c*) Torque/speed characteristics

operated with a leading power factor, has no starting torque. The synchronous induction motor combines the high starting torque of the induction motor with the leading power factor of the synchronous motor. The machine consists essentially of a wound-rotor induction motor, which has a longer air-gap than the normal induction motor to reduce the effect of armature m.m.f. in the machine when it is running as a synchronous motor. It is started by

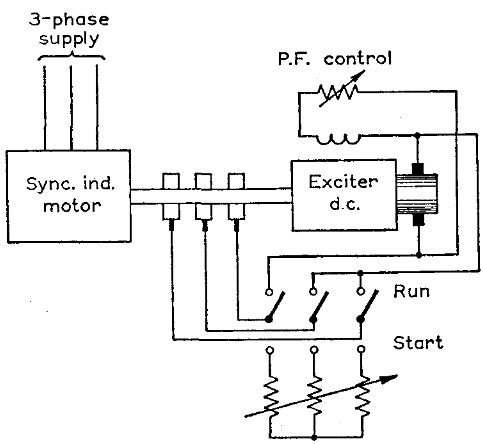


Fig. 13.29 SYNCHRONOUS INDUCTION MOTOR

resistance starting as an induction motor, and when it has run up to speed the starting resistance is disconnected, and direct current from a small exciter on the same shaft as the motor is fed into the rotor (Fig. 13.29). The machine then runs as a synchronous motor, the power factor being varied by controlling the direct current in the rotor.

The most common method of connecting the rotor for direct current is shown in Fig. 13.30, from which it will be seen that one phase carries the total current I_d, while the two other phases carry half each in the opposite direction. Since the rotor runs at synchronous speed, the rotor currents will move with the rotating field, and will produce a synchronously rotating field. Thus the rotor current may be represented on the stator complexor diagram. For the connexion shown the direct currents correspond to the instant when a 3-phase system has positive maximum current in one phase, and half

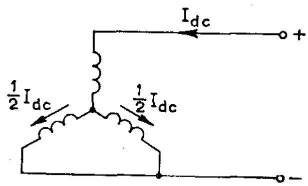


Fig. 13.30 CONNEXION OF A 3-PHASE WINDING FOR DIRECT CURRENT

the negative maximum current in the two other phases. The equivalent r.m.s. rotor current will thus be

$$I_{ac} = \frac{1}{\sqrt{2}} I_{dc} \tag{13.62}$$

This current must be multiplied by the rotor/stator turns ratio, to obtain the equivalent stator current.

13.17 Circle Diagram for a Synchronous Induction Motor

For starting and running up to speed the current locus will be the same as the circle diagram for the induction motor. When the direct current is switched into the rotor, there will be an additional stator current (at supply frequency) due to transformer action. The resultant stator current will be the complexor sum of the no-load current and this additional current, so that, for a constant direct current, the locus of the stator current I_1 (Fig. 13.31), will be a circle, centre D, and of a radius representing the rotor current; i.e.

$$I_1 = \frac{I_d}{\sqrt{2}} \times \frac{\text{Rotor turns}}{\text{Stator turns}}$$

The operating point for a given power output may be found (approximately) by the following procedure.

1. Calculate the active component of the total current which will correspond to the required output power. Mark this component from D parallel to the voltage complexor:

$$DQ' = \frac{\text{Power output (watts)}}{3V_{ph}}$$

(The entire diagram should preferably be drawn for phase values of voltage and current.)

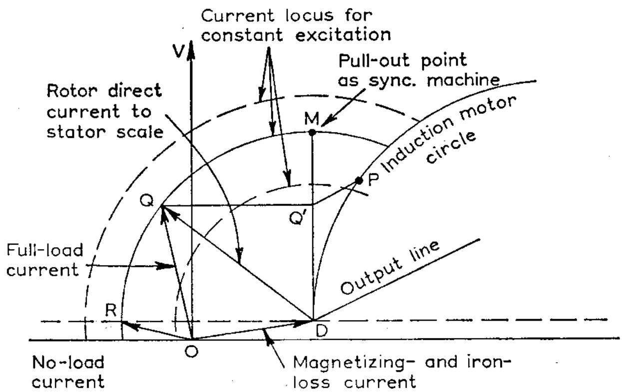


Fig. 13.31 CIRCLE DIAGRAM FOR SYNCHRONOUS INDUCTION MOTOR

2. Draw a line at right angles to DQ′ from Q′ to cut the circle representing the rotor current in use at the point Q. Q is then the operating point and OQ represents the total stator phase current.

It will be seen that for this load and rotor excitation current the machine has a leading power factor. DM represents the maximum possible load which the motor could supply with this particular d.c. excitation without falling out of step as a synchronous machine.

PROBLEMS

13.1 A 4-pole induction motor runs from a 50 Hz supply at 1,450 rev/min. Determine the frequency of the rotor current and the percentage slip.

Ans. 1·67, 3·3 per cent.

13.2 The rotor of a 3-phase 440 V 50 Hz 4-pole induction motor has a resistance of 0·3 Ω/phase and an inductance of 0·008 H/phase. The ratio of stator to rotor turns is 2:1. Determine the voltage between the rotor slip-rings on open-circuit, if the stator is delta connected and the rotor is star-connected. Also calculate the standstill rotor current and the rotor current when running with 3 per cent slip.

Ans. 381 V, 87A; 21·4 A.

13.3 A 6-pole 3-phase 440 V 50 Hz induction motor has an output of 14·9 kW the rotor e.m.f. making 90 cycles per minute. Calculate the motor input power and efficiency if the rotational losses absorb a torque of 16·4 N-m, and the stator losses are 800 W. (*H.N.C.*)

Ans. 17·9 kW, 84 per cent.

13.4 The power input to a 3-phase induction motor is 40 kW. The stator losses total 1 kW, and the mechanical losses are 1·6 kW. Determine the brake output power, the rotor winding loss per phase, and the efficiency if the motor runs with a slip of 3·5 per cent. (*H.N.C.*)

Ans. 36 kW; 445 W; 0·9.

13.5 A 3-phase wound-rotor induction motor has six poles, and operates on a 50 Hz supply. The rotor resistance is 0·1 Ω/phase and the standstill reactance 0·5 Ω/phase. Draw on graph paper the torque/slip curve of the motor. Calculate:

1. The speed for maximum torque.
2. The torque, expressed as a percentage of maximum torque, when the slip is 4 per cent.
3. The value of the external rotor starter resistance to obtain maximum torque on starting. (*H.N.C.*)

Ans. 800 rev/min; 38·5 per cent; 0·4 Ω/phase.

13.6 A 200 V 50 Hz 3-phase induction motor has a 4-pole star-connected stator winding. The rotor resistance and standstill reactance per phase are 0·1 Ω and 0·9 Ω respectively. The ratio of rotor to stator turns is 2:3. Calculate the total torque developed when the slip is 4 per cent. Neglect stator resistance and leakage reactance. (*H.N.C.*)

Ans. 39·8 N-m.

13.7 A 3-phase 50 Hz induction motor with its rotor star connected gives 500 V (r.m.s.) at standstill between slip rings on open circuit. Calculate the current and power factor in each phase of the rotor winding at standstill when joined to a star-connected circuit each limb of which has a resistance of 10 Ω and an inductance of 0·6 H. The resistance per phase of the rotor is 0·2 Ω and its inductance per phase is 0·03 H. Calculate also the current and power factor in each rotor phase when the slip rings are short-circuited and the motor is running with a slip of 4 per cent. Neglect the impedance of the stator. (*L.U.*)

Ans. 1·46 A; 0·0515 lagging; 27·1 A; 0·4 lagging.

13.8 Explain the advantage gained by using a slip-ring rotor instead of a cage rotor for a 3-phase induction motor.

A 3-phase 4-pole induction motor works at 200 V, 50 Hz on full load of 7·5 kW; its speed is 1,440 rev/min. (Frictional losses total 373 W.) Determine approximately (*a*) its speed at 200 V and half load, and (*b*) its speed with an output of 7·5 kW at 190 V, 50 Hz. (*L.U.*)

Ans. 1,470 rev/min; 1,433 rev/min.

13.9 A 3-phase 4-pole induction motor connected to a 50 Hz supply has a constant flux per pole of 0·05 Wb.

The 3-phase star-connected rotor winding has 48 slots and 6 conductors per slot, all the conductors of each phase being connected in series. The resistance of each phase is 0·2Ω and the inductance 0·006 H.

Assuming the flux to remain constant in magnitude, determine the rotor current and the torque in newton-metres (*a*) when the rotor is stationary and short-circuited, and (*b*) when the rotor is running at 5 per cent slip.

Assume that the distribution factor for the rotor winding is 0·96.

(*L.U.*)

Ans. 270 A, 279 N-m; 116 A, 1,030 N-m.

13.10 Explain the action of a 3-phase induction motor (i) at starting, (ii) when running. Deduce expressions for the rotor current, torque and slip when running, assuming the resistance of the rotor circuit to be constant. Sketch a typical slip/torque characteristic from standstill to synchronous speed.

A 6-pole 3-phase induction motor runs at a speed of 960 rev/min when the shaft torque is 136 N-m and the frequency 50 Hz. Calculate the rotor I^2R loss if the friction and windage losses are 150 W. (*L.U.*)

Ans. 574 W.

13.11 Explain the action of a 3-phase induction motor when running loaded. Derive the expression for the torque as a function of the slip assuming the impedance of the stator winding to be negligible.

If the star-connected rotor winding of such a motor has a resistance per phase of 0·1 Ω and a leakage reactance per phase at standstill of 0·9 Ω, calculate: (*a*) the slip at which maximum torque occurs, and (*b*) the additional resistance required to obtain maximum torque at starting. (*L.U.*)

Ans. 0·111; 0·8 Ω.

13.12 A 33·5 kW 4-pole 440 V 3-phase 50 Hz induction motor gave the following results on test:

(*a*) No load—440 V, 22 A, 1,500 W
(*b*) Standstill—220 V, 140 A, 12,000 W

Draw the circle diagram and determine the power factor, load current, and efficiency on full load.

If the rotor winding loss is 45 per cent of the total winding loss, find the slip, and the ratio of starting to full load torque. (*H.N.C.*)

Ans. 0·85, 56 A, 95 per cent; 0·022, 0·69.

13.13 Tests on a 440 V 50 Hz 3-phase 4-pole induction motor showed that on no-load the motor took 22 A at a power factor of 0·2, while at standstill the current was 135 A at a power factor of 0·4 with the applied voltage reduced to 200 V. Determine (*a*) the maximum power output, (*b*) the maximum torque, (*c*) the starting torque on normal voltage. Assume a star-connected stator of resistance 0·164 Ω/phase.

Ans. 75 kW, 585 N-m, 280 N-m.

13.14 A 6-pole 50 Hz 3-phase induction motor develops a useful full-load torque of 163 N-m when the rotor e.m.f. makes 120 complete cycles per minute. Calculate the brake power. If the mechanical torque lost in friction is 17·6 N-m and the stator loss is 750 W, find the winding loss in the rotor circuit, the input to the motor, and its efficiency. (*E.E.P.*)

Ans. 16·4 kW, 770 W, 20 kW, 82 per cent.

13.15 A 75 kW 440 V 3-phase 50 Hz 2-pole synchronous induction motor takes a no-load current of 30 A at a power factor of 0·2 lagging when operating as an induction motor. If the stator-to-rotor turns ratio is 1:2, determine the pull-out torque for a direct current of 99 A. What is the power factor on full load if the direct current is maintained constant? (*H.N.C.*)

Ans. 340 N-m, 0·84 leading.

13.16 Use the test data of Problem 13.12 to determine the parameters of the approximate equivalent circuit of the induction motor.

Ans. $R_c = 388\ \Omega$; $X_m = 35{\cdot}2\ \Omega$; $X_s + X_r' = 2{\cdot}65\ \Omega$; $R_s = 0{\cdot}34\ \Omega$ $R_r' = 0{\cdot}278\ \Omega$.

13.17 Tests on a 3-phase 4-pole 50 Hz induction motor established the following values for the parameters of the approximate equivalent circuit (per phase) as shown in Fig. 13.21: $V = 415$ V, $R_c = 1{,}600\,\Omega$, $X_s + X_r' = 5{\cdot}4\,\Omega$, $X_m = 120\,\Omega$, $R_s = 0{\cdot}48\,\Omega$, $R_r' = 0{\cdot}41\,\Omega$. Determine the input current and power factor and the gross torque and mechanical power developed when the slip is 0·05 p.u. The stator is mesh connected.

Ans. 114 A; 0·53 lagging; 260 N-m; 38·8 kW.

13.18 For the 3-phase induction machine whose equivalent circuit parameters are given in Problem 13.17, (*a*) draw the impedance locus, and (*b*) draw the current locus. (Shift the origin of the current locus by I_0 as explained in Section 13.9.)

(*c*) Use the locus diagram to find the input current and power factor at a slip of 0·05. Compare the values obtained with those found in Problem 13.17. (*Hint.* Find the point on the impedance locus corresponding to $R_s + R_r'/0{\cdot}05$.)

13.19 A single-phase 230 V 4-pole 50 Hz 0·4 kW induction motor gave the following test results:

Locked-rotor test	60 V	1 A	36 W
No-load test	230 V	0·35 A	16 W

Determine the approximate equivalent circuit for the machine. Find also the torque developed, the power output, the input current and power factor when the machine runs with a fractional slip of 0·04. Estimate the value of slip when full-load torque is applied. Assume the stator and rotor winding losses at standstill equal and the rotor leakage reactance referred to the stator and the stator leakage reactance equal.

Ans. 1·03 N-m; 155 W; 0·99 A; 0·89 lagging.

13.20 A 3-phase 440 V 4-pole 50 Hz mesh-connected double-cage induction motor has a stator leakage impedance of $(1 + j5)\,\Omega$/phase. The impedance of the inner cage referred to the stator is $(1 + j10)\,\Omega$/phase at standstill, while that of the outer cage is $(5 + j0)\,\Omega$/phase at standstill. Determine (*a*) the starting torque for full stator voltage applied, and (*b*) the gross torque developed when the fractional slip is 0·04.

Ans. 105 N-m; 123 N-m.

13.21 Tests on a 3-phase 440 V 33·5 kW synchronous induction motor gave the following results:

No-load	440 V	50 A	3,400 W
Locked-rotor	220 V	152 A	13,000 W

When running synchronously at full load the excitation is adjusted to give unity power factor.

$$\textit{If} \quad \frac{\text{Stator turns per phase}}{\text{Rotor turns per phase}} = \frac{1}{2}$$

Calculate the d.c. field current required.

For synchronous operation the rotor is connected with one phase joined in series with two in parallel. Draw the locus of stator current for both synchronous and induction running. From the diagram determine the maximum power output and the corresponding input currents and power factors for (*a*) synchronous running at the above excitation, and (*b*) induction running.

Ans. 47·6 A; 51·5 kW, 87·2 A, 0·83 lagging; 77·6 kW, 199 A, 0·67 lagging.

Chapter 14

THE D.C. CROSS-FIELD MACHINE

The d.c. cross-field machine is a two-axis machine consisting in effect of two machines connected in cascade. The armature winding plays a double role, acting as a field winding on one axis and as an output winding on the second axis. The machine as a whole operates as a two-stage d.c. amplifier suitable for use as a relatively fast-acting exciter to control large machines. The general arrangement of d.c. machines is described in Chapter 10 and is not repeated here. To reduce the effects of eddy currents in slowing response times, the entire magnetic circuit of a cross-field machine is often laminated.

14.1 E.M.F. and Torque Equations of a Single-axis D.C. Machine

Let E_A = Average e.m.f. induced in any one parallel path in the armature winding

Φ = Flux per pole

Z = Number of armature conductors

$2p$ = Number of poles

$2a$ = Number of parallel paths in the armature winding

n_r = Speed, rev/s

N = Number of turns in series in any one of the parallel paths in the armature winding

I_A = Armature winding current

T_A = Torque developed by armature

$$E_A = N\frac{\Delta\Phi}{\Delta t} \tag{14.1}$$

where $\Delta\Phi$ is the change in the flux linked with a coil in time Δt.

A coil whose coil sides are at A and A′ in Fig. 14.1 links the flux per pole, Φ. When the coil moves through a pole-pitch from this position it links the flux per pole in the reverse direction so that the

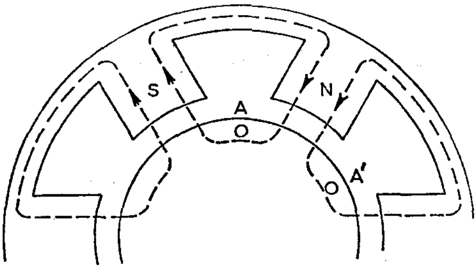


Fig. 14.1 INDUCTION OF E.M.F. IN A SINGLE-AXIS D.C. MACHINE

change in flux linked with coil is 2Φ, i.e.

$$\Delta\Phi = 2\Phi$$

This change takes place in the time required for the coil to move through one pole-pitch, i.e.

$$\Delta t = \frac{1}{2pn_r}$$

Since it requires two conductors to make a turn,

$$N = \frac{1}{2a}\frac{Z}{2}$$

Substituting in eqn. (14.1),

$$E_A = \frac{1}{2a}\frac{Z}{2}\frac{2\Phi}{\frac{1}{2pn_r}} = \Phi Z n_r \frac{p}{a} \tag{14.2}$$

This is the average e.m.f. induced in any one of the parallel paths in the armature winding, and is the voltage measured between the positive and negative brushes when they are positioned to give maximum output voltage and the machine acts as a generator on open-circuit.

Since energy is conserved,

$$2\pi n_r T_A = E_A I_A$$

Therefore the torque developed by the armature is

$$T_A = \frac{1}{2\pi n_r} \Phi Z n_r \frac{p}{a} I_A$$

$$= \frac{\Phi Z I_A}{2\pi} \frac{p}{a} \tag{14.3}$$

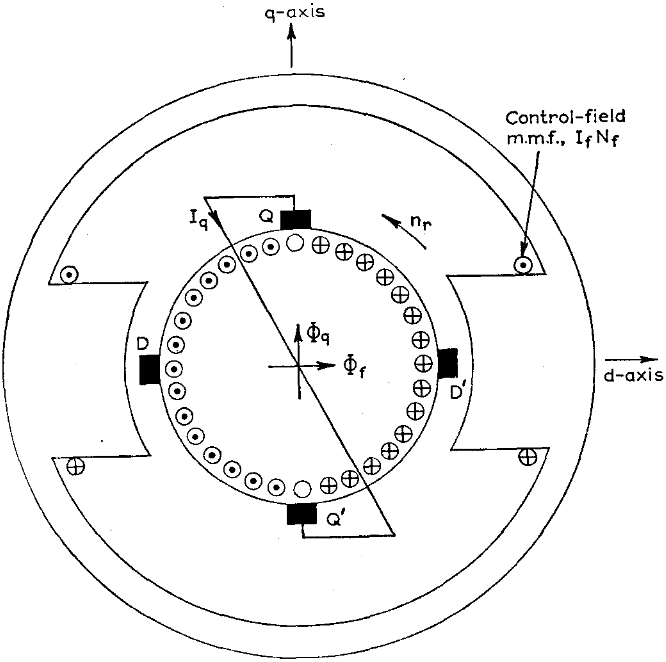


Fig. 14.2 TWO-POLE SINGLE-AXIS D.C. MACHINE

14.2 Performance Equations of a Generalized D.C. Cross-field Generator

Consider a 2-pole single-axis d.c. machine driven by a prime mover at constant speed as shown in Fig. 14.2. It is assumed that the magnetic circuit of the machine is unsaturated and that the flux per pole is proportional to the m.m.f. causing it.

A main field or control winding m.m.f., $I_f N_f$, gives rise to a direct-axis flux

$$\Phi_f = \Lambda_d I_f N_f \tag{14.4}$$

where Λ_d is the permeance of the direct-axis magnetic circuit. The control winding is sometimes called the *variator winding*.

When the field is excited, voltages are induced in the armature conductors in the direction indicated by the dots and crosses in Fig. 14.2 for anticlockwise rotation. A voltage will be obtained at a pair of diametral brushes, and this brush voltage has its greatest value when the brushes are in the position shown in Fig. 14.2. Brushes in the position DD′, on the other hand, have zero generated voltage and would be absent in a conventional single-axis machine.

If the quadrature axis brushes QQ′ are joined together, an armature current I_q will flow which will be large relative to the control-field current, I_f. The distribution of the armature current I_q corresponds to the dots and crosses of Fig. 14.2. This armature current gives rise to an armature flux Φ_q in the direction shown in Fig. 14.2, where

$$\Phi_q = \Lambda_q I_q N_A' \tag{14.5}$$

and Λ_q is the permeance of the quadrature-axis magnetic circuit.

The effective number of turns in the armature winding, N_A', is the number of turns which when multiplied by the total armature flux, Φ_q, gives the total flux linkage. The effective number of turns is less than the actual number, since not all the armature turns link all the armature flux.

The armature flux, Φ_q, resulting from a given armature current, I_q, may be increased by increasing the value of Λ_q. This may be done by providing poles on the quadrature axis as shown in Fig. 14.3.

In a conventional d.c. machine an interpole winding would be provided. This would be excited so as to reduce Φ_q to zero or even to produce a commutating flux in the opposite direction. In a cross-field machine, where a relatively large flux is required, commutation may be critical. To improve commutation the pole may be bifurcated as shown in Fig. 14.4, leaving only a small flux in the commutating zone. High-resistance brushes are employed, and the current to be commutated is reduced by providing a stator quadrature winding g excited so as to aid Φ_q. Thus for a given quadrature axis flux the current I_q required is reduced by increasing the effective number of turns. Thus

$$\Phi_q + \Phi_g = \Lambda_q I_q (N_A' + N_g) \tag{14.6}$$

The stator quadrature winding is sometimes called the *ampliator winding*.

The total quadrature axis flux, $\Phi_q + \Phi_g$, will give rise to induced voltages in the armature conductors. The directions of these e.m.f.s are indicated by the dots and crosses placed beside the armature conductors in Fig. 14.3; the dots and crosses within the armature conductors indicate the induced voltages due to the control flux, Φ_f.

Due to the existence of the quadrature-axis flux, a voltage will occur at the direct-axis brushes DD′ (but the quadrature-axis flux does not, of course, give an e.m.f. at the brushes QQ′).

When the direct-axis brushes are joined to a load resistance, R_L, a direct-axis armature current, I_d, flows. The distribution of this

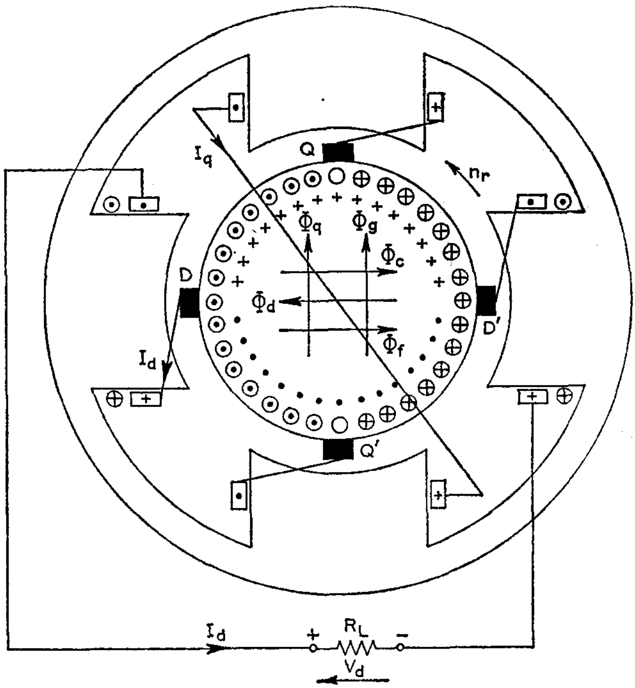


Fig. 14.3 TWO-POLE TWO-AXIS D.C. CROSS-FIELD MACHINE

current in the armature winding is also given by the dots and crosses placed beside the armature conductors. It will be noted that the m.m.f., $I_d N_A'$, due to this load current flowing through the armature winding is in opposition to the input m.m.f., $I_f N_f$, due to the control field. This means that the m.m.f. constitutes negative feedback that is proportional to the output current. In this mode of operation the machine will act as a constant-current generator. The component flux due to the direct-axis armature current is

$$\Phi_d = \Lambda_d I_d N_A' \tag{14.7}$$

The negative feedback due to the armature m.m.f. may be reduced or entirely eliminated by leading the output current I_d through a compensating winding that provides an m.m.f. I_dN_c acting in the opposite direction to that due to the armature. The component flux due to the compensating winding is

$$\Phi_c = \Lambda_d I_c N_c \tag{14.8}$$

The machine is exactly compensated when

$$I_d N_A' = I_d N_c \tag{14.9}$$

The compensating winding is more effective if it is distributed in a similar manner to the armature winding. In some cross-field machines

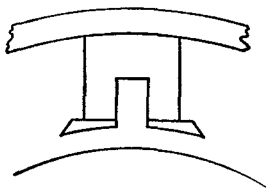


Fig. 14.4 BIFURCATED POLE STRUCTURE

the stator field windings are distributed in slots that are fewer and of a larger size than those on the armature.

The e.m.f. induced in the armature winding due to the component fluxes on the direct axis is

$$E_q = (\Phi_f - \Phi_d + \Phi_c)Zn_r\frac{p}{a} \tag{14.10}$$

The e.m.f. induced in the armature winding due to the component fluxes on the quadrature axis is

$$E_d = (\Phi_q + \Phi_g)Zn_r\frac{p}{a} \tag{14.11}$$

Substituting for the component fluxes in terms of the winding m.m.f.s in eqns. (14.10) and (14.11),

$$E_q = \Lambda_d\{I_fN_f - I_d(N_A' - N_c)\}Zn_r\frac{p}{a} \tag{14.12}$$

$$E_d = \Lambda_q I_q(N_A' + N_g)Zn_r\frac{p}{a} \tag{14.13}$$

Cross-field machines are commonly designed with the permeances of their direct and quadrature axes equal. Assuming this to be so, let

$$\lambda = \Lambda_d Zn_r\frac{p}{a} = \Lambda_q Zn_r\frac{p}{a}$$

Eqns. (14.12) and (14.13) then become

$$E_q = \lambda I_f N_f - \lambda(N_A' - N_c)I_d \tag{14.14}$$

$$E_d = \lambda(N_A' + N_g)I_q \tag{14.15}$$

These e.m.f.s are equal to the sums of the voltage drops in the quadrature-axis and direct-axis circuits respectively. Therefore, for generator action,

$$\lambda I_f N_f - \lambda(N_A' - N_c)I_d = I_q(R_A + R_g) \tag{14.16}$$

$$\lambda I_q(N_A' + N_g) = V_d + I_d(R_A + R_c) \tag{14.17}$$

where R_A = Armature winding resistance
R_g = Stator quadrature-winding resistance
R_c = Compensating-winding resistance

If the directions of the armature currents, I_d and I_q, are considered with respect to the total component fluxes on each of the axes in Fig. 14.3, it will be seen that:

1. The direct-axis armature current I_d produces no resultant torque with the direct-axis component fluxes.
2. The quadrature-axis armature current I_q produces no resultant torque with the quadrature-axis component fluxes.
3. The component torques due to the interaction of I_d and quadrature-axis component fluxes, and due to the interaction of I_q and the direct-axis component fluxes, are additive.

The torque due to the quadrature-axis armature current and the direct-axis flux is

$$T_q = \frac{\Phi_f - \Phi_d + \Phi_c}{2\pi} I_q Z \frac{p}{a} \tag{14.18}$$

and the torque due to the direct-axis armature current and the quadrature-axis flux is

$$T_d = \frac{\Phi_q + \Phi_g}{2\pi} I_d Z \frac{p}{a} \tag{14.19}$$

Substituting for the component fluxes in terms of the winding m.m.f.s,

$$T_q = \frac{\Lambda_d Z \frac{p}{a} I_q}{2\pi} (I_f N_f - I_d N_A' + I_d N_c)$$

$$T_d = \frac{\Lambda_q Z \frac{p}{a} I_d}{2\pi} (I_q N_A' + I_q N_g)$$

The total torque is

$$T = T_q + T_d$$

$$= \frac{\lambda}{2\pi n_r}\{I_q I_f N_f - I_q I_d(N_A' - N_c) + I_q I_d N_A' + I_q I_d N_g\}$$

$$= \frac{\lambda}{2\pi n_r}\{I_q I_f N_f + I_q I_d(N_c + N_g)\} \tag{14.20}$$

14.3 Fully Compensated Cross-field Generator

The cross-field machine is said to be fully compensated when the m.m.f. of the compensating winding is equal and opposite to the armature winding m.m.f. due to the direct-axis armature current, I_d. To ensure exact compensation, the compensating winding is usually overwound (i.e. provided with more turns than necessary), and fine adjustment is then made by connecting an adjustable diverter resistance in parallel with the compensating winding. When the machine is fully compensated, $I_d N_A' = I_d N_c$.

Substituting this condition in eqn. (14.16),

$$\lambda I_f N_f = I_q(R_A + R_g) \tag{14.21}$$

As previously,

$$\lambda I_q(N_A' + N_g) = V_d + I_d(R_A + R_c) \tag{14.17}$$

If the machine is driven at constant speed as a generator and is excited with a constant control-field current, the quadrature-axis armature current, from eqn. (14.21), is

$$I_q = \frac{\lambda I_f N_f}{R_A + R_g} = \text{constant}$$

From eqn. (14.17), therefore,

$$V_d + I_d(R_A + R_c) = \text{constant}$$

In this mode of operation the machine will act as an approximately constant-voltage source, provided that the internal voltage drop, $I_d(R_A + R_c)$, is small compared with the output voltage V_d, a condition which usually obtains since internal losses must be small for the machine to have reasonable efficiency. *Amplidyne generators* are fully compensated cross-field machines.

14.4 Uncompensated Cross-field Generator

When the cross-field machine has no compensating winding, $N_c = 0$ and $R_c = 0$, and eqns. (14.16), (14.17) and (14.20) become

$$\lambda I_f N_f - \lambda I_d N_A' = I_q(R_A + R_g) \tag{14.22}$$

$$\lambda I_q(N_A' + N_g) = V_d + I_d R_A \tag{14.23}$$

$$T = \frac{\lambda}{2\pi n_r}(I_q I_f N_f + I_q I_d N_g) \tag{14.24}$$

If the machine is driven at constant speed as a generator and is excited with a constant control-field current, then

$$\lambda I_f N_f = \text{constant}$$

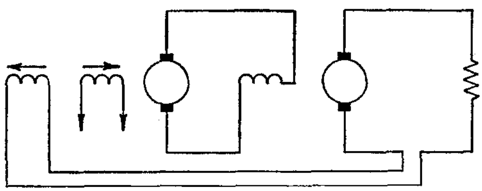


Fig. 14.5 TWO-MACHINE EQUIVALENT OF AN UNCOMPENSATED CROSS-FIELD GENERATOR

Provided that the internal voltage drop, $I_q(R_A + R_g)$, is small compared with either of the generated voltages, then

$$\lambda I_f N_f - \lambda I_d N_A' \rightarrow 0$$

or

$$\lambda I_d N_A' \rightarrow \lambda I_f N_f$$

i.e.

$$I_d \rightarrow \frac{\lambda I_f N_f}{N_A'} \rightarrow \text{constant}$$

In this mode of operation, therefore, the machine acts as an approximately constant-current source, a result already anticipated because of the existence of the negative-feedback term proportional to output current. *Metadyne generators* are uncompensated or undercompensated cross-field generators.

Fig. 14.5 shows the two-machine equivalent of the uncompensated cross-field generator. An external feedback loop is shown providing an m.m.f. proportional to output current at the input in opposition to the separately excited control-field m.m.f.

14.5 Determination of Winding Transfer Constants by Test

In order to predict the operation of a cross-field machine using eqns. (14.16) and (14.17) it is necessary to determine the values of the winding resistances R_A, R_g and R_c, and the values of the terms λN_f, $\lambda N_A'$, λN_c and λN_g. These latter quantities may be called the *winding transfer constants*. They can be determined by performing a series of open-circuit tests on the machine. All the winding interconnexions are removed, and with the machine running at rated speed one winding only is excited with a known current and the appropriate open-circuit brush voltage is measured.

For example, with the control-field winding only excited, the quadrature-axis brush voltage is, from eqn. (14.14),

$$E_q = \lambda I_f N_f \qquad \text{so that} \qquad \lambda N_f = \frac{E_q}{I_f}$$

The other winding transfer constants are found in a similar manner. To determine $\lambda N_A'$, a known current must be fed to one set of brushes and the open-circuit voltage measured at the other set. If the permeances of the direct and quadrature axes are equal, the same result will be obtained whether the current is fed in at the direct or quadrature-axis brushes.

EXAMPLE 14.1 A fully compensated d.c. cross-field generator has the following winding resistances:

Armature winding	$R_A = 0{\cdot}05\,\Omega$
Compensating winding	$R_c = 0{\cdot}05\,\Omega$
Quadrature winding	$R_g = 0{\cdot}15\,\Omega$

Separate excitation open-circuit tests gave the following constants at rated speed:

O.C. quadrature-axis brush voltage	750 V/control-winding ampere
O.C. direct-axis brush voltage	1·5 V/quadrature-winding ampere

The number of stator quadrature-axis turns is equal to the effective number of armature turns. The permeances of the direct and quadrature axes are equal. Magnetic saturation is negligible.

Determine for steady-state operation:

(i) The output voltage on open-circuit and for a load current of 50 A when the machine runs as a fully compensated generator at rated speed with a control current of 3·43 mA.

(ii) The output current on short-circuit and for an output voltage of 50 V when the machine runs without the compensating winding connected at rated speed and with a control current of 73·3 mA.

Sketch the output-voltage/output-current characteristics in each case.

Winding transfer constants are $\lambda N_f = 750\,\text{V/A}$ and $\lambda N_A' = 1{\cdot}5\,\text{V/A}$. Since the stator quadrature-winding turns are equal to the effective armature turns,

$$\lambda N_g = \lambda N_A' = 1{\cdot}5\,\text{V/A}$$

For the fully compensated machine,

$$\lambda N_c = \lambda N_A' = 1{\cdot}5\,\text{V/A}$$

For the uncompensated machine,

$$\lambda N_c = 0$$

(i) Fully compensated machine
Substituting the known constants in eqns. (14.16) and (14.17),

$$750 I_f = 0{\cdot}2 I_q \tag{i}$$

$$3 I_q = V_a + 0{\cdot}1 I_a \tag{ii}$$

Substituting the value $I_f = 3{\cdot}43 \times 10^{-3}$ in eqn. (i),

$$I_q = \frac{750 \times 3{\cdot}43 \times 10^{-3}}{0{\cdot}2} = 12{\cdot}9\,\text{A}$$

When the machine is open-circuited, $I_a = 0$ and from eqn. (ii),

$$V_a = 3 I_q = 3 \times 12{\cdot}9 = \underline{\underline{38{\cdot}6\,\text{V}}}$$

When the load current is 50 A, from eqn. (ii),

$$V_a = 3 I_q - 0{\cdot}1 I_a = (3 \times 12{\cdot}9) - (0{\cdot}1 \times 50) = \underline{\underline{33{\cdot}6\,\text{V}}}$$

(ii) Uncompensated machine
Substituting the known constants in eqns. (14.16) and (14.17),

$$750 I_f - 1{\cdot}5 I_a = 0{\cdot}2 I_q \tag{iii}$$

$$3 I_q = V_a + 0{\cdot}05 I_a \tag{iv}$$

Substituting the value $I_f = 73{\cdot}3 \times 10^{-3}$ in eqn. (iii),

$$55 - 1{\cdot}5 I_a = 0{\cdot}2 I_q \tag{v}$$

When the machine is short-circuited, $V_a = 0$. From eqn. (iv),

$$I_q = \frac{0{\cdot}05 I_a}{3}$$

Substituting for I_q in eqn. (v),

$$55 - 1{\cdot}5 I_a = \frac{0{\cdot}2 \times 0{\cdot}05 I_a}{3}$$

so that

$$I_a = \frac{55}{1{\cdot}503} = \underline{\underline{36{\cdot}5\,\text{A}}}$$

When the load voltage $V_d = 50\,\text{V}$, then from eqn. (iv),

$$I_q = \frac{50 - 0{\cdot}05I_d}{3}$$

Substituting for I_q in eqn. (v),

$$55 - 1{\cdot}5I_d = 0{\cdot}2\left(\frac{50 - 0{\cdot}05I_d}{3}\right)$$

This gives

$$I_d = \frac{55 - 3{\cdot}33}{1{\cdot}503} = \underline{\underline{34{\cdot}3\,\text{A}}}$$

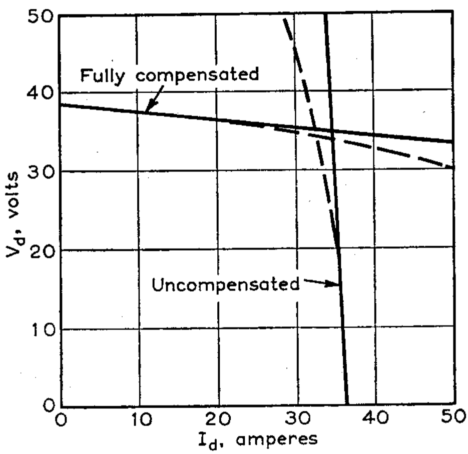


Fig. 14.6 V_d/I_d CHARACTERISTICS OF FULLY COMPENSATED AND UNCOMPENSATED CROSS-FIELD GENERATORS

It will be noted that I_q is constant in the operation of the fully compensated machine, and is approximately proportional to the output voltage in the uncompensated machine. The V_d/I_d characteristic for each mode of operation is shown in Fig. 14.6.

The V_d/I_d characteristics shown in Fig. 14.6 are straight lines since the operating equations were derived assuming that the machine was linear and that, in particular, magnetic saturation was absent.

In practical machines the magnetic circuits are subjected to magnetic saturation, and this causes the characteristics to be non-linear. As magnetic saturation grows the values of the direct-axis and quadrature-axis permeances, Λ_d and Λ_q, will fall, and this causes the V_d/I_d characteristics of the fully compensated and uncompensated machines to be modified in the way indicated by the dotted lines in Fig. 14.6.

PROBLEMS

14.1 A 220 V 0·8 kW fully compensated cross-field generator has the following winding resistances:

Armature winding	$R_A = 0{\cdot}5\,\Omega$
Compensating winding	$R_c = 0{\cdot}5\,\Omega$
Quadrature winding	$R_q = 1{\cdot}5\,\Omega$
Control field winding	$R_f = 340\,\Omega$

Separate excitation open-circuit tests at rated speed gave the following constants:

O.C. quadrature-axis brush voltage	1,000 V/control-winding ampere
O.C. direct-axis brush voltage	250 V/quadrature-winding ampere

The stator quadrature-axis turns are equal to the effective armature turns. The permeances of the direct and quadrature axes are equal. Magnetic saturation is negligible.

Determine for steady-state operation (*a*) the control-winding current required to give rated voltage output on open-circuit, and (*b*) the terminal voltage at rated full-load current and with a control-field current as calculated in (*a*).

Ans. 0·88 mA; 216 V.

14.2 The cross-field generator of Problem 14.1 is operated with only half the compensating winding connected ($R_c = 0{\cdot}25\,\Omega$ and $N_c = 0{\cdot}5N_A'$). Determine, for steady-state operation,

(*a*) The control-winding current required to give rated output current on short-circuit.

(*b*) The output current at rated full-load voltage and with a control-field current as calculated in (*a*).

Ans. 0·455 A; 3·64 A.

14.3 A cross-field machine has an armature resistance of R_A when measured either at the direct-axis or the quadrature-axis brushes. The open-circuit quadrature-axis brush voltage is $\lambda N_A'$ per ampere of the direct-axis brush current. The machine has no stator windings. The permeances of the direct- and quadrature-axis magnetic circuits are equal.

Show that, if a constant direct voltage, V_q, is applied to the quadrature axis, and the direct- and quadrature-axis brush currents are I_d and I_q, the operating equations of the machine are

$$\lambda N_A' I_d = V_q - I_q R_A$$
$$\lambda N_A' I_q = V_d + I_d R_A$$

Assume that magnetic saturation is negligible and that the speed is constant.

Find the value of I_d when $R_A = 0{\cdot}50\,\Omega$, $V_q = 500\,\text{V}$ and $\lambda N_A' = 50\,\text{V/A}$ (*a*) if $R_L = 0$, and (*b*) if $R_L = 60\,\Omega$.

The subject of this question is a simple example of a machine called a *metadyne transformer*.

Ans. 10 A; 9·88 A.

14.4 Starting from eqns. (14.16) and (14.17), show, by eliminating I_q between the equations, that the control-field current of a cross-field generator is related to the output voltage and current by the equation

$$I_f = \frac{R_A + R_g}{\lambda^2 N_f(N_A' + N_g)} V_d + \left\{\frac{N_A'(1-k)}{N_f} + \frac{(R_A + R_c)(R_A + R_g)}{\lambda^2 N_f(N_A' + N_g)}\right\} I_d$$

where the compensation ratio, $k = N_c/N_A'$.

14.5 Using the equation derived in Problem 14.4, show that the control-field current of a cross-field generator is related to the output voltage and current by the equation

$$I_f = c_o V_d + \{c_s(1-k) + c_o(R_A + R_c)\}I_d$$

where

$$c_o = \frac{\text{Output voltage on open-circuit}}{\text{Control-field current}} = \frac{R_A + R_g}{\lambda^2 N_F(N_A + N_g)}$$

and

$$c_s = \frac{\text{Output current on short-circuit with no compensation}}{\text{Control-field current}}$$

$$= \frac{I_d}{I_f} = \frac{N_F}{N_A'} \quad \text{when } k = 0$$

For the machine to be efficient $N_A'/N_f \gg c_o(R_A + R_c)$, and this should be assumed.

14.6 Open- and short-circuit tests on a cross-field generator running at rated speed gave the following results:

Open-circuit tests: voltage gain, $V_d/V_f = 25$
Short-circuit test (no compensation): current gain $I_d/I_f = 0{\cdot}95$

The armature winding resistance is 5Ω, the compensating winding resistance is 5Ω and the control-field winding resistance is 100Ω.

Use the equation found in Problem 14.5 to determine the output voltage and current of the generator when it is driven at constant speed with a control-field current of 4 mA and a load resistance of 100Ω, when the machine has a compensation ratio of 0·99.

Ans. 73·5 V; 0·735 A.

Chapter 15

INTERCONNECTED SYSTEMS

Fig. 15.1 is a line diagram of two power stations A and B joined by an interconnector, the interconnector being connected to the busbars

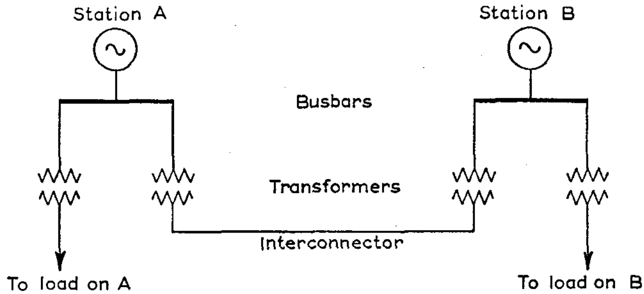


Fig. 15.1 INTERCONNEXION OF POWER STATIONS

of each station through transformers. Each station also has a feeder load connected through a transformer to its busbars.

The power sent across the interconnector will depend, ultimately, on the steam supply to the turbines of each station. For example, if the feeder loads on the busbars of A and B are each 50 MW and the output of the generators on A's busbars is 30 MW, the output of the generators on B's busbars must then be 70 MW, and 20 MW must be transmitted across the interconnector from B to A. As was shown in Chapter 12, the output of the generators depends only on the power supply to their prime movers. Thus the power transmitted from point to point in an interconnected network depends ultimately on the steam supplies to the prime movers. Where more than one path is available between interconnected points, the proportion of power transmitted by each path may be controlled, but the total remains dependent on the load conditions.

The control of the power transmitted over the National Grid in this country is centralized in the control rooms of the Generating Divisions and in the National Control Room. These maintain communication with the generating stations coming under their control and issue instructions to station engineers to increase or reduce station loadings. The control room engineers thus control the frequency and the loading of transmission links in the network.

Apart from the question of the control of the power flow there is the question of voltage regulation. When power is transmitted across the interconnector there will be a voltage drop in the interconnector, the magnitude of which will depend on the impedance of the interconnector and on the power factor at which the power is transmitted. This voltage drop may be accommodated in a number of ways. Assuming that power is being transmitted from B to A (Fig. 15.1) these are as follows.

1. The busbar voltage at B or at A may be so adjusted that the difference in the busbar voltages is equal to the voltage drop in the interconnector and associated transformers. The disadvantage of this method is that it affects the voltages at which the loads connected to the station busbars are supplied.

2. The interconnector transformers may be equipped with on-load tap-changing gear. The voltage drop in the interconnector may then be supplied by adjusting the secondary e.m.f.s of the interconnector transformers, and the busbar voltages may be maintained constant. This method is commonly used where main transformers are, in any case, necessary.

3. A voltage boost in the appropriate direction may be injected into the interconnector either by an induction regulator or by a series boosting transformer. The latter is now of less importance due to the modern practice of incorporating on-load tap-changing gear in main transformers which, in effect, performs the same function as the series boosting transformer.

4. The secondary terminal voltages of the interconnector transformers may be held constant and the voltage drop in the interconnector may be accommodated by adjusting the relative phase of the voltages at the sending and receiving ends of the interconnector by means of a synchronous phase modifier. Synchronous phase

modifiers are used only on transmission links some hundreds of miles in length.

A further use of voltage regulating equipment is to control the division of power between two or more feeders or transmission lines operating in parallel. In the absence of voltage regulating equipment the division of the load between two lines is determined by their respective impedances. This division of the load may be modified by the introduction of a voltage boost in one line.

The control of the power division between lines in parallel by voltage boosting has the important advantage that both lines may be utilized to maximum capacity. It was shown in Chapter 9 that, when lines are operated in parallel, one may become fully loaded before the other has taken up its full load because of disproportionate impedances. A voltage boost of the appropriate magnitude and direction in such under-loaded lines may allow them to take up their full load.

15.1 Tap-changing Transformers

Fig. 15.2(*a*) shows a transformer having variable tappings in the secondary winding. As the position of the tap is varied, the effective

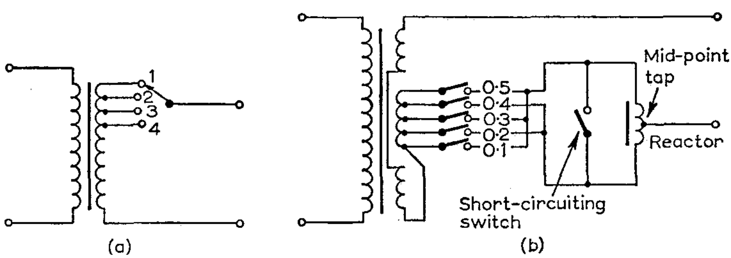


Fig. 15.2 TAP-CHANGING TRANSFORMER

number of secondary turns is varied, and hence the e.m.f. and output voltage of the secondary can be altered.

In supply networks, however, tap-changing has normally to be performed on load (that is, without causing an interruption to supply). The arrangement shown in Fig. 15.2(*a*) is unsuitable for this purpose. Suppose that the tapping is to be altered from position 1 to position 2. If contact with position 1 is broken before contact with position 2 is made, an open-circuit results. If, on the other hand contact with position 2 is made before contact with position 1 is broken, the coils connected between these two tapping points are short-circuited, and will carry damagingly heavy currents. Moreover, in both cases, switching would be accompanied by excessive arcing.

Fig. 15.2(*b*) shows diagrammatically one type of on-load tap-changing transformer. With switch 5 closed, all the secondary turns are in circuit. If the reactor short-circuiting switch is also closed, half the total current flows through each half of the reactor—since the currents in each half of the reactor are in opposition, no resultant flux is set up in the reactor and there is no inductive voltage-drop across it.

Suppose now it is desired to alter the tapping point to position 4. The reactor short-circuiting switch is opened. The load current now flows through one-half of the reactor coil only so that there is a voltage drop across the reactor. Switch 4 is now closed, so that the coils between tapping points 4 and 5 are now connected through the whole reactor winding. A circulating current will flow through this local circuit, but its value will be limited by the reactor. Switch 5 is now opened and the reactor short-circuiting switch is closed, thus completing the operation.

The tapping coils are placed physically in the centre of the transformer limb to avoid unbalanced axial forces acting on the coils, as would arise if they were placed at either end of the limb. Electrically, the tapped coils are at one end of the winding, the practice being to connect them at the earth-potential end.

15.2 Three-phase Induction Regulator

In construction, the 3-phase induction regulator resembles a 3-phase induction motor with a wound rotor. In the induction regulator, the rotor is locked, usually by means of a worm gear, to prevent its revolving under the action of the electromagnetic force operating on it. The position of the rotor winding relative to the stator winding is varied by means of the worm gear.

If the stator winding is connected to a constant-voltage constant-frequency supply, a rotating magnetic field is set up and will induce an e.m.f. in each phase of the rotor winding. The magnitude of the induced rotor e.m.f. per phase is independent of the rotor position, since the e.m.f. depends only on the speed of the rotating field and the strength of the flux, neither of which varies with rotor position. However, variation of the position of the rotor will affect the phase of the induced rotor e.m.f. with respect to that of the applied stator voltage.

Fig. 15.3(*a*) shows the star-connected stator winding of a 3-phase induction regulator with each of the rotor phase windings in series with one line of an interconnector. In Fig. 15.3(*b*) O*a*, O*b*, O*c* represent the input values of the line-to-neutral voltages of the interconnector. The circles drawn at the extremities *a, b, c* of these complexors represent the loci of the rotor phase e.m.f.s as the rotor position is varied with respect to the stator. The complexors *aa′*, *bb′* and *cc′* represent the voltage boosts introduced by the induction

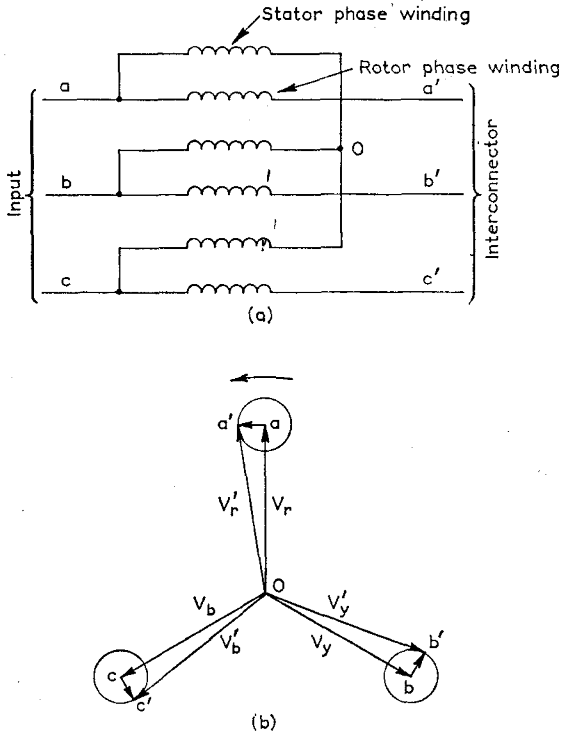


Fig. 15.3 POLYPHASE INDUCTION REGULATOR

regulator when the rotor position is such as to cause these voltage boosts to lead on their respective line-to-neutral voltages by 90°. O*a′*, O*b′* and O*c′* represent the resultant voltages V_r', V_y' and V_b'.

It will at once be seen that the induction regulator has altered the phase of the voltages as well as introducing a voltage boost.

To eliminate this phase displacement, a double polyphase induction regulator is employed, in which two rotors are assembled on a common shaft. The connexion diagram is shown in Fig. 15.4(*a*). The rotor windings of each regulator are connected in series with the interconnector. The stator windings are star-connected, but the phase sequence of one regulator stator is reversed with respect to the other. This reversal has the effect of eliminating any phase displacement in the resultant voltage boost in the interconnector. Thus, when the shaft of the double regulator is displaced, both rotors move by the same angular amount, but if the e.m.f. induced in one leads its former value, then the e.m.f. induced in the other lags by the same amount since the rotating fields in the regulators rotate in

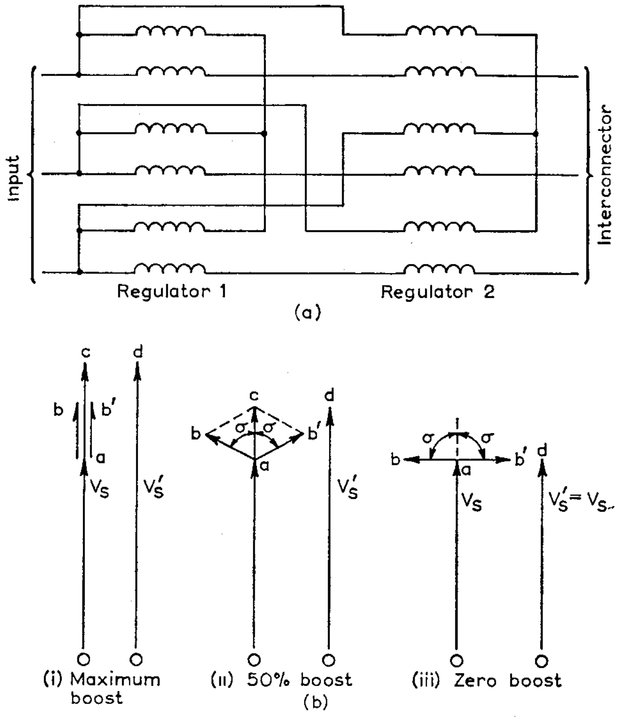


Fig. 15.4 DOUBLE POLYPHASE INDUCTION REGULATOR

opposite directions. Fig. 15.4(*b*) is a complexor diagram for various rotor positions. One phase only is shown for clarity. O*a* represents the unboosted input-end voltage, V_s, *ab* and *ab′* represent the voltage boosts supplied by each rotor, *ac* represents the resultant voltage boost, and O*d* represents the resultant voltage V_s'.

It is often convenient to reverse the functions of the stator and the rotor windings in induction regulators used for boosting. The rotor then carries the primary winding. This has the advantage of requiring only three connexions to the rotor instead of six, and the interconnector current flows in the stator instead of the rotor.

15.3 Synchronous Phase Modifier

In Chapter 12, it was shown that variation in the excitation of a synchronous motor alters the power factor at which the machine works. As the excitation of the machine is increased, the power factor passes from a lagging, through unity, to a leading power factor.

Use is made of this characteristic of the synchronous motor to correct the power factors of loads taking a lagging current. When so used the motor always acts with a leading power factor and is often

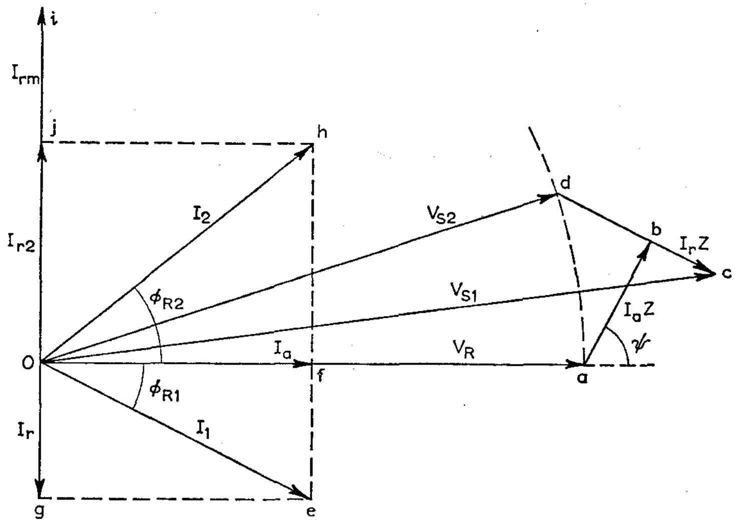


Fig. 15.5 VOLTAGE REGULATION BY SYNCHRONOUS PHASE MODIFIER

called a *synchronous capacitor*. When the synchronous motor is used as a means of controlling the voltage of a transmission line the term *synchronous phase modifier*, or *synchronous compensator* is usually preferred, since, in this application, the machine may be adjusted to take either a leading or a lagging current. The machine is connected in parallel with the load at the receiving end of the line.

The action of the synchronous phase modifier in controlling the voltage of a transmission line is best understood by reference to the complexor diagram shown in Fig. 15.5. For simplicity the diagram is that of a short line where the effects of capacitance are neglected, but it should be understood that this method of control is mostly applied to long lines where, with other methods of control, the voltage drop along the line would be excessive.

In Fig. 15.5, O*a* represents the receiving-end voltage V_R, and O*e* represents the receiving-end current, I_1, lagging behind the receiving-end voltage by a phase angle ϕ_{R1}. O*f* and O*g* represent the active and reactive components (I_a and I_r) of current, respectively, *ab* represents the voltage drop I_aZ caused by the active component of current, which leads V_R by the phase angle of the line impedance, ψ ($\tan^{-1} X_L/R$), Z being the line impedance. *bc* represents the voltage drop I_rZ caused by the reactive component of current. *bc* lags I_aZ by 90°, since I_r lags behind I_a by 90° when the load power factor is lagging. In an unregulated line the sending-end voltage, V_{s1}, is the complexor sum of V_R, I_aZ and I_rZ.

Suppose now that the sending-end and receiving-end voltages are to be held constant at the same value; then the extremity of O*d* representing the new value of the sending-end voltage V_{s2} must be at some point along the arc *ad*, whose centre is O and radius is OD. Moreover, if the same power is to be sent along the line as previously, the I_aZ drop will remain the same since the active component of current must remain the same. However, if the excitation of a synchronous phase modifier connected to the receiving end is adjusted so that it takes a leading current—the current will lead by almost 90° since the modifier works on no-load—then as this leading current is increased the lagging reactive current drawn along the line will be reduced and the voltage drop I_rZ will be reduced. The extremity of the complexor representing the sending-end voltage will move along the line *cb* towards *b*. When the leading reactive current taken by the modifier is equal to the lagging reactive current of the load, there will be no reactive current drawn along the line and no I_rZ drop, and hence the extremity of the complexor representing the sending-end voltage will be at *b*. If the leading current taken by the modifier is further increased, the overall power factor of the load and the modifier together becomes leading and the extremity of the complexor representing the sending-end voltage lies along the line *bd* between *b* and *d*. Thus if the leading current taken by the modifier is made sufficiently great the sending-end voltage complexor takes up the position O*d*.

The synchronous phase modifier may therefore be used to control the voltage drop of a transmission line. If the sending-end voltage is maintained constant, then on full-load at a lagging power factor (the usual condition) the modifier will be over-excited to take a leading current. The receiving-end voltage will thus increase compared with its value had the line been unregulated. On no-load, on the other hand, the modifier would be under-excited and would take a lagging current in order to offset the voltage rise which occurs at the receiving end of a long unregulated line when the load is removed.

The power which may be sent along a transmission line is limited by either the power loss in the line reaching its permissible maximum value or by the voltage drop along the line reaching the maximum value which can be conveniently dealt with. On long transmission lines it is the voltage drop which limits the power which can be sent. Thus, if synchronous phase modifiers are used to regulate the voltage, more power can be dealt with by the line. Moreover, since voltage drop in the line and associated plant is not the first consideration when synchronous phase modifiers are used to control the voltage, current-limiting reactors may be incorporated in the system to reduce the maximum short-circuit current should a fault occur.

The principal disadvantage, apart from cost, of using synchronous phase modifiers is the possibility of their breaking from synchronism and causing an interruption to the supply.

15.4 Sending-end Voltage

In constant-voltage transmission systems using synchronous phase modifiers, the sending-end and receiving-end voltages are held constant, but they do not necessarily have to be equal. There is, indeed, an advantage in having the sending-end voltage higher than the receiving-end voltage, particularly with short lines, since under such conditions a smaller synchronous phase modifier capacity will satisfactorily regulate the voltage. For example, referring to Fig. 15.5, if the sending-end voltage had been greater than the receiving-end voltage, the reactive voltage drop *cd*, due to the reactive current of the synchronous phase modifier, would have been smaller and a synchronous phase modifier of smaller capacity would have been sufficient.

On longer lines, the capacitive effect tends to cause a voltage rise on light loads and no load, and the synchronous phase modifier has to work with a lagging power factor in order to hold the voltage constant. Thus the longer the line the less is the advantage of having the sending-end voltage higher than the receiving-end voltage.

EXAMPLE 15.1 A 3-phase transmission line has a resistance of $8{\cdot}75\,\Omega$ and an inductive reactance of $15\,\Omega$ per phase. The line supplies a load of 10 MW at 0·8 power factor lagging and 33 kV. Determine the kVA rating of a synchronous phase modifier operating at zero power factor such that the sending-end voltage may be maintained at 33 kV. The effect of the capacitance of the line may be neglected. Determine, also, the maximum power which may be transmitted by the line at these voltages.

The problem is most easily tackled graphically.

The line current drawn by the load is

$$I = \frac{10^7}{\sqrt{3} \times 33 \times 10^3 \times 0.8} \underline{/-\cos^{-1} 0.8}$$

$$= 219\underline{/-36.9^\circ} = (175 - j132)\text{A}$$

Phase impedance $= 8.75 + j15 = 17.4\underline{/59.7}\,\Omega$

The voltage drop caused by the active component of the load current is

$I_aZ = 175 \times 17.4 = 3.04 \times 10^3\text{V}$

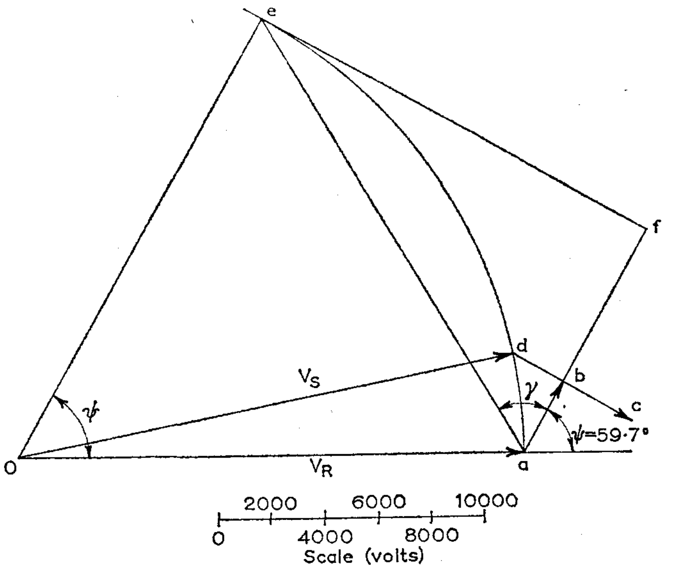


Fig. 15.6

The voltage drop caused by the reactive component of the load current is

$I_rZ = 132 \times 17.4 = 2.3 \times 10^3\text{V}$

Receiving-end voltage (phase value), $V_R = 19.1 \times 10^3\text{V}$

Phase angle of line impedance, $\psi = \tan^{-1}\dfrac{X_L}{R} = \tan^{-1}\dfrac{15}{8.75} = 59.7^\circ$

The graphical construction is shown in Fig. 15.6.

O*a* is drawn to represent V_R to a suitable scale.

ab, equivalent in length to 3.04×10^3V to scale, is drawn making an angle ψ ($=59.7^\circ$) with O*a* to represent I_aZ.

bc, equivalent in length to 2.3×10^3V to scale, is drawn lagging behind *ab* by 90° to represent I_rZ.

The line joining O and *c* would then represent the necessary sending-end voltage in an unregulated line. With centre O and radius O*a* (equivalent in length to $19{\cdot}1 \times 10^3$V) an arc *ae* is drawn. The extremity of the complexor representing the sending-end voltage must lie on this arc. *cb* is now produced until it cuts arc *ae* in *d*. O*d* represents the sending-end voltage.

cd represents the voltage drop due to the reactive current of the modifier alone.

It will be observed that the resultant power factor is leading and the resultant voltage drop, *bd*, due to the reactive current leads the voltage drop, I_aZ, due to the active component of current by 90°.

If I_r' is the reactive current of the synchronous phase modifier,

$$I_r'Z = dc = 4{\cdot}46 \times 10^3\,\text{V to scale}$$

Therefore

$$I_r' = \frac{4{,}460}{17{\cdot}4} = 257\,\text{A}$$

$$\text{Reactive MVA drawn by modifier} = 3 \times \frac{V_R I_r'}{10^6}$$

$$= \frac{3 \times 19{\cdot}1 \times 10^3 \times 257}{10^6}$$

$$= \underline{\underline{14{\cdot}6\,\text{MVA}}}$$

To find the maximum power which can be sent along the line it is necessary to determine the maximum value which the active component of current, I_a, may have. This may be done from the diagram by determining the maximum value of the voltage drop I_aZ.

With centre O and radius O*a* (equal in magnitude to the sending-end voltage V_S), an arc, *ae*, is described. O*e* is drawn parallel to *ab* to cut this arc in *e*. From *e* a perpendicular to *ab* is drawn meeting *ab* in *f*. *af* is the maximum value which voltage drop I_aZ may have; *fe* is the corresponding voltage drop caused by the reactive component of current.

The maximum power is transmitted when the phase angle between the receiving-end and sending-end voltages is equal to that of the line impedance. If the sending-end voltage V_S were to lead the receiving-end voltage V_R by more than ψ the power sent would decrease, since the projection of the sending-end voltage on *ab* would decrease. From Fig. 15.6,

$$I_aZ = af = 9{\cdot}6\,\text{kV} \qquad \text{so that} \qquad I_a = \frac{9{\cdot}6 \times 10^3}{17{\cdot}4} = 552\,\text{A}$$

$$\text{Power sent} = \frac{\sqrt{3}VI_a}{10^6} = \frac{\sqrt{3} \times 33{,}000 \times 552}{10^6} = \underline{\underline{31{\cdot}5\,\text{MW}}}$$

EXAMPLE 15.2 A 3-phase 50 Hz 132 kV transmission line 100 km long has the following constants per km: resistance $0{\cdot}2\,\Omega$, inductance 2 mH, capacitance $0{\cdot}015\,\mu$F.

The sending-end voltage is 132 kV and the receiving-end voltage is held constant at 132 kV by means of a synchronous phase modifier. Determine the reactive kVA of the synchronous phase modifier when the load at the receiving end is 50 MW at a power factor of 0·8 lagging.

One method of taking the effect of line capacitance into account is to use a nominal-π equivalent circuit. The advantage of doing this is that the graphical method developed for the short line is still applicable if, instead of the receiving-end current, I_R, the current in the mid-section of the π-circuit, I', is used.

Fig. 15.7(*a*) shows the nominal-π equivalent circuit, and Fig. 15.7(*b*) shows the complexor diagram of this circuit. The voltage drop in the line impedance is shown as the sum of $I_a'Z$ (the voltage drop caused by the active component of I' and $I_r'Z$ (the voltage drop caused by the reactive component of I'). The voltage diagram showing the relationship between V_R and V_S is now similar to that for the short line.

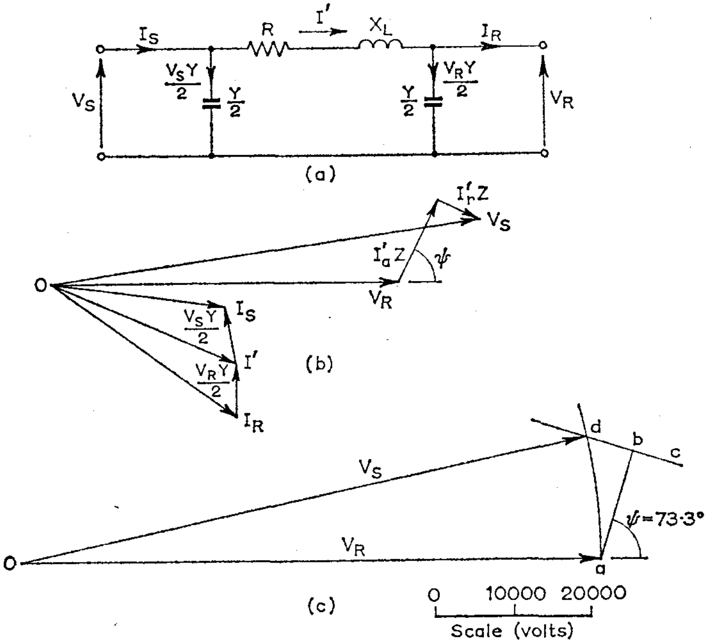


Fig. 15.7

Taking the receiving-end voltage as reference complexor and working with phase values,

$$V_R = \frac{132 \times 10^3}{\sqrt{3}} \underline{/0^\circ} = 76{\cdot}2 \times 10^3\underline{/0^\circ}\,\mathrm{V}$$

$$I_R = \frac{50 \times 10^6}{\sqrt{3} \times 132 \times 10^3 \times 0{\cdot}8} \underline{/-\cos^{-1} 0{\cdot}8} = 273\underline{/-36{\cdot}9^\circ}\,\mathrm{A}$$

$$Z = 0{\cdot}2 \times 100 + j2\pi \times 50 \times 2 \times 10^{-3} \times 100$$

$$= 20 + j62{\cdot}8 = 66{\cdot}2\underline{/72{\cdot}4^\circ}\,\Omega$$

$$\frac{Y}{2} = j\frac{2\pi \times 50 \times 0{\cdot}015 \times 10^{-6} \times 100}{2}$$

$$= j0{\cdot}236 \times 10^{-3} = 0{\cdot}236 \times 10^{-3}\underline{/90^\circ}\,\mathrm{S}$$

Before proceeding to the normal graphical construction, I' should first be calculated with its active and reactive components.

$$I' = I_R + \frac{V_R Y}{2}$$

$$= (273\underline{/-36{\cdot}9^\circ}) + (76{\cdot}2 \times 10^3 \times 0{\cdot}236 \times 10^{-3}\underline{/90^\circ})$$

$$= 219 - j146\,\text{A}$$

$$I_a'Z = 219 \times 66{\cdot}2 = 14{\cdot}5 \times 10^3\,\text{V}$$

$$I_r'Z = 146 \times 66{\cdot}2 = 9{\cdot}65 \times 10^3\,\text{V}$$

The graphical construction shown in Fig. 15.7(*c*) is the same as that described in Example 15.1.

If I_r'' is the reactive current of the modifier, then

$$I_r''Z = cd = 15{\cdot}7 \times 10^3\,\text{V} \qquad \text{so that} \qquad I_r'' = \frac{15{\cdot}7 \times 10^3}{66{\cdot}2} = 236\,\text{A}$$

$$\text{Modifier MVAr} = \frac{\sqrt{3}VI_r''}{10^6} = \frac{\sqrt{3} \times 132 \times 10^3 \times 236}{10^6}$$

$$= \underline{\underline{54\,\text{MVAr}}}$$

15.5 Power/Angle Diagram for a Short Line

In a constant-voltage system of transmission where the sending-end and receiving-end voltages, V_S and V_R, are maintained constant,

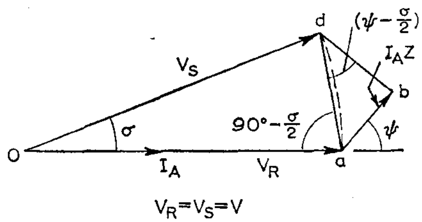


Fig. 15.8 COMPLEXOR DIAGRAM FOR A SHORT LINE

the power transmitted depends on the phase displacement between these voltages. It is possible to derive an expression for the power sent as a function of the phase displacement between V_S and V_R. To simplify the derivation it will be assumed (i) that the line is a short one, and (ii) that $V_S = V_R = V$ (phase values).

Fig. 15.8 shows the complexor diagram in which the receiving-end voltage V_R is taken as the reference. By geometry,

$$\angle adb = (\psi - \sigma/2) \qquad \text{so that} \qquad ab = ad \sin (\psi - \sigma/2)$$

But $ad = 2V \sin \sigma/2$; therefore

$$ab = 2V \sin (\sigma/2) \sin (\psi - \sigma/2) = V\{\cos (\sigma - \psi) - \cos \psi\}$$

Therefore

$$I_A = \frac{V}{Z}\{\cos (\sigma - \psi) - \cos \psi\}$$

and

$$\text{Power transmitted per phase} = VI_A = \frac{V^2}{Z}\{\cos (\sigma - \psi) - \cos \psi\}$$

When $\sigma = 0$ the power transmitted is zero. When $\sigma = \psi$ the power transmitted is a maximum (compare with Example 15.1). The power/angle diagram is shown in Fig. 15.9.

15.6 Stability of Operation of Synchronous Systems

Experience in operating transmission lines with synchronous machinery at both ends has shown that there are definite limits

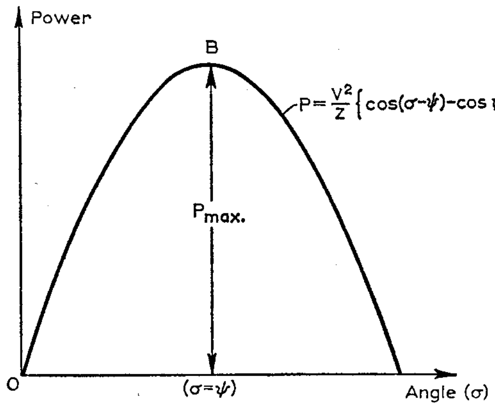


Fig. 15.9 POWER/ANGLE DIAGRAM FOR A SHORT LINE WITH EQUAL SENDING- AND RECEIVING-END VOLTAGES

beyond which operation becomes unstable, resulting in a loss of synchronism between the sending and receiving ends. It is possible to distinguish two limits of stability, a static limit and a dynamic or transient limit.

For given constant values of sending-end and receiving-end voltages, the load on a transmission line can be gradually increased until a condition is reached corresponding to B in Fig. 15.9. At this point the power transmitted is a maximum and corresponds to an angle of phase difference between the sending-end and receiving-end voltages of ψ ($\tan^{-1} X_L/R$). This point represents the static limit of stability (i.e. for a gradually applied load), and any attempt to impose further load on the line results in loss of synchronism between the ends of the line.

Short lines, where constant-voltage transmission using synchronous phase modifiers is not used, cannot be operated near the limit of stability, since the voltage drop along the line would become excessive. On long lines, however, where synchronous phase modifiers are used to control the voltage this limit may be approached and becomes of practical importance. Because of the high capital cost of transmission lines, the more power which can be transmitted over a given line the more economical becomes the operation.

The limit of stability of transmission lines is analogous to the limit of stability of a synchronous motor, where, as load is imposed on the machine, the rotor shifts backwards relative to the rotating field of the stator, and the angle of phase difference between the applied voltage and the e.m.f. increases. The limit of stability is reached when the e.m.f. lags behind the applied voltage by an angle equal to $\tan^{-1} X_s/R$.

This analogy may be made use of in examining the transient stability of a transmission line by considering a line loaded by a synchronous motor. Fig. 15.10 shows the power/angle diagram of a transmission line with the power axis scaled in units of the torque output of the synchronous motor at the receiving end of the line. This may be done conveniently since the synchronous motor is a constant-speed machine and consequently the torque output is proportional to the power transmitted.

When a load is suddenly applied to a synchronous motor, the inertia of the rotor prevents it from immediately falling back by the appropriate electrical angle. Suppose that, in Fig. 15.10, a motor is operating stably at a point A, where the torque developed, T_A, is equal to the load torque. If the load torque is suddenly increased to T_B, the angular displacement of the rotor remains momentarily at σ_A so that the electrically developed torque remains at T_A. The difference torque ($T_B - T_A$) must therefore slow down the rotor causing its angular displacement to increase. The rotor, in slowing down, loses an amount of kinetic energy proportional to the area of the triangle DBA (energy = torque × angular displacement), this energy being transferred to the load.

At B the rotor is running below synchronous speed, so that the angular displacement will continue to increase. However, for angular displacements greater than σ_B the generated torque will exceed the load torque and the rotor will be accelerated; it will regain synchronous speed at an angular displacement σ_C. During this time the load torque remains constant at T_B, so that the torque represented by the difference between the curve BC and the line BE must be that which accelerates the rotor. The area BCE is then proportional to the energy stored in the rotor due to its acceleration. The areas ADB and BCE must be equal if the kinetic energy taken from the rotor during deceleration is all to be returned during the accelerating

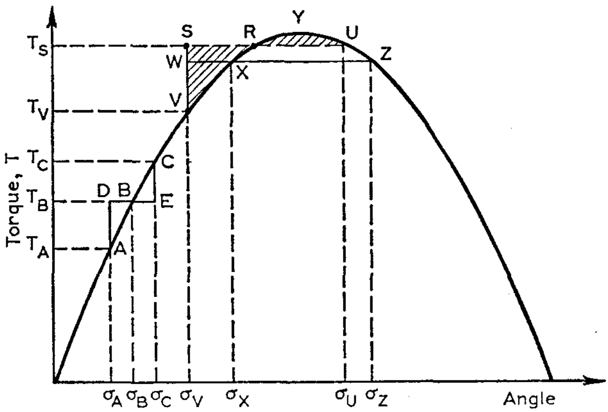


Fig. 15.10 TRANSIENT STABILITY OF A TRANSMISSION LINE

period. This is necessary for the operation to be again stable. This is called the *equal-area criterion for stability.*

At C the electrically developed torque, T_C, exceeds the load torque, so that the machine will continue to accelerate and its speed will rise above synchronous speed. The angular displacement will therefore be reduced, and the electrically developed torque will fall. Beyond B the rotor will experience a synchronizing torque tending to decelerate it. It will fall in speed until it again runs at synchronous speed at A. The whole cycle of events will then be repeated, giving rise to phase swinging or hunting between A and C.

In actual fact the damping which takes place means that the swing of the rotor will becomes less and less, and the machine will eventually run stably at B. The frequency of the oscillation may be determined from eqn. (12.48).

Consider now the motor operating at V with a load torque T_V. If the load torque suddenly increases to T_S, then the shaded area VSR represents the energy which is taken from the rotor as it slows down to accommodate the increased torque. The energy returned to the rotor while it accelerates under the influence of the synchronizing torque is the shaded area RUY. This is less than area VSR, so that the rotor cannot have as much energy returned to it as has been taken from it. It must therefore fall out of step (i.e. lose synchronism).

The maximum load which can be suddenly applied at V is represented by VW, where area VWX is just equal to area XYZ. Note that in this case it is possible for the rotor to swing through an angle which is greater than the maximum angular displacement for static stability. It is, of course, not possible for the operating point to lie beyond the peak point Y on the power/angle diagram.

PROBLEMS

15.1 Explain, with the aid of a complexor diagram, how the power factor of the load influences the voltage drop in a transmission line.

A 3-phase load of 10 MW at 0·8 p.f. lagging is supplied at 33 kV by an overhead line, each conductor of which has a resistance of 2·9 Ω and an inductive reactance of 6·5 Ω. A 3-phase bank of capacitors is connected to the load end of the line so that the voltage at the sending end is equal to that at the load. Calculate the MVA rating of the capacitors. (*L.U.*)

Ans. 12·2 MVA.

15.2 A 3-phase transmission line 25 km in length supplies a load of 10 MW at 0·8 p.f. lagging at a voltage of 33 kV. The resistance and reactance per km per conductor are 0·35 Ω and 0·6 Ω respectively. Neglecting the capacitance of the line, determine the rating of a synchronous capacitor, operating at zero power factor, connected at the load end of the line such that the sending-end voltage may be 33 kV. (*L.U.*)

Ans. 14·7 MVA.

15.3 A 3-phase transmission line is automatically regulated to zero voltage regulation by means of a synchronous phase modifier at the load end. If the full-load output is 50 MW at 0·8 p.f. lagging delivered at 200 kV and the line-to-neutral impedance is $(20 + j60)\,\Omega$, find the input to the synchronous set under these conditions. Deduce any formula employed. (*H.N.C.*)

Ans. 60·3 MVAr.

15.4 A 3-phase overhead line has resistance and reactance of 12 Ω and 40 Ω respectively per phase. The supply voltage is 132 kV and the load-end voltage is maintained constant at 132 kV for all loads by an automatically controlled synchronous phase modifier. Determine the kVAr of the modifier when the load at the receiving end is 120 MW at power factor 0·8 lagging. (*H.N.C.*)

Ans. 145 MVAr.

15.5 A 3-phase transmission line has a resistance of 6 Ω/phase and a reactance of 20 Ω/phase. The sending-end voltage is 66 kV and the voltage at the receiving end is maintained constant by a synchronous phase modifier.

Determine the MVAr of the synchronous phase modifier when the load at the receiving end is 75 MW at 0·8 p.f. lagging, and also the maximum load which can be transmitted over the line, the voltage being 66 kV at both ends. (*L.U.*)

Ans. 96·8 MVAr; 148 MW.

15.6 A 3-phase transmission line has a resistance per phase of 5 Ω and an inductive reactance per phase of 12 Ω, and the line voltage at the receiving end is 33 kV.

(*a*) Determine the voltage at the sending end when the load at the receiving end is 20 MVA at 0·8 p.f. lagging.

(*b*) The voltage at the sending end is maintained constant at 36 kV by means of a synchronous phase modifier at the receiving end, which has the same rating at zero load at the receiving end as for the full load of 16 MW. Determine the power factor of the full-load output and the rating of the synchronous phase modifier. (*L.U.*)

Ans. 40·1 kV; 0·90 lagging; 7·92 MVAr.

15.7 The "constants" per kilometre per conductor of a 150 km 3-phase line are as follows:

Resistance, 0·25 Ω; inductance, 2×10^{-3} H; capacitance to neutral, 0·015 μF.

A balanced 3-phase load of 40 MVA at 0·8 p.f. lagging is connected to the receiving end, and a synchronous capacitor operating at zero power factor leading, is connected to the mid-point of the line. The frequency is 50 Hz.

If the voltage at the load is 120 kV, determine the MVA rating of the synchronous capacitor in order that the voltage at the sending end may be equal in magnitude to that at the mid-point. The nominal-T circuit is to be used for the calculations. (*L.U.*)

Ans. 31·9 MVA.

15.8 A 3-phase 50 Hz transmission line has the following values per phase per km: $R = 0{\cdot}25\,\Omega$; $L = 2{\cdot}0\,\text{mH}$; $C = 0{\cdot}014\,\mu\text{F}$. The line is 50 km long, the voltage at the receiving end is 132 kV and the power delivered is 80 MVA at 0·8 power factor lagging.

If the voltage at the sending end is maintained at 140 kV by a synchronous phase modifier at the receiving end, determine the kVAr of this machine (i) with no load, (ii) with full load at the receiving end. (*L.U.*)

Ans. 30·4 MVAr lagging; 42·6 MVAr leading.

15.9 A 3-phase transmission line has a resistance of 10 Ω per phase and a reactance of 30 Ω per phase.

Determine the maximum power which could be delivered if 132 kV were maintained at each end.

Derive a curve showing the relationship between the power delivered and the angle between the voltage at the sending and receiving ends, and explain how this curve could be used to determine the maximum additional load which could be suddenly switched on without loss of stability if the line were already carrying, say, 50 MW. (*L.U.*)

Ans. 380 MW.

15.10 Develop an expression connecting the power received with the angle between the sending-end and receiving-end voltages of a regulated transmission line in terms of these voltages and the line constants, ignoring capacitance between lines.

Sketch the curve which this equation represents and use it to describe briefly what is meant by (*a*) static, and (*b*) transient stability.

Find the maximum power which can be transmitted over the following line:

Impedance per conductor: $(24 + j45)\,\Omega$
Receiving-end voltage: 110 kV
Regulation: Zero (*H.N.C.*)

Ans. 126 MW.

15.11 Describe the necessary conditions under which power can be transmitted between two interconnected power stations.

Two power stations are linked by a 3-phase interconnector the impedance per line of which (including transformers and reactors) is $(10 + j40)\,\Omega$. The busbar voltage of each station is 66 kV. Calculate the angular displacement between the two station voltages in order to transmit 8 MW from one station to the other. (*H.N.C.*)

Ans. 4·8°.

Chapter 16

TRAVELLING VOLTAGE SURGES

A voltage surge in a transmission system consists of a sudden voltage rise at some point in the system, and the transmission of this voltage to other parts of the system, at a velocity which depends on the medium in which the voltage wave is travelling.

The initiation of voltage surges on overhead transmission lines is frequently caused by lightning discharges. The voltage may be induced in the line due to a lightning discharge in the vicinity without the discharge occurring directly to the line. A most severe voltage rise may be caused where the lightning discharge is direct to a line conductor; such direct strokes, however, are not common.

Voltage surges may also be initiated by switching. Such transient disturbances are due to the rapid redistribution of the energy associated with electric and magnetic fields. For example when the current in an inductive circuit is interrupted the energy stored in the magnetic field must be rapidly transferred to the associated electric field and will give rise to a sudden increase in voltage.

Fig. 16.1 shows the waveform of a typical surge voltage. This is, in effect, a graph of the build-up of voltage at a particular point to a base of time. The steepness of the wavefront is of great importance, since the steeper the wavefront the more rapid is the build-up of voltage at any point on the network, and the properties of insulators depend on the rate of rise of voltage. In most cases the build-up is comparatively rapid, being of the order of 1–5 μs.

Surge voltages are usually specified in terms of the rise time and the time to decay to half maximum value. For example, a $1/50\ \mu s$ wave is one which reaches its maximum value in $1\ \mu s$ and decays to

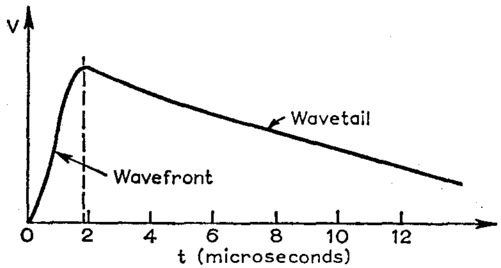


Fig. 16.1 WAVEFORM OF A VOLTAGE SURGE

half its maximum value in $50\ \mu s$. Impulse voltage tests are usually carried out with a wave of this shape.

16.1 Velocity of Propagation of a Surge

In the circuit of Fig. 16.2 the high-voltage source is assumed to give a constant high voltage E, which is applied to one end of the

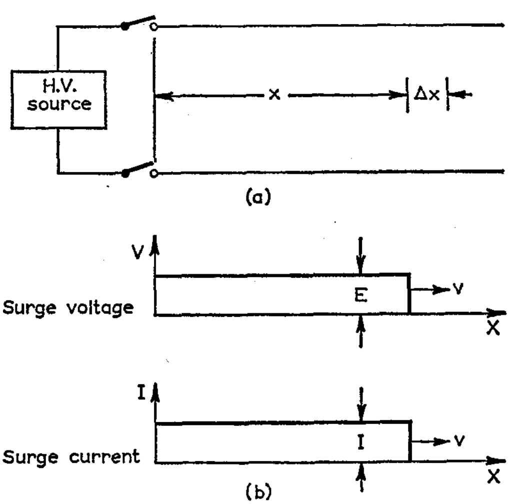


Fig. 16.2 CONSTANT VOLTAGE SURGE

line when the switch is closed. This will simulate a lightning stroke with a rapid rise of voltage and a long tail. It will also represent 50 Hz phenomena on short lines for reasons which will be discussed later. The lines will be assumed loss-free, i.e. no conductor resistance or leakage conductance. This greatly simplifies the theory and errs on the right side, since the effect of losses is to reduce the size of a surge and the rate of rise of voltage. Real surges then will not be quite so severe as those calculated by the present theory.

Let C = Line capacitance per unit length (F/m)
L = Line inductance per unit length (H/m)

When the switch is closed the whole line will not immediately become charged to a voltage E, for raising the voltage of a length of line entails charging the capacitance of that length and this entails a current flow from the source. The current cannot immediately flow due to the inductance of the line and therefore the line voltage cannot immediately rise to the voltage E, though in time it will do so.

Suppose that at an instant t after the closing of the switch a length x of the line has become charged to voltage E, and that those parts of the line beyond a further distance Δx are not yet charged at all. The intermediate region of length Δx will be called the "disturbance" since it is only in this region that the voltage is changing—in front of the disturbance the voltage is zero; behind it the voltage is constant at E. The disturbance will move along the line from the switch with a uniform velocity since the line is uniform.

Let v be the velocity of the disturbance; i.e. the *surge velocity*, or *speed of propagation.*

Since the velocity is uniform equal lengths (v units) of the line will be charged up each second as the disturbance moves along the line. Therefore

$$\begin{aligned}\text{Charge required per second} &= E \times \text{capacitance of length } v\\ &= ECv\end{aligned}$$

i.e.

$$\text{Charging current flowing along line, } I = ECv \qquad (16.1)$$

I is called the *surge current.*

As the surge proceeds along the line a new length v will carry the current I each second, i.e. a length which has an inductance Lv has the current in it changed from zero to I in 1 sec.

$$\begin{aligned}&\text{Potential required to increase current}\\ &\qquad = \text{Inductance} \times \text{Rate of change of current}\\ &\qquad = LvI \text{ volts}\end{aligned}$$

The potential applied to increase the current is E since in front of the disturbance the potential is zero and behind it the potential is E. Therefore

$$E = LvI \tag{16.2}$$

From eqns. (16.1) and (16.2),

$$\text{Surge velocity, } v = \frac{1}{\sqrt{(LC)}} \tag{16.3}$$

Consider, for example, the velocity of propagation in a concentric cable.

$$\text{Inductance/unit length} = \frac{\mu}{2\pi} \log_e \frac{b}{a} \text{ henrys/metre} \tag{7.29}$$

$$\text{Capacitance/unit length} = 2\pi \frac{\epsilon}{\log_e \dfrac{b}{a}} \text{ farads/metre} \tag{7.7}$$

where b and a are the sheath and core radii respectively, and μ and ϵ are the permeability and permittivity of the dielectric material respectively. The internal sheath linkages are not considered since a surge is a high-speed effect equivalent to a high-frequency effect and skin effect will reduce the internal linkages. If the resistance is zero, the depth of penetration is also necessarily zero.

$\mu = \mu_0$ if the dielectric is non-magnetic
$\epsilon = \epsilon_r\epsilon_0$ where ϵ_r is the relative permittivity

Therefore

$$v = \frac{1}{\sqrt{(LC)}} = \frac{1}{\sqrt{(\mu_0\epsilon_0\epsilon_r)}} \tag{16.4}$$

i.e. the velocity is independent of the size and spacing of the conductors. This applies to all configurations of lossless conductors, e.g. a core that is not concentric, or a twin-line system. Substituting numerical values for μ_0 and ϵ_0,

$$v = \frac{1}{\sqrt{\left(4\pi \times 10^{-7} \times \dfrac{\epsilon_r}{36\pi \times 10^9}\right)}} = \frac{3 \times 10^8}{\sqrt{\epsilon_r}} \text{ metres/second} \tag{16.5}$$

This is the velocity of electromagnetic waves in the medium concerned.

The surge velocity is seen to be extremely fast (3×10^8 m/s in air). Even in a cable with a relative permittivity of 9 (fairly high) the velocity is 10^8 m/s. If, say, 5 μs is the time of rise of the surge voltage (length of the wavefront), then the length of the disturbance is

$5 \times 10^{-6} \times 10^8 = 500$ m, which is short for power transmission-line distances (not, of course, for high-frequency transmission-line distances). For power transmission lines, surges are often represented by a "vertical front block" as in Fig. 16.2(*b*) since the length of the disturbance is small.

In $\frac{1}{50}$th second a surge will travel $\frac{1}{50} \times 3 \times 10^8 = 6 \times 10^6$ m (3,730 miles). Thus in $\frac{1}{10}$th of a cycle, for a 50 Hz system, a surge will travel the whole length of a 600 km line (about 400 miles). In this period the voltage of the 50 Hz source will not have greatly changed, so that if a 50 Hz source is suddenly connected to a line the surge may be examined by the present theory taking the surge voltage as the instantaneous voltage when the switch is closed. Naturally the peak voltage should be considered, since this will be at the most dangerous instant at which the switch might be closed.

16.2 Surge Impedance

Dividing and simplifying eqns. (16.1) and (16.2),

$$\frac{E}{I} = \sqrt{\frac{L}{C}} \tag{16.6}$$

$\sqrt{(L/C)}$ is called the *surge impedance*, or *characteristic impedance*, Z_0, of a line. For a lossless line the surge impedance is evidently a pure resistance.

It should be noted that the surge current I, consequent on a voltage surge E, is related to the voltage surge by eqn. (16.6), i.e. by the properties of the line in which the surge travels.

The magnitude of the surge impedance for a particular conductor configuration may be determined from

$$Z_0 = \sqrt{\frac{L}{C}} \text{ ohms} \tag{16.7}$$

For an overhead power line the surge impedance is usually about $300\,\Omega$; for a power cable it usually is about $50\,\Omega$. The inductance per unit length of a line increases with the spacing of the conductors while the capacitance per unit length decreases as the spacing is increased. Thus, by eqn. (16.7), the surge impedance will increase with the spacing.

16.3 Power Input and Energy Storage

When the switch of Fig. 16.2(*a*) is closed, the high-voltage source maintains a potential E volts across the input to the line and supplies a current I amperes:

Power input to line $= EI$ watts

This energy input must become the energy stored in the line, for until the surge reaches the termination, there can be no output of energy from the line.

If v is the surge velocity, then in one second a length v stores electrostatic energy $\frac{1}{2}CvE^2$ and electromagnetic energy $\frac{1}{2}LvI^2$.

Energy input per second = Energy stored per second

i.e.

$$EI = \tfrac{1}{2}CvE^2 + \tfrac{1}{2}LvI^2 \tag{16.8}$$

Now,

$$\tfrac{1}{2}CvE^2 = \tfrac{1}{2}E\frac{C}{\sqrt{(LC)}}E = \tfrac{1}{2}E\frac{E}{\sqrt{\dfrac{L}{C}}} = \tfrac{1}{2}EI$$

and

$$\tfrac{1}{2}LvI^2 = \tfrac{1}{2}I\frac{L}{\sqrt{(LC)}}I = \tfrac{1}{2}I\sqrt{\frac{L}{C}}I = \tfrac{1}{2}EI$$

Thus the electrostatic and electromagnetic stored energies are equal.

16.4 Terminations

Consider a surge, of voltage E_i, impinging on the termination of a transmission line; if the characteristic or surge impedance of the line is Z_0, then the surge current I_i in the transmission line is given by

$$I_i = \frac{E_i}{Z_0} \tag{16.6a}$$

This surge impinging on the termination is called the *incident surge*; E_i is the *incident surge voltage* and I_i is the *incident surge current*.

Power conveyed to termination with incident surge $= E_iI_i$

Consider the particular case of a line terminated in a pure resistor equal to the surge impedance of the line. When the surge arrives the current in the resistor is E_i/Z_0.

$$\text{Power absorbed by resistor} = E_i \times \frac{E_i}{Z_0} = E_iI_i$$

$$= \text{Power transmitted by the surge}$$

In this case the surge power is exactly absorbed by the terminating impedance and there will be no further changes, i.e. the line will continue to be charged to potential E_i, and will continue to carry a current I_i.

If the terminating resistor has a resistance R, either higher or lower than the surge impedance Z_0, then changes in both voltage and current will occur. For instance if $R > Z_0$, then, on the arrival of the surge at the termination, the current through the terminating resistor will be E_i/R, which will be less than E_i/Z_0. The incident current on the line is in excess of the current which can be absorbed by the terminating resistor at the surge voltage. Since the current cannot instantaneously decrease due to the inductance of the line, the excess current will increase the charge on the capacitance at the end of the line. This increases the voltage at the termination to a value higher than the incident voltage E_i.

Let the voltage at the termination rise to E_T. Then

$$\text{Current through terminating resistor} = \frac{E_T}{R} = I_T$$

$$\text{Excess voltage appearing at termination} = E_T - E_i$$

$$\text{Excess current at termination} = I_T - I_i$$

The whole line is now charged to a potential E_i, but there has been a further rise of potential to E_T at the end of the line, i.e. an excess potential suddenly appears at the termination of the line. This is a similar condition to the initial closure of the switch which produced the incident surge, and so, in a similar manner, a surge will now travel from the termination back along the line. This is called the *reflected surge.*

$$\text{Reflected surge voltage, } E_r = E_T - E_i$$

Therefore

$$E_i + E_r = E_T \tag{16.9}$$

$$\text{Reflected surge current, } I_r = I_T - I_i$$

Thus

$$I_i + I_r = I_T \tag{16.10}$$

where $I_i = E_i/Z_0$, $I_r = -E_r/Z_0$, $I_T = E_T/R$.

The positive current direction is the direction of the current in the incident surge. Since the reflected current has the opposite direction the minus sign is necessary.

Substituting in eqn. (16.10),

$$\frac{E_i}{Z_0} - \frac{E_r}{Z_0} = \frac{E_T}{R}$$

Substituting for E_T from eqn. (16.9),

$$\frac{E_i}{Z_0} - \frac{E_r}{Z_0} = \frac{E_i + E_r}{R}$$

Therefore

$$RE_i - RE_r = Z_0 E_i + Z_0 E_r$$

and

$$E_r = \frac{R - Z_0}{R + Z_0} E_i \tag{16.11}$$

Substituting for E_r in eqn. (16.9),

$$E_i + \frac{R - Z_0}{R + Z_0} E_i = E_T$$

Thus

$$E_T = \frac{2R}{R + Z_0} E_i \tag{16.12}$$

Also

$$I_r = -\frac{E_r}{Z_0} = -\frac{1}{Z_0}\frac{R - Z_0}{R + Z_0} E_i \tag{16.13}$$

$$I_T = \frac{E_T}{R} = \frac{2}{R + Z_0} E_i \tag{16.14}$$

Graphs of the voltage and current distributions along the line are shown in Fig. 16.3 for instants before and after the incident surge reaches the termination. Since $R > Z_0$, the current at the termination is reduced and the reflected current surge is negative.

If the terminating resistance R is less than Z_0, then on the arrival of the incident surge the current, E_i/R, in the terminating resistor will be greater than the surge current $I_i = E_i/Z_0$. Due to the line inductance the line current cannot suddenly increase. The capacitance at the end of the line is then discharged, so reducing the voltage at the termination. The reduction of voltage at the termination is equivalent to the sudden application of a negative voltage surge at the termination which travels back along the line. This is the *reflected surge*.

Since the reflected voltage is negative the reflected current will be a negative current in the negative direction, i.e. equivalent to a positive current.

All the previous equations apply without change. The voltage and current distributions are shown in Fig. 16.4.

For an open-circuited line, $R \to \infty$, and hence the terminating current tends to zero. Therefore, by eqn. (16.11),

$$\text{Reflected voltage, } E_r = \frac{\infty - Z_0}{\infty + Z_0} E_i = E_i$$

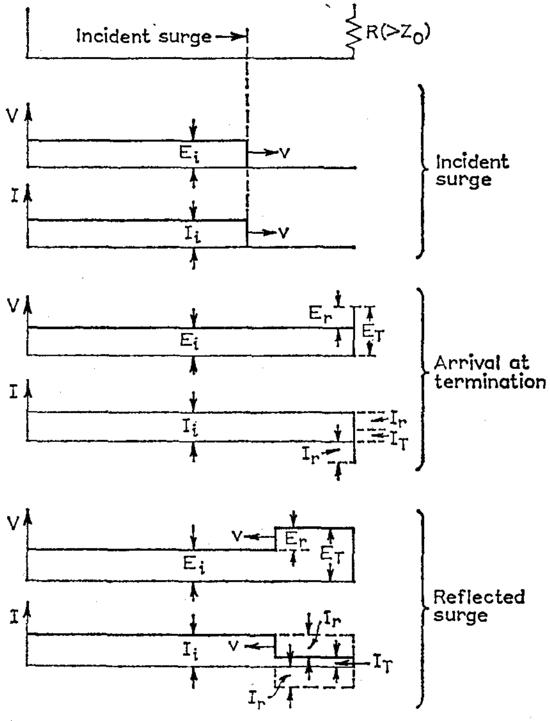


Fig. 16.3 REFLECTION AT A TERMINATION WHERE $R > Z_0$

and

$$\text{Termination voltage, } E_T = E_i + E_i = 2E_i$$

$$\text{Reflected current, } I_r = -\frac{E_r}{Z_0} = -\frac{E_i}{Z_0} = -I_i$$

Thus

$$\text{Termination current, } I_T = I_r + I_i = 0$$

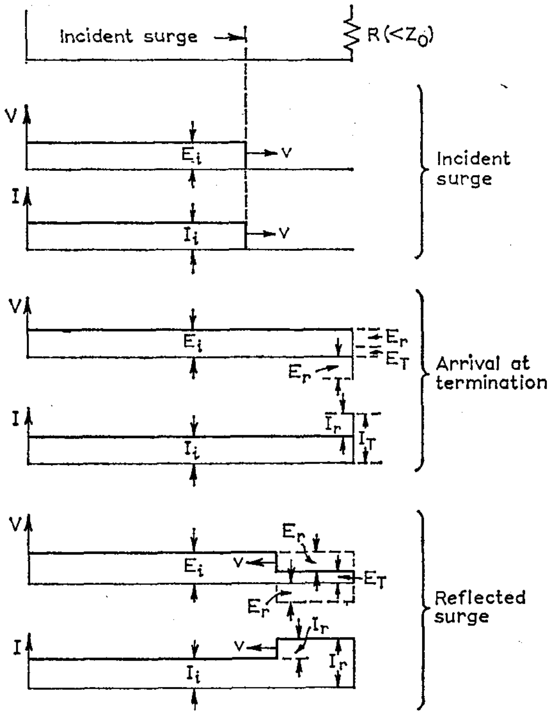


Fig. 16.4 REFLECTION AT A TERMINATION WHERE $R < Z_0$

For a short-circuited line, $R = 0$; therefore by eqn. (16.11),

$$\text{Reflected voltage, } E_r = \frac{0 - Z_0}{0 + Z_0} E_i = -E_i$$

and

$$\text{Termination voltage, } E_T = E_i - E_i = 0$$

$$\text{Reflected current, } I_r = -\frac{E_r}{Z_0} = \frac{E_i}{Z_0} = I_i$$

Thus

$$\text{Termination current, } I_T = I_i + I_i = 2I_i$$

Obviously there can be no voltage at the short-circuit. It should be noted that the terminating voltage and current are often called the *transmitted voltage* and *current.*

To summarize, if $R > Z_0$ the voltage surge is reflected unchanged in sign, while the current surge is reflected with the opposite sign. If $R < Z_0$ the sign of the reflected voltage surge is reversed, while that of the current surge is unchanged.

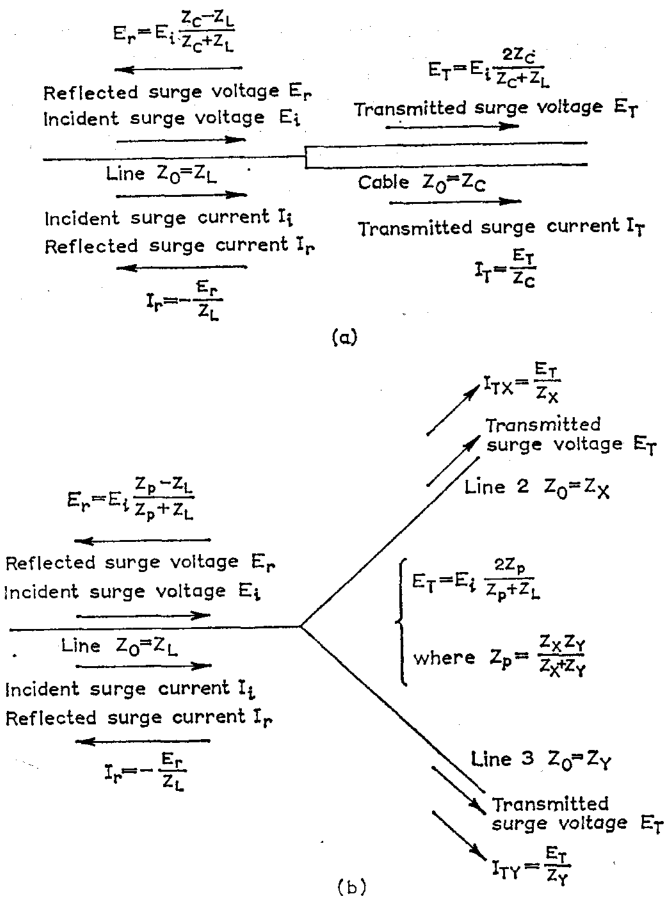


Fig. 16.5 SURGES AT LINE JUNCTIONS

16.5 Junctions of Lines having Different Characteristic Impedances

If a voltage surge travels along a line towards a junction where the characteristic impedance of the line changes, reflexion will take place at the junction and the surge transmitted into the section beyond the junction will be modified in value (Fig. 16.5(*a*)).

Consider the case of a junction between an overhead line and a cable. Let the characteristic impedance of the line be Z_L and that of the cable be Z_C. Suppose a surge voltage E_i is initiated in the line and travels down the line to the junction between the line and the cable. The cable presents a terminal impedance of Z_C to the line, since the cable voltage will be Z_C times the cable current under surge conditions.

Initial surge voltage in line = E_i

Initial surge current in line = $I_i = \dfrac{E_i}{Z_L}$

The final voltage in the line after reflexion from the junction is the voltage surge transmitted into the cable:

$$E_T = \frac{2Z_C}{Z_C + Z_L} E_i \tag{16.12}$$

The final current in the line after reflexion from the junction is the current surge transmitted into the cable:

$$I_T = \frac{2E_i}{Z_C + Z_L} \tag{16.14}$$

Reflected voltage in line, $E_r = \dfrac{Z_C - Z_L}{Z_C + Z_L} E_i$ (16.11)

Reflected current in line, $I_r = -\dfrac{1}{Z_0}\dfrac{Z_C - Z_L}{Z_C + Z_L} E_i$ (16.13)

It should be noted that, since the characteristic impedance of the cable Z_C is likely to be much less than the characteristic impedance of the line Z_L, the magnitude of the voltage surge in the cable will be much less than that in the line.

If there is a junction between three lines or a "tee" junction on one line (Fig. 16.5(*b*)), the reflected and transmitted surges may be calculated by the above equations with Z_p replacing Z_C, where

$$Z_p = \frac{Z_X Z_Y}{Z_X + Z_Y}$$

EXAMPLE 16.1 An underground cable having an inductance of 0·3 mH/km and a capacitance of 0·4 μF/km is connected in series with an overhead line having an inductance of 2·0 mH/km and a capacitance of 0·014 μF/km.

Calculate the values of the transmitted and reflected waves of voltage and current at the junction, due to a voltage surge of 100 kV travelling to the junction (*a*) along the cable, and (*b*) along the line. (*L.U.*)

Characteristic impedance of cable, $Z_C = \sqrt{\dfrac{L}{C}}$

$$= \sqrt{\frac{0.3 \times 10^{-3}}{0.4 \times 10^{-6}}} = 27.4\,\Omega$$

Characteristic impedance of line, $Z_L = \sqrt{\dfrac{L}{C}}$

$$= \sqrt{\frac{2.0 \times 10^{-3}}{0.014 \times 10^{-6}}} = 378\,\Omega$$

(*a*) 100 kV surge initiated in cable

Initial value of surge voltage, $E_i = 100\,\text{kV}$

Initial value of surge current, $I_i = \dfrac{E_i}{Z_C} = \dfrac{100}{27.4} = 3.65\,\text{kA}$

Surge voltage transmitted into line, $E_T = \dfrac{2Z_L}{Z_L + Z_C} E_i$

$$= \frac{2 \times 378}{378 + 27.4} \times 100 = \underline{\underline{186\,\text{kV}}}$$

Surge current transmitted into line, $I_T = \dfrac{E_T}{Z_T} = \dfrac{E_T}{Z_L}$

$$= \frac{186}{378} = \underline{\underline{0.492\,\text{kA}}}$$

Reflected surge voltage in cable, $E_r = E_i \dfrac{Z_L - Z_C}{Z_L + Z_C}$

$$= 100 \times \frac{378 - 27.4}{378 + 27.4} = \underline{\underline{86\,\text{kV}}}$$

Reflected surge current in cable, $I_r = -\dfrac{E_r}{Z_C} = -\dfrac{86}{27.4} = \underline{\underline{-3.16\,\text{kA}}}$

(*b*) 100 kV surge initiated in line

Initial value of surge voltage, $E_i = 100\,\text{kV}$

Initial value of surge current, $I_i = \dfrac{E_i}{Z_L} = \dfrac{100}{378} = 0.264\,\text{kA}$

Surge voltage transmitted into cable, $E_T = \dfrac{2Z_C}{Z_C + Z_L} E_i$

$$= \frac{2 \times 27.4}{27.4 + 378} = 100 = \underline{\underline{13.5\,\text{kV}}}$$

Surge current transmitted into cable, $I_T = \dfrac{E_T}{Z_C} = \dfrac{13.5}{27.4} = \underline{\underline{0.494\,\text{kA}}}$

Reflected surge voltage in line, $E_r = E_T - E_i = \underline{\underline{-86.5\,\text{kV}}}$

Reflected surge current in line $= -\dfrac{E_r}{Z_L} = +\dfrac{86.5}{378} = \underline{\underline{0.23\,\text{kA}}}$

EXAMPLE 16.2 A single-phase overhead line is 50km long and has a surge impedance of 300Ω. The line has a circuit-breaker at the input end. A dead

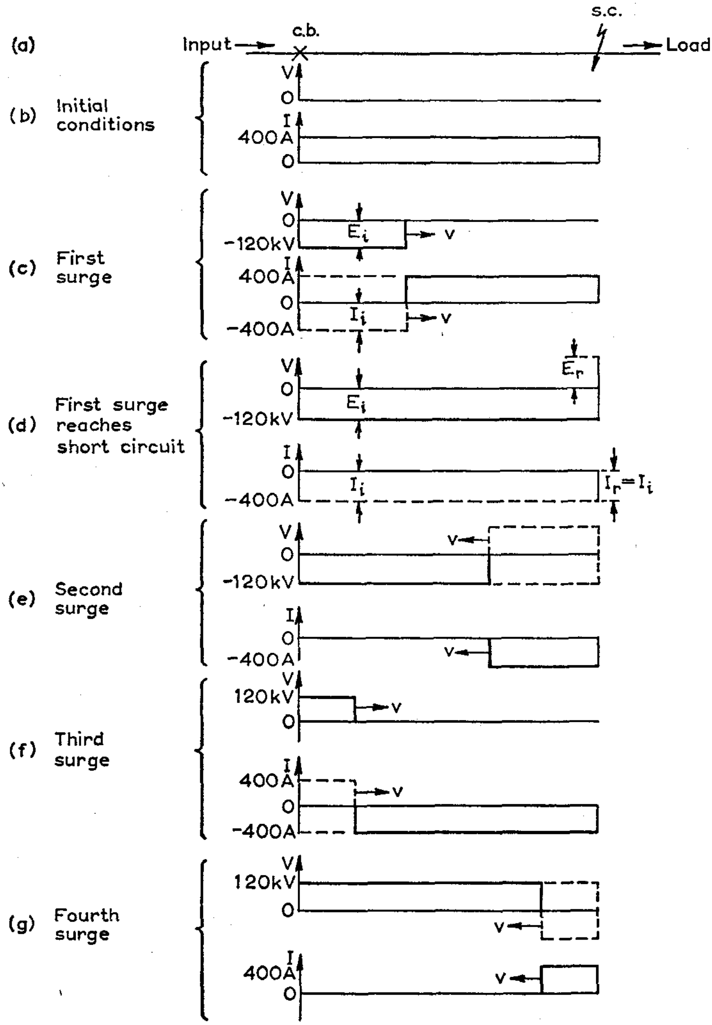


Fig. 16.6

short-circuit occurs at the terminating end and the circuit-breaker suddenly interrupts the short-circuit current when it has an instantaneous value of 400 A. Describe the surge phenomena which will occur in the line.

The line is shown in Fig. 16.6(*a*). Under the initial short-circuit conditions the voltage along the line may be taken as zero and the current as uniform at 400 A.

When the circuit-breaker opens, the current of 400 A cannot immediately cease due to the line inductance; thus the line capacitance at the circuit-breaker end must become negatively charged by the instantaneous continuation of the 400 A current. A negative voltage then arises at the circuit-breaker end while the rest of the line is uncharged; hence a negative voltage surge travels from the circuit-breaker end. The current associated with this surge will be −400 A, since there can be no resultant current at the open-circuited circuit-breaker end. Therefore

$$\text{Voltage of surge} = -400 \times Z_0 = -400 \times 300\,\text{V} = -120\,\text{kV}$$

i.e. a surge of −120 kV and −400 A travels down the line to the short-circuit termination.

At the short-circuit the termination voltage is necessarily zero; thus a reflected surge arises with a voltage of +120 kV and a current of −400 A. (At a short-circuit the surge voltage is reflected with change of sign and the surge current is reflected without change of sign, see Section 16.4.) When the reflected surge reaches the open-circuited circuit-breaker end it will in turn be reflected back down the line. The third surge voltage will be +120 kV and the third surge current will be +400 A. (At an open-circuit the surge voltage is reflected without change of sign and the surge current is reflected with change of sign; see Section 16.4.) These reflexions obey the rules: (i) the resultant voltage at a short-circuit must be zero, and (ii) the resultant current at an open-circuit must be zero.

The fourth surge stage is shown in Fig. 16.6(*g*).

It will be seen that after the fourth stage the resultant voltage and current are the same as the initial voltage and current. The fifth surge would then be the same as the first surge. In all real lines the losses will continually be reducing the magnitude of the surges so that the later surges are much smaller than the first ones.

Since this line is an overhead line, the surge velocity will approach 3×10^8 m/s. Thus the time required for a surge to travel the length of the line will be $50{,}000/(3 \times 10^8)$ s, i.e. 0·167 ms.

16.6 Surges of Short Duration

The previous theory has been developed on the assumption of a sudden rise of voltage followed by the steady application of the same voltage. In practice the sudden rise of voltage is usually followed by a slow fall of voltage back to normal. If the surge front reaches the far end of the line before the input voltage falls appreciably, then the previous theory gives a good representation of the surge conditions on the line. If the sudden rise of voltage is followed by a rapid fall so that the voltage surge becomes a voltage pulse, the previous theory does not give a clear representation of the conditions on the line. The theory is, however, still applicable and a further simple assumption gives a good representation of the actual conditions.

Fig. 16.7(*a*) shows a pulse or short-duration-surge source connected to a line. The output voltage of the source is shown in Fig. 16.7(*b*).

This output voltage may be considered as (i) a positive voltage surge of infinite duration followed at a discrete interval by (ii) a negative voltage surge of infinite duration and the same magnitude. The negative surge will then cancel the "tail" of the positive surge. Since both the positive and negative long-duration surges will obey the previous equations, the pulse voltage will also obey them.

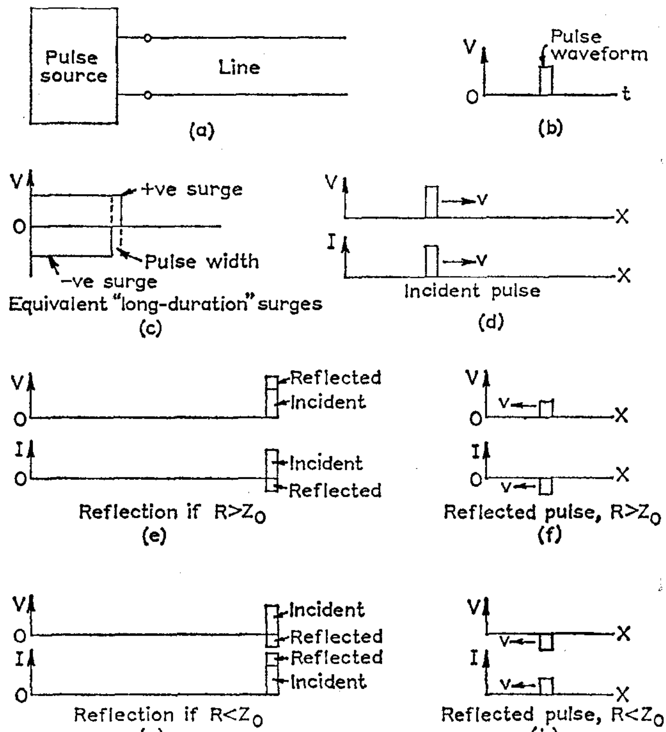


Fig. 16.7. SURGES OF SHORT DURATION

Incident and reflected pulses are shown in Figs. 16.7(*d*)–(*h*). Losses along the line will tend to decrease the pulse magnitude and to broaden the pulse front and pulse tail.

EXAMPLE 16.3 A short-duration pulse of magnitude 10kV travels along a very long line of characteristic impedance 300Ω. The line is joined to a similar very long line by a cable of characteristic impedance 30Ω and with a relative permittivity of 4. If the cable is 1·5km long, calculate the magnitude of the first and second pulses entering the second line. What is the time interval between the two pulses?

$$\text{Magnitude of pulse transmitted into cable} = 10 \times \frac{2 \times 30}{30 + 300} \qquad (16.12)$$
$$= 1{\cdot}82\,\text{kV}$$

Part of the input pulse will be reflected back along the first very long line and thus conveyed away from the cable. Assuming no losses in the cable, the pulse of 1·82 kV will be incident on the junction of the cable with the second length of line.

Magnitude of pulse transmitted into second line from cable

$$= 1{\cdot}82 \times \frac{2 \times 300}{300 + 30} = \underline{\underline{3{\cdot}31\,\text{kV}}}$$

This is the first pulse to travel along the second length of line.

Magnitude of pulse reflected back into cable from junction with second length of line

$$= 1{\cdot}82 \times \frac{300 - 30}{300 + 30} = 1{\cdot}49\,\text{kV} \qquad (16.11)$$

This reflected pulse will travel back along the cable to the junction with the first length of line. Part of the reflected pulse will be transmitted into the first length of line and part of it will be reflected and will again pass down the cable toward the second length of line.

Magnitude of pulse reflected back into cable from junction with first length of line

$$= 1{\cdot}49 \times \frac{300 - 30}{300 + 30} = 1{\cdot}22\,\text{kV} \qquad (16.11)$$

This pulse will form a second pulse incident on the junction of the cable and the second length of line. Therefore

Magnitude of second pulse transmitted into second line

$$= 1{\cdot}22 \times \frac{2 \times 300}{300 + 30} = \underline{\underline{2{\cdot}22\,\text{kV}}}$$

The time interval between the two pulses will be the time required for a pulse to travel first back and then forward along the length of cable, i.e. a total distance of 3 km in the cable.

$$\text{Velocity of pulse in cable} = \frac{3 \times 10^8}{\sqrt{4}} = 1{\cdot}5 \times 10^8\,\text{m/s}$$

Therefore

$$\text{Time interval} = \frac{3 \times 10^3}{1{\cdot}5 \times 10^8} = \underline{\underline{2 \times 10^{-5}\,\text{s}}}$$

Note. The surge pulses in the second line were much smaller than those in the first. Equipment in stations at the ends of overhead lines is sometimes protected from overvoltage surges by bringing the overhead lines through a short length of cable before reaching the station.

16.7 Mitigation of High-voltage Surges

High-voltage surges are mainly due to either (*a*) lightning discharges, or (*b*) switching. The voltages set up by lightning discharges are reduced by stringing one or two earth wires above the main conductors. The voltages set up by switching are reduced by using *resistance switching*.

A direct lightning stroke to a line causes enormous voltages and there is little possibility of preventing these. Fortunately direct lightning strokes are rare. It is more common for high-voltage surges to be caused by the charge induced on the conductors of an overhead line when a charged cloud passes over or near to the line. The charge induced on the conductors will have the opposite polarity to that in the cloud. If the cloud passes slowly away, the charges induced on the conductors will gradually flow to earth and no disturbance will be caused. If, however, the cloud is suddenly discharged by a lightning stroke to earth or another cloud, then the induced charge on the line will be suddenly released and a surge voltage will travel along the line in either direction.

The charge induced on an overhead transmission line system by a charged cloud is mainly concentrated in the uppermost conductor, since the other conductors are to some extent electrostatically shielded by the uppermost one. If the uppermost conductor is made an earth wire and not one of the system conductors, then the charges induced on the system conductors are considerably reduced. This reduces the surges in the system conductors. It also affords some mitigation of the effects of a direct lightning stroke.

Where an earth wire is present the resistance of the tower footings must be kept low or *back flashover* from the earth wire may occur. For example, in the extreme case of the resistance to earth of the tower footing being infinite, any surge voltage wavefront reaching the tower base is doubled and reflected back to the earth wire. Eventual build-up of earth-wire potential to a value well above that of the system conductors may result in a discharge from the earth wire to the system conductors.

Transient currents and voltages naturally occur with most switching operations. Generally switching-in and disconnecting can be performed without dangerous disturbances arising. The interruption of a high short-circuit current by an efficient circuit-breaker does, however, tend to give high-voltage surges. These can be mitigated by arranging that the circuit-breaker will be opened in stages. During the stages one or more resistance sections carry the current which is being interrupted and part, at least, of the energy stored in the line inductance is dissipated in the switch resistors.

16.8 Protection of Insulation

The insulators which support an overhead line and the insulation of cables, switches or transformers will, under some surge conditions, have voltages impressed on them which are greater than the breakdown strength of the insulators or insulation. To prevent the breakdown of these costly units and to prevent the interruption of the supply which would result from their breakdown, the insulation is usually protected by an air-gap so arranged that the surge voltage will produce breakdown of the air-gap rather than of the insulation. A string of insulators for an overhead line, or the bushing of a transformer, has frequently a rod gap across it (Fig. 16.8), so that a

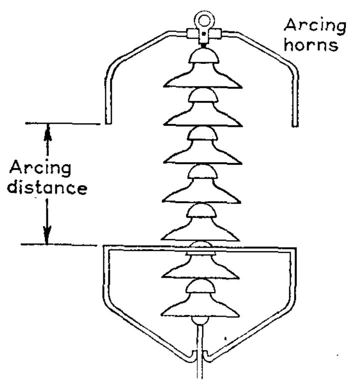


Fig. 16.8 ROD GAP PROTECTING AN INSULATOR STRING

spark or an arc will jump across the rod gap rather than down the insulator or the bushing.

Alternatively a metal ring concentric with the insulator string and about level with the third insulator shed may be used as the lower electrode in place of a rod electrode.

When setting the rod gap two factors must be taken into account: (*a*) impulse ratio and (*b*) time factor.

Impulse ratio

$$= \frac{\text{Breakdown voltage under surge conditions}}{\text{Breakdown voltage under low-frequency conditions}}$$

It is found that the breakdown voltage under surge, i.e. rapidly changing or high-frequency conditions, is often higher than the breakdown voltage under steady or low-frequency conditions. The *impulse ratio* is a measure of this difference. Supposing the breakdown voltage of a string of insulators is, say, 300 kV at 50 Hz and that the string is protected by a rod gap with a breakdown voltage of, say, 200 kV at 50 Hz. If the impulse ratio for the insulators is, say, 1·3, then the surge breakdown voltage for the insulators will be 390 kV; and if the impulse ratio for the rod gap is, say, 2·1, then the surge breakdown voltage for the rod gap will be 420 kV. The rod gap does not then protect the insulators under surge conditions. Either the impulse ratio for the rod gap must be improved or the 50 Hz setting for the rod gap must be reduced. The impulse ratio is found to depend on the geometry of the air-gap. A sphere gap, with relatively close spacing, has an impulse ratio of unity—a needle gap may have an impulse ratio of between 2 and 3.

Time Factor. The breakdown of insulation or an air-gap does not occur instantaneously on the application of the excess voltage. The time for the complete breakdown to develop depends (i) on the magnitude of the excess voltage, (ii) on the material in the breakdown path, and (iii) on the shape and spacing of the electrodes. Naturally the greater the excess voltage the shorter is the time required for breakdown to develop—a typical characteristic is shown in Fig. 16.9(*a*).

Fig. 16.9(*b*) compares the characteristics of a rod gap and an insulator which are used in conjunction with one another. If the voltage across them were slowly increased, the rod gap would correctly break down first, i.e. at the lower voltage. If a voltage greater than V_c were suddenly applied, the insulator would break down first and thus, under a steep-wavefront surge condition, the rod gap does not protect the insulator. The correct relative characteristics for a rod gap to protect the insulation for all surge voltages is shown in Fig. 16.9(*c*).

The time delay is relatively short for sphere gaps and relatively long for needle gaps.

16.9 Surge Diverters

Rather than permit a surge to impinge on the terminal apparatus it is advantageous to eliminate the surge if possible. The elimination may be carried out in two ways, either (*a*) the surge may be diverted to earth, i.e. short-circuited, or (*b*) the surge energy may be absorbed. The latter method is not now used.

Modern surge diverters consist essentially of elements having a non-linear volt/ampere characteristic, and made of a ceramic material consisting of silicon carbide bonded with clay.

To protect plant successfully against high-voltage travelling waves the surge diverter must operate, as far as possible, simultaneously with the incidence of the surge. Since the surge is travelling at

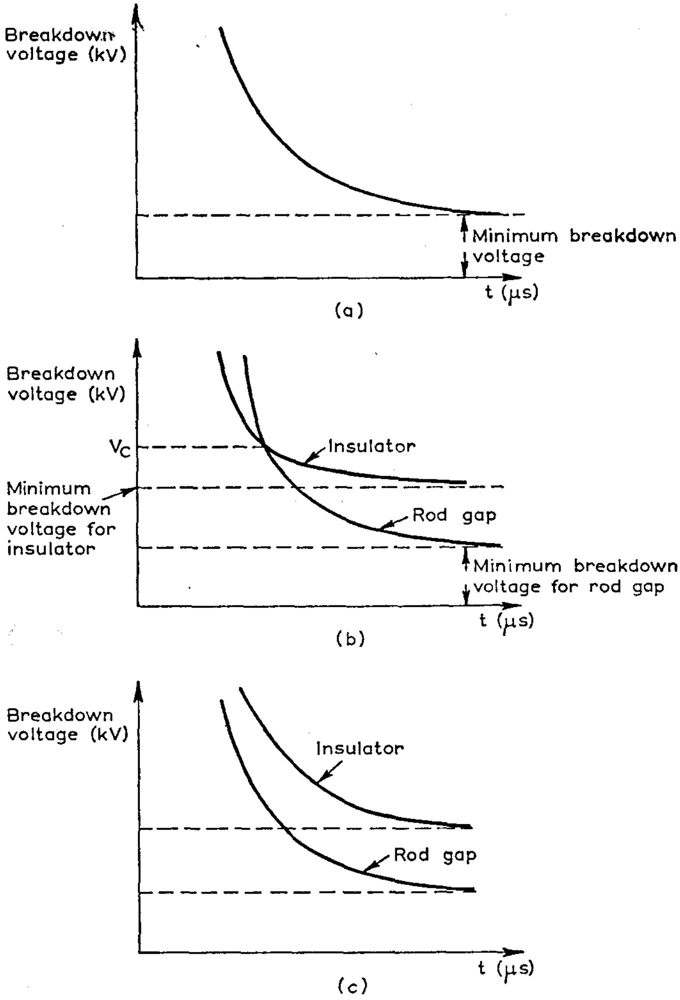


Fig. 16.9 BREAKDOWN-VOLTAGE/TIME CHARACTERISTICS

(*a*) Typical rod gap characteristic
(*b*) Unsuitable combination of insulator and rod gap
(*c*) Correct relative characteristics for insulator and rod gap protection

approximately 3×10^8 m/s a short delay will permit the surge to pass the diverter and be transmitted into the plant which the diverter is intended to protect. Moreover, the surge diverter should be placed as close as possible to the plant to be protected to obviate the risk of a surge being initiated in the line between the horn gap and the plant. The line between the surge diverter and the plant should be protected by an earth wire or wires.

Fig. 16.10 shows approximately a typical characteristic of a Metrosil disc suitable for incorporation in a surge diverter designed

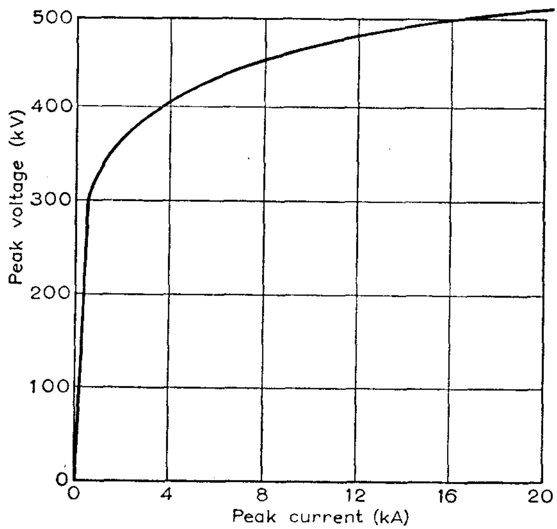


Fig. 16.10 TYPICAL METROSIL VOLT/AMPERE CHARACTERISTIC

to operate on a high-voltage transmission system to protect transformers and other plant from high-voltage surges.

The law connecting the applied voltage and the current is of the form

$$V = kI^{\beta}$$

where k is a constant depending on the geometrical form and β is a constant depending on the composition and treatment of the substance. Ideally β should be zero, so that whatever the value of the surge current the voltage would be constant. In practice values for β of the order of 0·2 are achieved.

The principle of operation is that a stack of Metrosil discs is connected between line and earth close to the transformer (or other plant) to be protected. Because of the nature of the volt/ampere characteristic, at normal voltage the diverter passes only a very small current to earth, but when a high over-voltage occurs the resistance

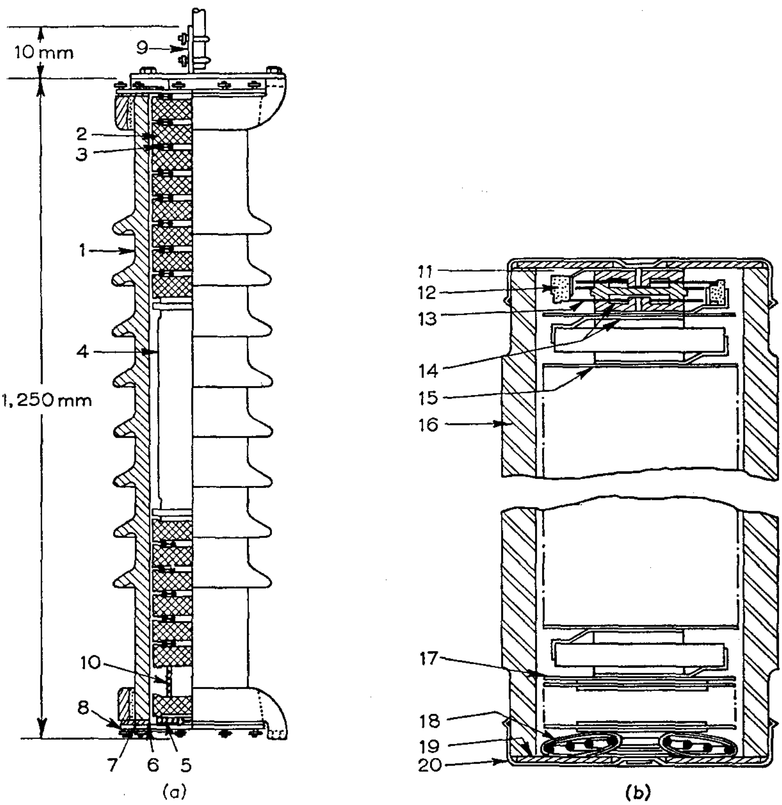


Fig. 16.11 METROSIL DIVERTER AND SPARK-GAP ASSEMBLY

(*AEI Ltd.*)

(*a*) Surge diverter

1. Glazed porcelain housing
2. Metrosil disc
3. Metallic spacers
4. Spark-gap assembly
5. Compression spring
6. Inner sealing gasket
7. Outer sealing gasket
8. Sealing plate
9. Terminal assembly
10. Spacing tube

(*b*) Section through spark-gap assembly

11. Contact clips
12. Metrosil grading ring
13. Mica disc
14. Electrode
15. Locating disc
16. Porcelain housing
17. Metal spacers
18. Compression spring
19. Contact plate
20. Sealing cap

of the diverter falls and the diverter passes a high current, diverting the surge energy to earth.

In practice, in a diverter suitable for operation on a 132kV system, a stack of Metrosil discs 6in. in diameter is assembled in a glazed porcelain housing which is provided on its exterior with rain sheds, which may be of a special shape for operation in dirt-laden atmospheres. Spark gaps are incorporated to prevent current flow to earth under normal-voltage conditions. The air in the interior of the porcelain housing containing the spark-gap assembly is evacuated and then the housing is filled with nitrogen. Fig. 16.11 shows the details of the arrangement of a 33kV Metrosil diverter and spark-gap assembly.

The diverters are mounted vertically, mechanical support being also provided at the top of the assembly for larger ratings. External stress rings are provided for ratings above 110kV.

The diverters are normally set to operate on twice normal voltage, it being undesirable for them to operate on small over-voltages. Operation is extremely rapid, taking less than a microsecond. The impulse ratio is practically unity.

PROBLEMS

16.1 An overhead line of surge impedance 500Ω terminates in a transformer of surge impedance 3,500Ω. Find the amplitudes of the current and voltage surge transmitted to the transformer due to an incident voltage of 30kV. *(H.N.C.)*

Ans. 52·5kV; 0·015kA.

16.2 Derive an expression for the surge impedance of a transmission line.

A transmission line has a capacitance of 0·012 μF per km and an inductance of 1·8 mH per km. This overhead line is continued by an underground cable with a capacitance of 0·45 μF per km and an inductance of 0·3 mH per km. Calculate the maximum voltage occurring at the junction of line and cable when a 20kV surge travels along the cable towards the overhead line. *(H.N.C.)*

Ans. 37·5kV.

16.3 Obtain an expression for the surge impedance of a transmission line and for the velocity of propagation of electric waves in terms of the line inductance and capacitance.

A cable having an inductance 0·3 mH per km and a capacitance of 0·4 μF per km is connected in series with a transmission line having an inductance of 1·5 mH per km and a capacitance of 0·012 μF per km. A surge of peak value 50kV originates in the line and progresses towards the cable. Find the voltage transmitted into the cable. Use the result to explain the practice sometimes adopted of terminating a line by a short length of cable before connecting to reactive apparatus. *(H.N.C.)*

Ans. 7·2kV.

16.4 An overhead transmission line 300 km long, having a surge impedance of 500 Ω is short-circuited at one end and a steady voltage of 3 kV is suddenly applied at the other end.

Neglecting the resistance of the line explain, with the aid of diagrams, how the current and voltage change at different parts of the line, and calculate the current at the end of the line 0·0015 s after the voltage is applied.

Ans. 0.

16.5 Two stations are connected together by an underground cable having a capacitance of 0·15 μF/km and an inductance of 0·35 mH/km joined to an overhead line having a capacitance of 0·01 μF/km and an inductance of 2·0 mH/km.

If a surge having a steady value of 100 kV travels along the cable towards the junction with an overhead line, determine the values of the reflected and transmitted waves of voltage and current at the junction.

State briefly how the transmitted waves would be modified along the overhead line if the line were of considerable length. (*L.U.*)

Ans. 81 kV; 181 kV; 1·57 kA; 0·404 kA.

16.6 Derive an expression for the velocity with which a disturbance will be transmitted along a transmission line.

A disturbance, due to lightning, travels along an overhead line of characteristic impedance 200 Ω. After travelling 30 km along the line the disturbance reaches the end of the line where it is joined to a cable of surge impedance 50 Ω and dielectric constant [relative permittivity] 6. Calculate the relative magnitude of the energy of the disturbance in the cable and the time taken between initiation and arrival at a point 15 km along the cable from the junction. (*H.N.C.*)

Ans. 0·64; 225 μs.

16.7 An overhead transmission line having a surge impedance of 500 Ω is connected at one end to two underground cables, one having a surge impedance of 40 Ω and the other one of 60 Ω. A rectangular wave having a value of 100 kV travels along the overhead line to the junction.

Deduce expressions for, and determine the magnitude of, the voltage and current waves reflected from and transmitted beyond the junction.

If the rectangular wave originated at a long distance from the junction, state how and why it would be modified in its passage along the line. (*L.U.*)

Ans. 90·8 kV; 0·182 kA; 9·16 kV; 0·229 kA; 0·153 kA.

16.8 Two single transmission lines A and B with earth return are connected in series and at the junction a resistance of 2,000 Ω is connected between the lines and earth. The surge impedance of line A is 400 Ω and of B 600 Ω. A rectangular wave having an amplitude of 100 kV travels along line A to the junction.

Develop expressions for and determine the magnitude of the voltage and current waves reflected from and transmitted beyond the junction. What value of resistance at the junction would make the magnitude of the transmitted wave 100 kV? (*L.U.*)

Ans. 7 kV; 0·018 kA; 107 kV; 0·178 kA; 1,200 Ω.

Chapter 17

SHORT-CIRCUIT PROTECTION

The possibility of a short-circuit to earth or between the phases of any transmission system due to mechanical and/or electrical breakdown cannot be ignored. Hence it is essential that the maximum fault currents which could exist in such a system be calculated and that adequate measures be taken to reduce to a minimum the effects which these currents would have. It is also necessary to ensure that the circuit-breakers which are installed in the system will be capable of interrupting these fault currents.

Modern switchgear ratings are approximately:

35,000 MVA for the 400 kV system
15,000 MVA for the 275 kV system
7,000 MVA for the 132 kV system
2,500 MVA for the 66 kV system

Usually the internal leakage reactances of the interconnected generators and transformers are sufficient to limit the prospective fault MVA within these ratings, but in exceptional cases separate reactors may be used to limit fault levels. In older power stations the use of external reactors was standard practice. The methods of reactor connection are discussed below.

17.1 Reactor Control of Short-circuit Currents

The circuit-breakers connected in a transmission network must be capable of dealing with the maximum possible short-circuit current that can occur at their points of connexion. If no steps are taken to limit the value of the possible short-circuit currents, not only will the duty required of circuit-breakers be excessively heavy, but damage to the lines, cables and interconnected plant will almost certainly occur.

Reactors may be connected in power station circuits to limit the maximum possible short-circuit current occurring at any point to a value which will not cause damage to plant for the short time during which it flows, and to relieve the duty required of circuit-breakers. The use of reactors may also increase the chance of continuity of supply by making it possible for healthy sections of the power station busbars to continue in operation.

There are three types of reactor in general use.

IRON-CORED MAGNETICALLY-SHIELDED REACTOR (Fig. 17.1)

The turns of the reactor are surrounded by a number of iron cores providing a complete magnetic path. The cross-sectional area of the cores is made as generous as possible to limit the change in reactance due to magnetic saturation of the cores. The disposition of the cores gives almost complete magnetic shielding. The whole unit is reinforced mechanically and immersed in oil to enable it to withstand high voltages, to permit of good cooling, and to make outdoor operation possible.

AIR-CORED RING-SHIELDED REACTOR (Fig. 17.2)

A reactance coil having an air core is surrounded by short-circuited shielding rings mounted sufficiently far from the coil not to reduce the inductance unduly. The shielding achieved is sufficiently good to permit the assembly to be inserted in an oil-filled tank. The chief advantage of this type is that saturation difficulties are avoided, but, compared with the iron-cored type, the L/R ratios obtainable are smaller.

AIR-CORED NON-SHIELDED REACTOR (Fig. 17.3)

In this type of reactor, the air-cored coil is supported by cast concrete pillars. The reactor is mounted on porcelain pedestal-type insulators. The construction is cheap and robust, but since no shielding is provided the reactor is not suitable for immersion in an oil tank. The reactor is not suitable for outdoor operation and may not be placed near metal objects.

17.2 Location of Current-limiting Reactors

Current-limiting reactors may be connected in series with each generator, in series with each feeder, or between busbar sections.

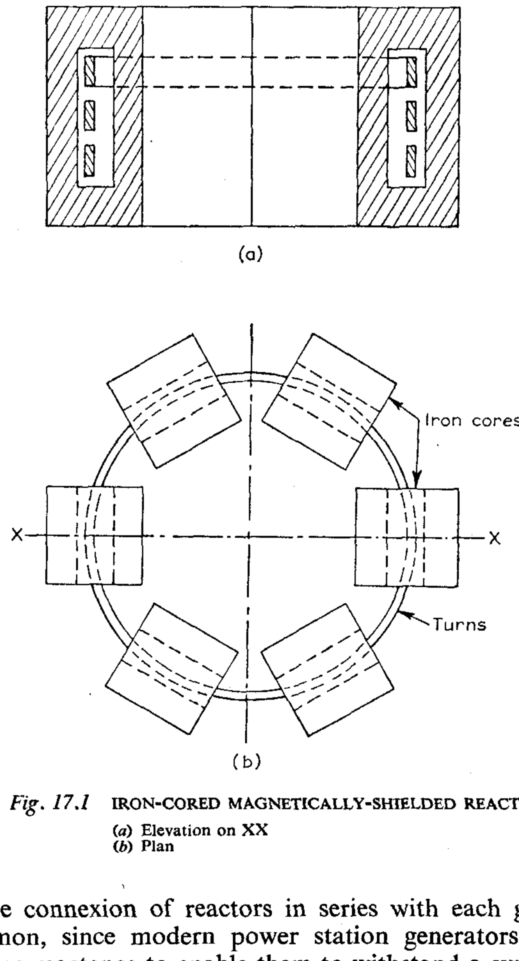


Fig. 17.1 IRON-CORED MAGNETICALLY-SHIELDED REACTOR

(*a*) Elevation on XX
(*b*) Plan

The connexion of reactors in series with each generator is not common, since modern power station generators have sufficient leakage reactance to enable them to withstand a symmetrical short-circuit across their terminals. The disadvantages of connecting reactors in series with the generators are (i) that there is a relatively large power loss and voltage drop in the reactor, (ii) that a busbar or feeder fault close to the busbar will reduce the busbar voltage to a low value thereby causing the generators to fall out of step, and (iii) that a fault on one feeder is likely to affect the continuity of supply to others.

When the reactors are connected in series with each feeder, a feeder fault will not seriously affect the busbar voltage so that

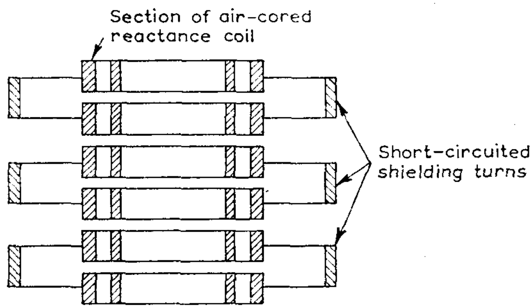


Fig. 17.2 AIR-CORED RING-SHIELDED REACTOR

there is little tendency for the generators to lose synchronism, other feeders will be little affected and the effects of the fault will be localized. The disadvantages of this method of connexion are (i) that

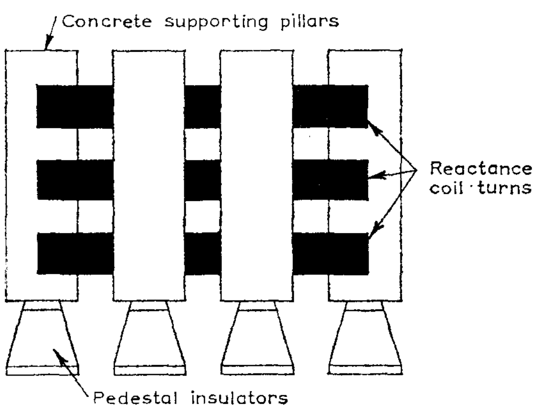


Fig. 17.3 AIR-CORED NON-SHIELDED REACTOR

there is a relatively large power loss and voltage drop in each reactor, since the reactors are in series with the feeder currents, (ii) if the number of generators is increased so the size of the feeder reactors will also have to be increased to keep the short-circuit current within the rating of the feeder circuit-breakers, and (iii) no protection is given against busbar faults.

The most common system of connecting reactors is between busbar sections. There are two methods, the ring system and the tie-bar system, and these are shown in Fig. 17.4. Under normal operating conditions, each generator will supply its own section of the load and no current will flow through the reactors, thus reducing power loss and voltage drop in the reactors. Even when power has to be transferred from one busbar section to another, a large voltage difference between sections is not necessary, since the voltage drop

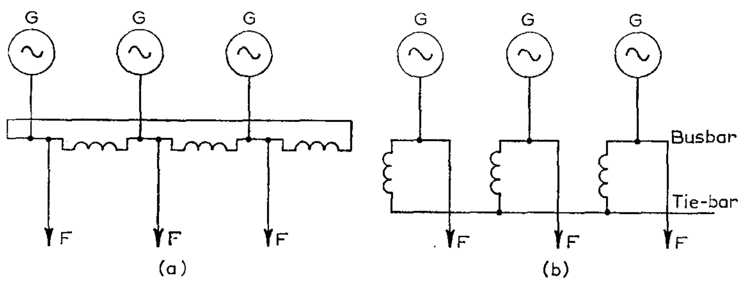


Fig. 17.4 BUSBAR REACTORS
(*a*) Ring reactors
(*b*) Tie-bar reactors

in the reactor will be almost in quadrature with the busbar voltage. The transfer of reactive current between sections is, however, difficult and would necessitate large voltage differences between sections, so that each generator must supply the reactive current for its own load.

The presence of the reactors between busbar sections has the effect of tending to localize any fault which occurs. Thus a feeder fault is likely to affect only the busbar section to which it is connected, the other sections being able to continue in normal operation.

Comparing the ring system with the tie-bar system, it will be seen that in the tie-bar system there are effectively two reactors in series between sections so that the reactors must have approximately half the reactance of those used in a comparable ring system. The tie-bar system has the disadvantage of requiring an additional busbar, but has the advantage that additional generators may be added to the system without necessitating changes in the existing reactors. This is because there is a limiting value of short-circuit current which can be fed into a busbar fault when the tie-bar system is used (see Section 17.5).

17.3 Per-unit Reactance and Resistance*

The per-unit value of a quantity A was defined in Chapter 9 as

$$A_{pu} = \frac{A}{A_{base}} \tag{9.78}$$

The per-unit reactance is therefore

$$X_{pu} = \frac{X}{X_{base}} \tag{17.1}$$

From eqn. (9.80) the base reactance is

$$X_{base} = \frac{V_{base}}{I_{base}} \tag{9.80}$$

The rated full-load values are usually chosen as bases, so that

$$X_{base} = \frac{V_{fl}}{I_{fl}} \tag{17.2}$$

Substituting for X_{base} in eqn. (17.1),

$$X_{pu} = X\frac{I_{fl}}{V_{fl}} \tag{17.3}$$

Consider a supply system of reactance X ohms and negligible resistance supplying a load; if the load is suddenly short-circuited, the current I_{sc} which will flow is

$$I_{sc} = \frac{V_{fl}}{X} = \frac{V_{fl}}{X_{pu}\dfrac{V_{fl}}{I_{fl}}} = \frac{I_{fl}}{X_{pu}} \tag{17.4}$$

If the rated full-load volt-amperes is S_{fl}, then multiplying eqn. (17.4) by V_{fl} gives

$$V_{fl}I_{sc} = \frac{V_{fl}I_{fl}}{X_{pu}} = \frac{S_{fl}}{X_{pu}}$$

i.e.

$$\text{Short-circuit volt-amperes} = \frac{S_{fl}}{X_{pu}} \tag{17.5}$$

* Reactances and resistances are often quoted in percentage values. The corresponding per-unit values are obtained by dividing the percentage values by 100.

Per-unit resistance is dealt with in an analogous manner; thus

$$R_{pu} = R\frac{I_{fl}}{V_{fl}} \tag{17.6}$$

and

$$\text{Short-circuit volt-amperes} = \frac{S_{fl}}{R_{pu}} \tag{17.7}$$

In a supply system containing both series resistance and reactance,

$$\text{Short-circuit volt-amperes} = \frac{S_{fl}}{R_{pu} + jX_{pu}} = \frac{S_{fl}}{Z_{pu}} \tag{17.8}$$

17.4 Reference of Per-unit Impedance to a Common Base

Where a system contains units of different volt-ampere ratings the per-unit impedance of each unit may be based on one of the several ratings of the units involved. Before any calculations can be undertaken the per-unit impedances must all be referred to the same base volt-amperes.

Suppose the following details refer to two units A and B.

	kVA	Per-unit impedance	Full-load current
A	S_A	$Z_{A\,pu}$	$I_A = \frac{S_A}{V_{fl}}$
B	S_B	$Z_{B\,pu}$	$I_B = \frac{S_B}{V_{fl}}$

Following eqn. (9.77) the per-unit value of any quantity A to $A_{base\,1}$ is

$$A_{pu\,1} = \frac{A}{A_{base1}} \tag{9.77a}$$

Similarly the per-unit value of A to a second base value is

$$A_{pu\,2} = \frac{A}{A_{base2}} \tag{9.77b}$$

Combining the above equations,

$$A_{pu\,2} = A_{pu\,1}\frac{A_{base1}}{A_{base2}} \tag{9.82}$$

Let Z_{Apu}' be the per-unit value of Z_A referred to base C; i.e.

$$Z_{Apu}' = Z_{Apu}\frac{Z_{base\,A}}{Z_{base\,C}}$$

$$= Z_{Apu}\frac{V_{fl}}{I_A}\frac{I_C}{V_{fl}}$$

or

$$Z_{Apu}' = Z_{Apu}\frac{S_C}{S_A} \tag{17.9}$$

Similarly the per-unit impedance of B referred to base C is

$$Z_{Bpu}' = Z_{Bpu}\frac{S_C}{S_A} \tag{17.10}$$

EXAMPLE 17.1 A generating station is laid out as shown in Fig. 17.5(*a*). The ratings and per-unit reactances of the different elements are as indicated. Calculate the volt-amperes and the current fed into the following symmetrical 3-phase short circuits: (*a*) at a busbar section, e.g. at D; (*b*) at the distant end of a feeder, e.g. at I.

Refer to a base of 5 MVA.

$$\text{Reference value of generator per-unit reactance} = 0{\cdot}30 \times \frac{5}{10} = 0{\cdot}15\,\text{p.u.}$$

$$\text{Reference value of reactor per-unit reactance} = 0{\cdot}10 \times \frac{5}{10} = 0{\cdot}05\,\text{p.u.}$$

(*a*) The equivalent reactance diagram for a fault on a busbar section is shown in Fig. 17.5(*b*).

$$\text{Equivalent per-unit reactance of parallel branch} = \frac{0{\cdot}2 \times 0{\cdot}2}{0{\cdot}2 + 0{\cdot}2} = 0{\cdot}1\text{ p.u.}$$

$$\text{Equivalent per-unit reactance of branch 1} = 0{\cdot}1 + 0{\cdot}05 = 0{\cdot}15\,\text{p.u.}$$

$$\text{Total equivalent per-unit reactance} = \frac{0{\cdot}15 \times 0{\cdot}15}{0{\cdot}15 + 0{\cdot}15} = 0{\cdot}075\,\text{p.u.}$$

$$\text{Short-circuit MVA} = \frac{5}{0{\cdot}075} = \underline{\underline{66{\cdot}7\text{ MVA}}}$$

$$\text{Short-circuit current} = \frac{66{\cdot}7 \times 10^6}{\sqrt{3} \times 6{\cdot}6 \times 10^3} = \underline{\underline{5{\cdot}83\text{ kA}}}$$

(*b*) The equivalent impedance diagram for a fault at the distant end of a feeder is shown in Fig. 17.5(*c*).

Total equivalent per-unit impedance $= (0{\cdot}06 + j0{\cdot}205)$ p.u.
Magnitude of per-unit impedance $= 0{\cdot}214$ p.u.

$$\text{Short-circuit MVA} = \frac{5}{0{\cdot}214} = \underline{\underline{23{\cdot}4\text{ MVA}}}$$

$$\text{Short-circuit current} = \frac{23{\cdot}4 \times 10^6}{\sqrt{3} \times 11 \times 10^3} = \underline{\underline{1{\cdot}24\text{ kA}}}$$

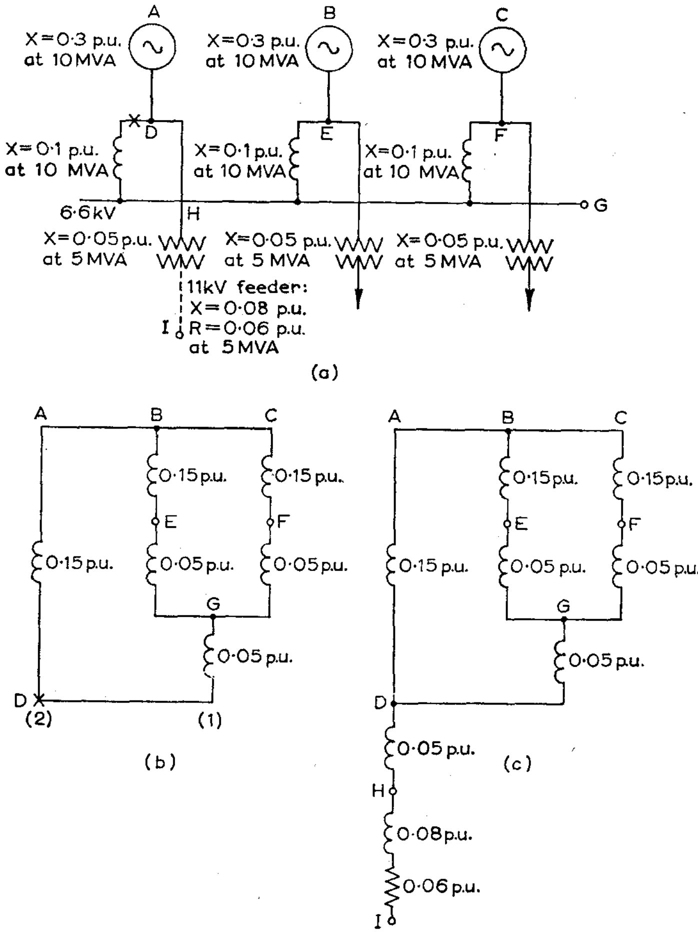


Fig. 17.5

17.5 Limiting Value of Short-circuit VA using Tie-bar Reactors

Consider a generating station having N busbar sections connected on the tie-bar system. Let each section have a generating capacity of S volt-amperes with internal per-unit reactances X_{gpu} on a basis of S volt-amperes. Let each tie-bar reactor have a per-unit reactance of X_{rpu} also on a basis of S volt-amperes. Assume a symmetrical 3-phase short-circuit occurs in a feeder connected to the busbars of section 1, as indicated by X in Fig. 17.6.

Per-unit reactance to fault from 1st generator $= X_{gpu}$

Short-circuit VA fed into fault by 1st generator $= \dfrac{S}{X_{gpu}}$

Per-unit reactance to fault from remaining $(N - 1)$ generators

$$= X_{rpu} + \frac{X_{rpu} + X_{gpu}}{N-1} = \frac{X_{gpu} + NX_{rpu}}{N-1}$$

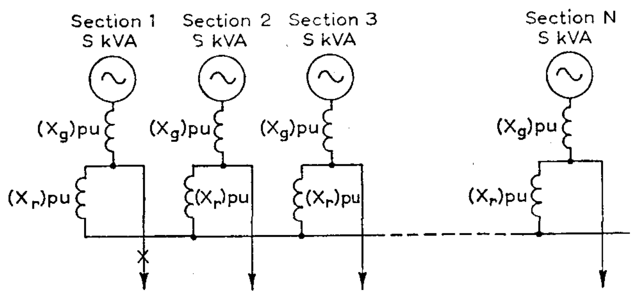


Fig. 17.6 CALCULATION OF SHORT-CIRCUIT VOLT-AMPERES ON TIE-BAR SYSTEM

Short-circuit VA fed into fault by $(N - 1)$ generators

$$= \frac{S(N-1)}{X_{gpu} + NX_{rpu}}$$

$$\text{Total short-circuit VA} = S\left(\frac{1}{X_{gpu}} + \frac{N-1}{X_{gpu} + NX_{rpu}}\right)$$

Therefore

$$\operatorname*{Lt}_{N\to\infty} (\text{s.c. VA}) = S\left(\frac{1}{X_{gpu}} + \frac{1}{X_{rpu}}\right) \tag{17.11}$$

This is the short-circuit volt-amperes at the fault if N is very large. A smaller number of generators would reduce the short-circuit volt-amperes. If the circuit-breaker on the faulty feeder has the rating derived from eqn. (17.11), then no matter how many extra generators and reactors are added, the circuit-breaker rating will remain adequate. This is a useful property of the tie-bar arrangement. With the ring arrangement extra generators may not be added without changing the existing circuit-breakers or increasing the existing reactance.

EXAMPLE 17.2 The busbars of a generating station are to be divided into three sections by the use of three reactors. A 60 MVA generator having 0·15 p.u.

leakage reactance is to be connected to each busbar section. Determine the minimum value of reactor reactance, in ohms, if the maximum MVA fed into a symmetrical 3-phase short-circuit at a section is to be 500, (*a*) if the three reactors are connected to a common tie-bar, and (*b*) if the three reactors are ring connected. The busbar voltage is 22 kV.

(*a*) Tie-bar Reactors (Fig. 17.7(*a*)).

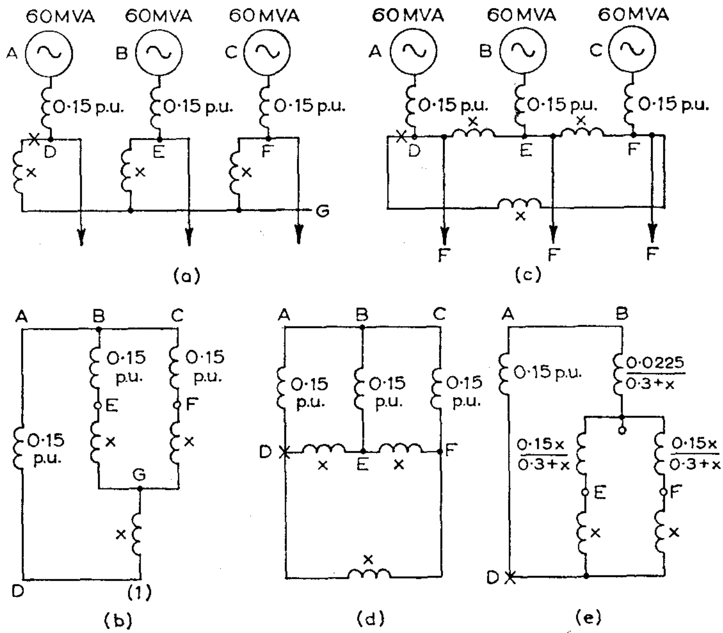


Fig. 17.7

The equivalent reactance diagram is shown in Fig. 17.7(*b*).

Let x be the per-unit reactance of each reactor on a 60 MVA base.

$$\text{Equivalent per-unit reactance of parallel branch} = \frac{0{\cdot}15 + x}{2}\ \text{p.u.}$$

$$\text{Equivalent per-unit reactance of branch 1} = x + \frac{0{\cdot}15 + x}{2}$$

$$= \frac{0{\cdot}15 + 3x}{2}\ \text{p.u.}$$

$$\text{Total equivalent per-unit reactance} = \frac{0{\cdot}15 \times \dfrac{0{\cdot}15 + 3x}{2}}{0{\cdot}15 + \dfrac{0{\cdot}15 + 3x}{2}}\ \text{p.u.}$$

$$= \frac{0{\cdot}15(0{\cdot}05 + x)}{0{\cdot}15 + x}\ \text{p.u.}$$

$$\text{Short-circuit MVA} = \frac{0{\cdot}15 + x}{0{\cdot}15(0{\cdot}05 + x)} = 500$$

since the short-circuit MVA is not to exceed 500 MVA. Solving gives

Per-unit reactance of each reactor, $x = 0{\cdot}35$ p.u.

$$\text{Full-load current, } I_{fl} = \frac{60 \times 10^6}{\sqrt{3} \times 22 \times 10^3} = 1{,}580\,\text{A}$$

$$\text{Reactance of each reactor, } X = \frac{V_{ph}}{I_{fl}} = \frac{22 \times 10^3}{\sqrt{3} \times 1{,}580} \times 0{\cdot}35 = \underline{\underline{2{\cdot}8\,\Omega}}$$

(*b*) Ring Reactors (Fig. 17.7(*c*)).
The equivalent reactance diagram is shown in Fig. 17.7(*d*).
Let x be the equivalent per-unit reactance of each reactor on a 60 MVA base. To obtain an expression for the total equivalent per-unit reactance, the mesh BCFE may be replaced by the equivalent star.
Let the star-point of the equivalent star be O.

$$X_{BOpu} = \frac{0{\cdot}15 \times 0{\cdot}15}{0{\cdot}15 + 0{\cdot}15 + x} = \frac{0{\cdot}0225}{0{\cdot}30 + x}\,\text{p.u.}$$

$$X_{FOpu} = \frac{0{\cdot}15x}{0{\cdot}15 + 0{\cdot}15 + x} = \frac{0{\cdot}15x}{0{\cdot}30 + x}\,\text{p.u.}$$

$$X_{EOpu} = \frac{0{\cdot}15x}{0{\cdot}15 + 0{\cdot}15 + x} = \frac{0{\cdot}15x}{0{\cdot}30 + x}\,\text{p.u.}$$

The equivalent reactance diagram is redrawn in Fig. 17.7(*e*).

$$X_{BDpu} = \frac{1}{2}\left(\frac{0{\cdot}15x}{0{\cdot}30 + x} + x\right) + \frac{0{\cdot}0225}{0{\cdot}30 + x} = \frac{x + 0{\cdot}15}{2}$$

$$\text{Total equivalent per-unit reactance} = \frac{0{\cdot}15 \times \dfrac{x + 0{\cdot}15}{2}}{0{\cdot}15 + \dfrac{x + 0{\cdot}15}{2}}$$

$$= \frac{0{\cdot}15(x + 0{\cdot}15)}{x + 0{\cdot}45}\,\text{p.u.}$$

$$\text{Short-circuit MVA} = \frac{60(x + 0{\cdot}45)}{15(x + 0{\cdot}15)} = 500\,\text{MVA}$$

since the short-circuit MVA is not to exceed 500 MVA. Hence

Per-unit reactance of each reactor, $x = 1{\cdot}05$ p.u.

As before, the full-load current is 1,580 A; hence

$$\text{Reactance of each reactor} = \frac{22 \times 10^3}{\sqrt{3} \times 1{,}580} \times \frac{105}{100} = \underline{\underline{8{\cdot}4\,\Omega}}$$

Note. In this particular problem, it is not absolutely necessary to use the delta-star transformation, since the symmetry of paths BED and CFD in Fig. 17.14(*d*) will mean that there is no current in the reactor EF.

17.6 Principles of Arc-extinction in Circuit-breakers

The calculation of the currents fed into symmetrical 3-phase short-circuits made in the previous sections, while giving an indication of the duty to which a particular circuit-breaker may be subjected, ignores the fact that there may be considerable asymmetry in the short-circuit current due to the presence of a d.c. component. It was seen in Section 6.4 that, when a short-circuit occurs in a circuit whose resistance is negligible compared with the inductive reactance, which is usually the case in transmission networks, the resulting short-circuit current has a d.c. component except when the short-circuit occurs at the instant at which the circuit voltage is a maximum. This d.c. component has a maximum value when the short-circuit occurs at the instant at which the circuit voltage is zero. Since in a 3-phase system there are six voltage zeros per cycle, it is certain that there will be considerable asymmetry in the current flowing in at least one of the phases. In considering the operation of circuit-breakers it is therefore necessary to take account of this asymmetry.

Fig. 17.8 indicates the process of arc extinction in a circuit-breaker. Fig. 17.8(*a*) shows the alternator line-to-neutral voltage. This voltage is shown as having constant amplitude, but it may be subjected to some decrement due to alternator armature reaction until the arc is extinguished, when the alternator voltage will slowly recover. Whether or not this effect is noticeable depends on whether the short-circuit endures for a sufficient length of time and on the distance of the fault from the alternator. If this distance is great (say at the distant end of a feeder), the alternator voltage decrement will not be marked.

In Fig. 17.8 it is assumed that the short-circuit occurs when the alternator phase voltage is passing through zero. As has been seen, this gives rise to maximum asymmetry when the circuit reactance is much greater than the resistance. The maximum peak current is about 1·8 times the peak symmetrical current, i.e. about $1{\cdot}8 \times \sqrt{2}$, or 2·55 times the r.m.s. value of the symmetrical short-circuit current. The d.c. component rapidly dies away, a typical value of decrement factor being 0·8 per half-cycle. Thus, under this decrement factor, the d.c. component will have fallen to about 30 per cent of its initial value after $2\frac{1}{2}$ cycles.

Fig. 17.8(*c*) shows the voltage across the contacts of the circuit-breaker. This is zero until the instant of separation. The voltage across the contacts after separation is the arc drop. The arc extinguishes each half-cycle and the principle of arc extinction in an a.c. circuit-breaker is to permit the arc to interrupt itself at a current zero. Whether or not the arc will restrike after a current zero depends on whether the insulation strength of the medium between the contacts (usually the medium is air or oil) builds up more rapidly or less rapidly than the voltage across the contacts. The build-up of insulation strength depends largely on the speed and thoroughness

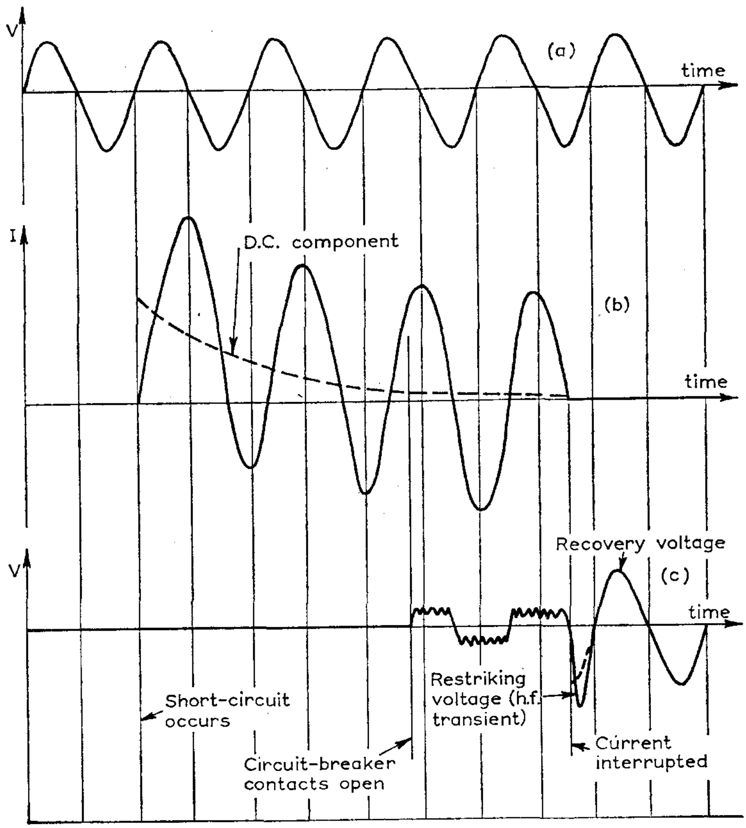


Fig. 17.8 INTERRUPTION OF AN ASYMMETRICAL SHORT-CIRCUIT CURRENT

(*a*) Alternator line-to-neutral voltage
(*b*) Short-circuit current
(*c*) Voltage across circuit-breaker contacts

with which the ionized gas (caused by the heat and electronic bombardment in the arc) is removed and replaced by un-ionized air or oil.

When the arc is finally extinguished, there is an extremely rapid rise of voltage across the circuit-breaker contacts. This voltage, which is called the *restriking voltage*, is a high-frequency transient voltage and is caused by the rapid redistribution of energy between the magnetic and electric fields associated with the plant and transmission lines of the system.

The *recovery voltage* is the 50 Hz voltage that appears across the circuit-breaker contacts when final extinction takes place and is approximately equal to the alternator phase voltage. Since the power factor of the fault circuit is low, the instantaneous recovery voltage will be almost equal to the maximum value of the alternator phase voltage.

If the alternator voltage has been subjected to decrement due to armature reaction, then after final extinction of the arc has taken place there will be a relatively slow growth in the value of recovery voltage. This is not shown in Fig. 17.8(*c*).

17.7 Rating of Circuit-breakers

It was seen in Section 17.6 that, when a symmetrical 3-phase short-circuit occurs in a transmission system, there will be a considerable asymmetry in the short-circuit current of at least one phase. The more rapidly a circuit-breaker operates after the occurrence of a short-circuit, therefore, the more onerous may be its duty. If the operation of the circuit-breaker is delayed for a few cycles, the asymmetry will be considerably reduced. On the other hand, the longer the short-circuit persists, the greater is the chance of synchronous plant losing synchronism and causing a serious interruption to the supply. Therefore circuit-breaker action is made as rapid as possible even though this may tend to make their duty more onerous.*

It is normal practice to specify the *rupturing capacity* of circuit breakers in kilovolt-amperes or megavolt-amperes. This practice is well established but may be criticized as not being logical, since the breaking capacity in megavolt-amperes is obtained from the product of short-circuit current and recovery voltage. While the short-circuit current is flowing, however, there is only a small voltage across the circuit-breaker contacts (the recovery voltage does not appear across them until final extinction has taken place). Thus the MVA rating is the product of two quantities which do not exist simultaneously in the circuit-breaker. It would appear more logical to have a current rather than an MVA rating for circuit-breakers. The agreed international standard method of specifying circuit-breaker rupturing capacity is defined as a rated symmetrical current at a rated voltage.

* Asymmetry in the short-circuit current may actually relieve circuit-breaker duty since current zeros occur closer to voltage zeros so that the value of the fundamental component of recovery voltage is reduced.

The symmetrical breaking current of a circuit-breaker is the current which the circuit-breaker will interrupt at a power factor of 0·15 for ratings up to 500 MVA and a power factor of 0·3 for ratings of 750 MVA or upwards with a recovery voltage of 95 per cent normal voltage.

The asymmetrical breaking current is the current the circuit-breaker will interrupt when there is asymmetry in one of the phases. It is assumed, for purposes of proving the capacity of a circuit-breaker, that the asymmetrical d.c. component is 50 per cent of the maximum value of the a.c. component. Normal British practice is to make the a.c. component equal to the rated symmetrical current. With a decrement factor of 0·8 per half-cycle, the maximum asymmetrical d.c. component, which is initially 80 per cent of the maximum value of the a.c. component, will have fallen to 50 per cent of this value in a little over one cycle.

The r.m.s. value of an asymmetrical current having a 50 per cent d.c. component is approximately 1·25 times that of the a.c. component. If the a.c. component (I_{ac}) is equal to the rated symmetrical component, then the asymmetrical breaking current (I) will be equal to 1·25 times the rated symmetrical breaking current. This may be shown from eqn. (5.19), since

$$I = \sqrt{\left(I_{dc}^2 + \frac{I_m^2}{2}\right)}$$

where I_m is the peak value of the a.c. component. Therefore

$$I = \sqrt{(0{\cdot}25 I_m^2 + 0{\cdot}5 I_m^2)} = 0{\cdot}866 I_m \approx 1{\cdot}25 I_{ac}$$

17.8 Types of A.C. Circuit-breaker

It is impossible here to give more than an outline of the many types of a.c. circuit-breaker in use. Circuit-breakers may be divided into (i) those which do not incorporate some form of arc control, and (ii) those which do. The latter class may be subdivided into (ii*a*) self-blast arc-control circuit-breakers, and (*iib*) circuit-breakers in which the arc control is provided by mechanical means external to the circuit-breaker. Fig. 17.9 shows a classification of circuit-breakers in the form of a diagram.

Circuit-breakers which do not have a form of arc control may be either plain air-break or plain oil-break. In these types the contacts separate either in air or in oil. In the plain oil-break circuit-breakers some assistance to arc extinction is afforded by the gas bubble generated around the arc. The gas bubble, by setting up turbulence in the oil, helps to eliminate ionized arc-products from the arc path. Long and inconsistent arcing times are obtained with these types of circuit-breaker and they are only suitable for low-current low-voltage operation.

Self-blast circuit-breakers are oil circuit-breakers in which the pressure of the gas bubble set up in the oil by the arc is utilized to force fresh, un-ionized oil into the arc path, thus materially increasing the rate of rise of insulation resistance in the circuit-breaker.

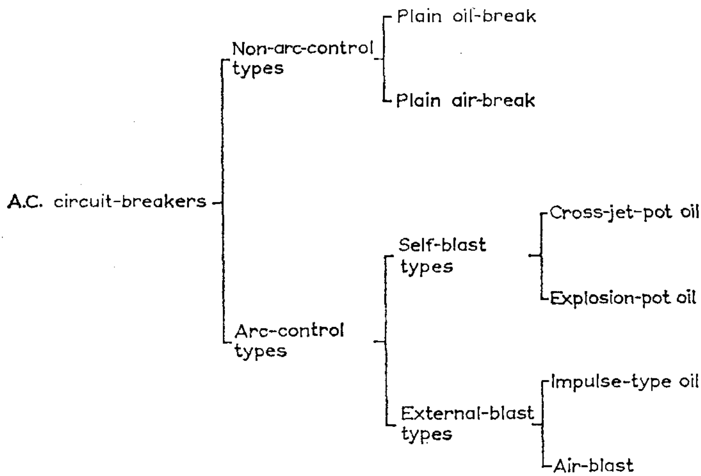


Fig. 17.9 TYPES OF CIRCUIT-BREAKER

Fig. 17.10 shows the explosion pot of a cross-jet oil circuit-breaker, which incorporates this principle of arc control. In this type of circuit-breaker, as the moving contact is withdrawn, the gas generated by the arc exerts pressure on the oil in the back passage, and as a result, when the moving contact uncovers the arc-splitting jets, fresh oil is forced across the arc path.

Circuit-breakers of this type are made with rupturing capacities of up to 2,500 MVA at 66 kV. For higher voltages and capacities multiple-break units have been developed. In these, two or more sets of cascaded contacts open simultaneously. The main difficulty associated with multiple-break units is to ensure uniform voltage distribution over the breaks. On open-circuit the voltage distribution is mainly controlled by the self-capacitance of the breaks, the insulation resistance being extremely high. Fig. 17.11 shows the opening sequence for a multi-break circuit-breaker with four cross-jet pots per phase. Four Metrosil resistors are used to control the voltages across the breaks and to eliminate over-voltages.

A difficulty experienced with self-blast oil circuit-breakers is that

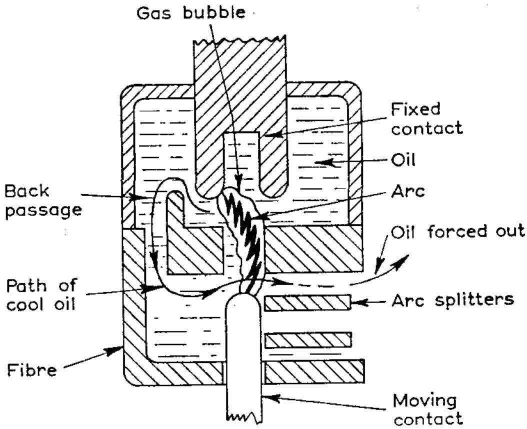


Fig. 17.10 EXPLOSION POT OF A CROSS-JET OIL CIRCUIT-BREAKER

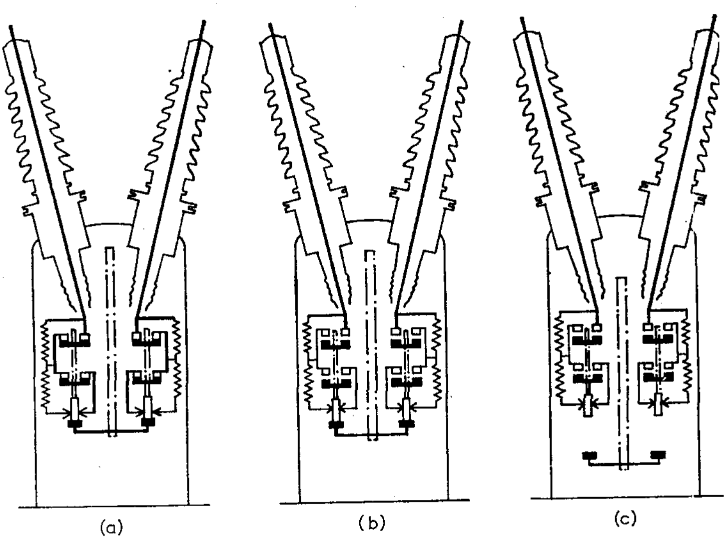


Fig. 17.11 OPENING SEQUENCE FOR A MULTI-BREAK CIRCUIT-BREAKER

(*AEI Ltd.*)

(*a*) Switch closed
(*b*) Resistors in circuit
(*c*) Switch completely open

arcing times tend to be long and inconsistent when operating against current considerably less than rated current—say about 30 per cent of rated current. This is so because the gas pressure generated is much reduced compared with that generated at rated current. Indeed the most onerous duty of the self-blast circuit-breaker is when it is called upon to break fault currents of the order of 30 per cent of its rated current.

This particular difficulty is overcome in circuit-breakers which utilize a form of arc control in which the blast is provided by external means and thus is independent of the value of fault current to be broken. Circuit-breakers using this form of arc control may be either impulse-oil circuit-breakers or air-blast circuit-breakers.

In impulse oil circuit-breakers, oil is forced across the arc-path; the necessary pressure, being produced by external mechanical means, does not in any way depend on the strength of current to be broken. Such circuit-breakers are usually multi-break and have capacitance or resistance shunts to control the voltage across the cascaded breaks. They are suitable for very-high-voltage systems of 200 kV and over.

A disadvantage attaching to all oil circuit-breakers is the risk of fire due to the inflammability of the oil. For this reason, and others, the air-blast circuit-breaker has been developed.

Air-blast circuit-breakers are similar to impulse oil circuit-breakers in that arc control, which takes the form of an air blast which may be across the arc (cross-blast), along the arc (axial blast), or radial to the arc (radial-blast), is provided by an external air compressor and is independent of the current to be interrupted.

Fig. 17.12 indicates one type of air-blast circuit-breaker. The type shown has four series breaks shunted by non-linear resistors.

Arc control is used basically to remove ionized gas, which acts as a conductor rather than an insulator, from the arc path. A recent development is to surround the circuit-breaker contacts with either sulphur hexafluoride (SF_6) or vacuum which makes it possible to dispense with external arc control.

In an SF_6 circuit-breaker the contacts are surrounded by SF_6 gas at a pressure of 45 lb/in.^{2} when the breaker is quiescent. Movement of the contacts is synchronized with the opening of a valve which allows additional gas to flow into the interrupter from a receiver containing gas at about 250 lb/in.^{2} The increased pressure assists in arc quenching.

The large SF_6 molecules rapidly absorb free electrons produced in the arc path to form heavy negative ions of low mobility which are ineffective as charge carriers. Rapid restrike-free arc interruption is thus achieved.

A typical SF_6 circuit-breaker consists of interrupter units each capable of dealing with currents of up to 60 kA and voltages of 50 to 80 kV. A number of units is connected in series according to the system voltage.

SF_6 circuit-breakers for lower voltages sometimes operate in the same manner as oil impulse breakers where the movement of the contacts operates a piston which produces a gas flow across the arc.

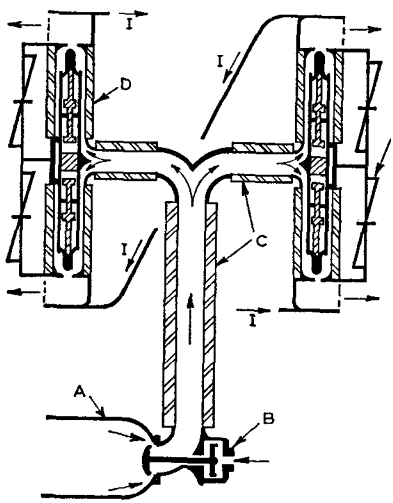


Fig. 17.12 AIR-BLAST CIRCUIT-BREAKER
(*AEI Ltd.*)

A. Air reservoir
B. Blast valve at earth potential
C. Porcelain blast tubes
D. Porcelain interrupting chamber
E. Non-linear resistor
I. Current paths

The advantageous features of SF_6 circuit-breakers are that SF_6 is odourless, non-toxic, inert and non-inflammable, and should any decomposition take place, the products (fluorine powders) possess high electric strength and are not a maintenance problem. The circuit-breakers are extremely quiet in operation, and maintenance costs are low.

Vacuum circuit-breakers are based upon the employment of a number of vacuum interrupter units in series. Each unit consists of a pair of separable contacts in a sealed, evacuated envelope of borosilicate glass. The moving contact is operated by flexible metal bellows, and it is essential that all occluded gases be removed from all components within the envelope during manufacture to maintain

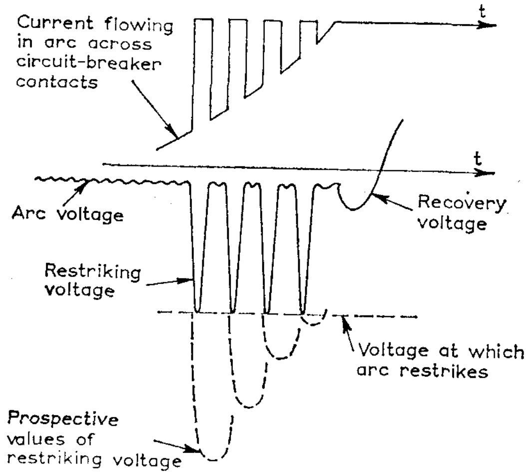


Fig. 17.13 PREMATURE EXTENTION OF ARC-CURRENT CHOPPING (SIMPLIFIED ACTION)

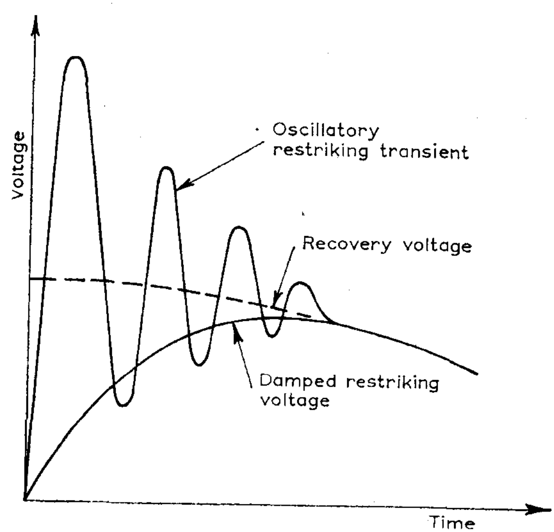


Fig. 17.14 CRITICAL DAMPING OF RESTRIKING VOLTAGE

vacuum. The high electric strength of the vacuum allows a short contact separation (0·25 in.) and rapid restrike-free interruption of the arc is achieved. A tubular shield around the contacts condenses any metallic vapours expelled from the contacts during operation, and effective interruption depends on the thoroughness and speed with which these arc products are condensed. The interrupter units are sealed into an outer envelope containing SF_6 gas. A typical 132 kV 3,500 MVA vacuum circuit-breaker uses eight interrupter units in series in each phase, each shunted by a capacitor to secure uniform voltage distribution across the units.

The advantageous features of vacuum circuit-breakers are that vacuum is non-inflammable and non-toxic, that the circuit-breaker has a light, simple, short-stroke mechanism giving high-speed performance. The relatively rapid rate of recovery of electric strength in vacuum circuit-breakers make them ideal for the interruption of large leading-power-factor currents encountered in underground cable systems. These circuit-breakers require little maintenance and are extremely quiet in operation.

17.9 Current Chopping

Current chopping is the name given to a phenomenon associated with circuit-breakers when they extinguish the arc before the natural current zero. This occurs when the circuit-breaker is breaking against a fault current considerably below its rated current. When the current is chopped in this manner the energy associated with the inductance of the circuit ($\frac{1}{2}Li^2$, where i is the instantaneous value of the current chopped) is rapidly transferred to the circuit capacitance and gives rise to a high voltage of high frequency, the frequency being determined by the inductance and capacitance of the circuit.

Current chopping occurs in all types of circuit-breaker but is probably more marked in those in which the arc control is not proportional to the fault current, i.e. in impulse oil circuit-breakers and air circuit-breakers. In air circuit-breakers, where the electric strength is usually lower than in oil circuit-breakers, the rise in voltage across the contacts may cause the arc to restrike, thus relieving the electric stress on the system insulation and preventing the transmission of a high-voltage surge. Indeed the arc may restrike several times as indicated in Fig. 17.13. In oil circuit-breakers, where electric strengths are normally higher than in air circuit-breakers, restriking may not take place and other methods of protection against the effects of current chopping such as resistance switching are necessary. Such methods may also be applied to air circuit-breakers.

17.10 Resistance Switching

The value which the restriking voltage reaches after arc extinction may be limited by the use of a shunt resistance connected across the main contacts of the circuit-breakers. The effect of the resistance is to prevent the oscillatory growth of voltage and to cause the voltage to grow exponentially up to the recovery value. Fig. 17.14 shows oscillatory growth and exponential growth when the circuit is critically damped. The dotted line represents the generated e.m.f., which is approximately at maximum value when the current zero occurs. The value of resistance required to obtain critical damping is $\frac{1}{2}\sqrt{(L/C)}$. When this technique is used some form of auxiliary contact is required to break the current through the resistor.

PROBLEMS

17.1 Explain what is meant by the percentage leakage reactance of an alternator

Two 18 MVA alternators each with a 25 per cent leakage reactance, run in parallel on a section busbar A which is connected through a 36 MVA reactor having 30 per cent reactance to a busbar B having two alternators similar to those on A. A 9 MVA feeder having 10 per cent reactance is connected to busbar A. Estimate the initial fault MVA if a short-circuit occurs between the three conductors at the far end of the feeder. (*L.U.*)

Ans. 62·8 MVA.

17.2 The busbars of a station are in two sections, A and B, separated by a reactor. Connected to section A are two 15 MVA generators of 12 per cent reactance each, and to B one 8 MVA generator of 10 per cent reactance. The reactor is rated at 10 MVA and 15 per cent reactance. Feeders are connected to the busbar A through transformers, each rated at 5 MVA and 4 per cent reactance. Determine the maximum short-circuit kVA with which oil-switches on the outgoing side of the transformer have to deal. (*H.N.C.*)

Ans. 87·1 MVA.

17.3 The 3-phase busbars of a station are divided into two sections, A and B, joined by a reactor having a reactance of 10 per cent at 20 MVA.

A 60 MVA generator with 12 per cent reactance is connected to section A and a 50 MVA generator with 8 per cent reactance is joined to section B. Each section supplies a transmission line through a 50 MVA transformer with 6 per cent reactance which steps up the voltage to 66 kV.

If a 3-phase short-circuit occurs on the high-voltage side of the transformer connected to section A, calculate the maximum initial fault current which can flow into the fault.

Explain how you would estimate the current to be interrupted by a circuit-breaker opening after 0·3 s, and show why this value would differ from the maximum value calculated. (*L.U.*)

Ans. 3,210 A.

17.4 Explain the object of (i) sectionalizing the busbars of a large generating station, and (ii) connecting reactance coils between the sections. Sketch alternative arrangements of connecting the sections and reactance coils, and discuss their relative advantages and disadvantages.

A station has three busbar sections A, B and C, which are interconnected by reactance coils (one coil between A and B, and one coil between B and C), each coil being rated at 14 per cent reactance on a basis of 60 MVA. Generators are connected as follows: one 50 MVA of 18 per cent reactance to A; one 60 MVA of 20 per cent reactance to B; one 75 MVA of 20 per cent reactance to C. Calculate the MVA which would be fed into a short-circuit on section B when all the generators are running. (*L.U.*)

Ans. 668 MVA.

17.5 A generating station has three section busbars with the following plant:

Busbar section	*Plant*	*Percentage reactance*
1	10 MVA generator	10
2	7 MVA generator	5
3	8 MVA grid transformer	8

Sections 1 and 2 are connected through a 6 MVA, 5 per cent busbar reactor, and sections 2 and 3 by a 6 MVA, 6 per cent reactor. Calculate the MVA fed into a short-circuit at the distant end of one of the outgoing feeders on section 2. The reactance of the feeder may be taken as 10 per cent on a basis of 5 MVA. Treat the grid networks as having infinite capacity. (*H.N.C.*)

Ans. 40·2 MVA.

17.6 Write a short account of the use of reactors in busbar layouts.

Generators aggregating 10 MVA and 20 per cent reactance are connected to each section busbar of a system consisting of three sections which later are connected in ring formation by a 4 MVA, 8 per cent reactor between each section. A feeder having a rating (including transformer) of 8 MVA and 10 per cent reactance is connected to one section. If a fault occurs at the distant end of the feeder, find the percentage of normal voltage at each busbar section. (*H.N.C.*)

Ans. 45 per cent; 90·1 per cent; 90·1 per cent.

17.7 Describe the "star" or "tie-bar" method of interconnecting the busbar sections in a generating station and compare it with other busbar arrangements.

A station contains 4 busbar sections to each of which is connected a generating unit of 30 MVA having 12 per cent leakage reactance, the busbar reactors having a reactance of 10 per cent.

Calculate the maximum MVA fed into a fault on any busbar section and also the maximum MVA if the number of similar busbar sections were increased to infinity. Deduce any formula used. (*L.U.*)

Ans. 422 MVA; 550 MVA.

17.8 What are the objects of sectionalizing the busbars in a large power station?

The busbars of a generating station are divided into three sections to each of which is connected a 16 MVA generator having 30 per cent reactance. The busbars are connected to a common tie-bar through 12 MVA reactors having 15 per cent reactance. To each busbar is connected a 10 MVA transformer having 10 per cent reactance and a 20 km 10 MVA feeder having a reactance of 0·7 Ω/km and negligible resistance is connected to the 33 kV secondary of one transformer.

Find the current fed into a symmetrical 3-phase short-circuit at the distant end of the feeder. (*H.N.C.*)

Ans. 513 A.

17.9 In a generating station there are four busbar sections with a 60 MVA 33 kV 3-phase generator having 15 per cent leakage reactance connected to each section. The sections are connected through a 10 per cent reactor to a common tie-bar. A 1 MVA 3-phase feeder joined to one of the busbar sections has a resistance of 60 Ω/phase and a reactance of 70 Ω/phase.

If a symmetrical 3-phase short-circuit occurs at the receiving end of the feeder, determine the MVA and also the voltage on the four busbar sections. (*L.U.*)

Ans. 11·6 MVA, 3·83 MVA; 32·3 kV, 32·7 kV.

17.10 Define the terms "symmetrical breaking current" and "asymmetrical breaking current" as applied to oil circuit-breakers. Show how these quantities are determined from oscillograms of short-circuit tests on a circuit-breaker. Explain why, on symmetrical 3-phase short-circuit tests, some initial asymmetry in the current always occurs.

A 50 MVA generator of 18 per cent leakage reactance and a 60 MVA generator of 20 per cent leakage reactance are connected to separate busbars which are interconnected by a 50 MVA reactor. Calculate the percentage reactance this reactor must possess in order that switches rated at 500 MVA may be employed on feeders connected to each of the busbars. (*L.U.*)

Ans. 7 per cent, or 0·07 p.u.

17.11 Enumerate the positions in which current-limiting reactors may be connected. What advantage does the tie-bar system have over the ring system?

A station busbar has three sections A, B and C, to each of which is connected a 20 MVA generator of reactance 8 per cent. Two similar reactors are to be connected, one between the busbar sections A and B and one between the busbar sections B and C. Calculate the percentage reactance of these reactors if the MVA fed into a symmetrical short-circuit on section A busbar is not to exceed 400. The reactors are to be rated at 10 MVA. (*H.N.C.*)

Ans. 4 per cent, or 0·04 p.u.

17.12 Three 11 kV 40 MVA alternators are connected to three sets of 33 kV busbars A, B and C by means of three 11/33 kV 40 MVA transformers. The busbars are joined by two similar reactor sets, one set being connected between A and B and the other between B and C. The reactance of each alternator is 20 per cent and that of each transformer is 6 per cent at 40 MVA.

Determine the percentage reactance of each set of reactors at 10 MVA in order that the symmetrical short-circuit on busbars A should be limited to 350 MVA. (*L.U.*)

Ans. 1·5 per cent, or 0·015 p.u.

17.13 The 33 kV busbars of a generating station are divided into three sections A, B and C, which are connected to a common tie-bar by a reactance X ohms. To section A is connected a 60 MVA generator having a leakage reactance of 15 per cent, to B a 40 MVA generator of leakage reactance 12 per cent, and to C a 30 MVA generator of leakage reactance 10 per cent.

If the breaking capacity of the circuit-breaker connected to section A is not to exceed 500 MVA, determine the minimum value of the reactance X.

Ans. 6·12 Ω.

Chapter 18

CLOSED-LOOP CONTROL SYSTEMS

In almost every sphere of human endeavour there is a need to exercise control of physical quantities. Manual control, involving a human operator, suffers from several disadvantages among which may be numbered fatigue, slow reaction time (some 0·3 s), lack of exact reproducibility, limited power, tendency to step-by-step action and variations between one operator and another. The demand for precision control of physical quantities has led to the development of *automatic control systems* or *servo systems*. It is the purpose of this chapter to examine some simple servo systems, as an introduction to a subject which is of ever-growing importance.

All precision control involves the feedback of information about the controlled quantity, in such a way that if the controlled quantity differs from the desired value an error is observed, and the control system operates to reduce this error. This type of control is called *closed-loop control* and can be either manual or automatic. In simple regulating systems there is no feedback of information, and precise control cannot be achieved. This is called *open-loop control*, since there is no feedback loop.

18.1 Open-loop Control

The operation of an open-loop regulating system may be understood by considering one or two illustrations of such systems. For example, the flow of water in a pipe may be controlled by a valve. The opening of the valve can be measured on a scale, but for any one setting the actual flow of water will depend on the available head at the inlet to the valve, and on the loading at the outlet, as well as on the valve setting. The accuracy of the setting is thus dependent on external disturbances.

Again, consider the speed control of a d.c. shunt motor by field resistance. Increasing the field resistance increases the motor speed, but at any setting of the field rheostat the actual speed will depend on the supply voltage and on the load on the machine. The rheostat cannot be calibrated accurately in terms of speed, and the system is an open-loop control. If now we connect a tachometer to the shaft, and mark on its scale the desired speed, then a human operator may adjust the field rheostat as required to keep the actual speed as near as he can to the desired speed. The operator then acts as the feedback loop, and the system has become a closed-loop system, where the control action depends on the observed error between actual and desired speed. An increased accuracy of control is thus achieved. An automatic control is achieved by replacing the operator by an error-measuring device and output controller.

18.2 Basic Closed-loop Control

The basic elements of a simple closed-loop control system are illustrated in Fig. 18.1. In this case the output of some industrial

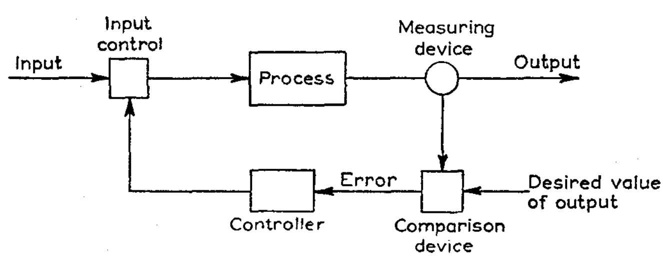


Fig. 18.1 SIMPLE CLOSED-LOOP PROCESS CONTROL

process is being controlled by a control element. The actual output from the process is measured and compared with a desired value in the comparison device. The magnitude and sense of any difference between the desired and actual values of the output is fed as an error signal to the controller which in turn actuates the correcting device in such a way as to tend to reduce the error. Note that correction takes place no matter how the error arises, e.g. by external disturbances or changes in input conditions. The control gear forms a closed loop with the process.

The essential elements of the automatic control system are thus (*a*) a measuring device, which can often be combined with (*b*) a comparison device to produce an error signal, (*c*) a controller, which normally incorporates power amplification, and (*d*) a correction device.

One very important type of control system is that in which the angular position of a shaft has to be controlled from some remote position with great accuracy. Such a system is called a *remote position control* (or r.p.c.) servo, and has applications including the automatic control of gun positions, servo-assisted steering of vehicles and ships, positioning of control rods in nuclear reactors, and automatic control of machine tools. In the following sections an electrical r.p.c. servo will be considered in more detail. In such a system, the shaft position is measured electrically, an electrical error signal is generated, amplified, and used to control an electric positioning motor.

18.3 The Summing Junction

In electrical servos it is often required to apply the sum of or difference between two or more signals to an amplifier. This can conveniently be done by means of a summing junction, as illustrated

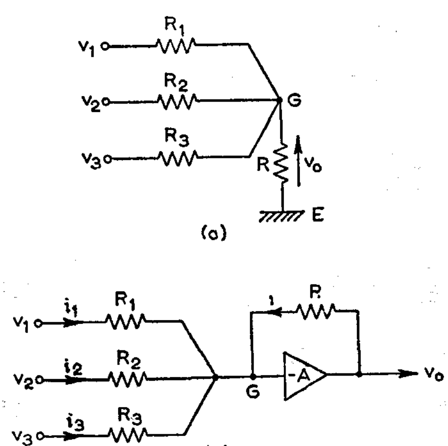


Fig. 18.2 THE SUMMING JUNCTION

in Fig. 18.2(*a*). Thus, if the free ends of the input resistors R_1, R_2, R_3 have voltages v_1, v_2, v_3 to earth, then by Millman's theorem,

$$v_0 = v_{GE} = \frac{\Sigma v_{ne} Y_n}{\Sigma Y_n} = \frac{\dfrac{v_1}{R_1} + \dfrac{v_2}{R_2} + \dfrac{v_3}{R_3}}{\dfrac{1}{R_1} + \dfrac{1}{R_2} + \dfrac{1}{R_3} + \dfrac{1}{R}} \tag{18.1}$$

The output voltage is thus dependent on the sum of the input voltages taken in proportions which depend on the ratios of the resistors. Note that, if (as is not uncommon) $R_1 = R_2 = R_3$ and $R \gg R_1$, then

$$v_0 = \tfrac{1}{3}(v_1 + v_2 + v_3) \tag{18.2}$$

For two inputs (i.e. if R_3 were disconnected) this would reduce to

$$v_0 = \tfrac{1}{2}(v_1 + v_2)$$

A more sophisticated version of the summing junction is obtained by using a high-gain d.c. amplifier as shown in Fig. 18.2(*b*). If the gain of the amplifier is $-A$ and its input impedance is very high, then

$$i = -(i_1 + i_2 + i_3)$$

and the potential of point G is $-v_0/A$ which will be very small if A is large (typically A can be of the order of 10^7). Thus G can be considered to be almost at earth potential, and is called a *virtual earth*, so that

$$i_1 = \frac{v_1}{R_1} \qquad i_2 = \frac{v_2}{R_2} \qquad i_3 = \frac{v_3}{R_3} \qquad i = \frac{v_0}{R}$$

Hence

$$\frac{v_0}{R} = -\left(\frac{v_1}{R_1} + \frac{v_2}{R_2} + \frac{v_3}{R_3}\right)$$

or

$$v_0 = -\left\{v_1 \frac{R}{R_1} + v_2 \frac{R}{R_2} + v_3 \frac{R}{R_3}\right\} \tag{18.3}$$

In the case where $R_1 = R_2 = R_3 = R$,

$$v_0 = -(v_1 + v_2 + v_3) \tag{18.4}$$

This arrangement is called an *operational amplifier*, and finds an application in analogue computers as well as in servo systems.

18.4 Measurement of Shaft Position Error by Voltage Dividers

Two methods of obtaining an electrical signal which will give a measure of the size and sense of the difference between the actual angular position of a shaft and the desired angular position will be considered. The first of these methods involves voltage dividers whose sliders are fixed to a reference and an actual output shaft respectively.

Consider the two linearly wound voltage dividers shown in Fig. 18.3, connected to a summing junction, G, though resistors which are of such high values that they give negligible loading on the voltage dividers. The slider of the reference voltage divider is set at the desired angular shaft position θ_i, so that, assuming the divider to be wound over 300°,

$$v_1 = \frac{\theta_i}{300} \times (-V)$$

The slider of the output voltage divider is connected to the output shaft, and for an output shaft angular position of θ_o,

$$v_2 = \frac{\theta_o}{300} \times (V)$$

It follows from eqn. (18.4) that

$$v_e = -\left(\frac{-\theta_i V}{300} + \frac{\theta_o V}{300}\right) = \frac{V}{300}(\theta_i - \theta_o) \tag{18.5}$$

The voltage v_e is thus proportional to the shaft error (i.e. the difference between the desired and actual shaft positions $(\theta_i - \theta_o)$ and is called the *error voltage*.

18.5 Synchros as Shaft Position Error Detectors

The synchro gives an a.c. error voltage whose amplitude is proportional to the shaft error and whose phase depends on the sense of the error. The principle of operation can be seen from Fig. 18.4. The stator houses three windings whose centre-lines are 120° apart. The rotors may be of the salient-pole type or of the wound-rotor type.

Two units are used. The transmitter has its rotor supplied from an a.c. source, the rotor shaft being set to the desired angular position. The rotor current sets up an air-gap flux which links the three stator windings and induces e.m.f.s in them according to the relative position of the rotor. The stator windings of the transmitter are linked to those of the receiver as shown in Fig. 18.4,

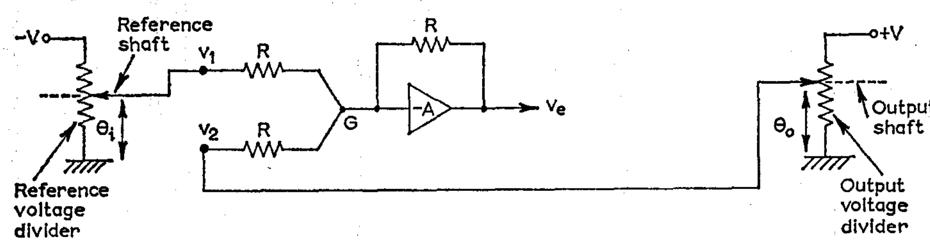


Fig. 18.3 USE OF VOLTAGE DIVIDERS FOR SHAFT ERROR MEASUREMENT

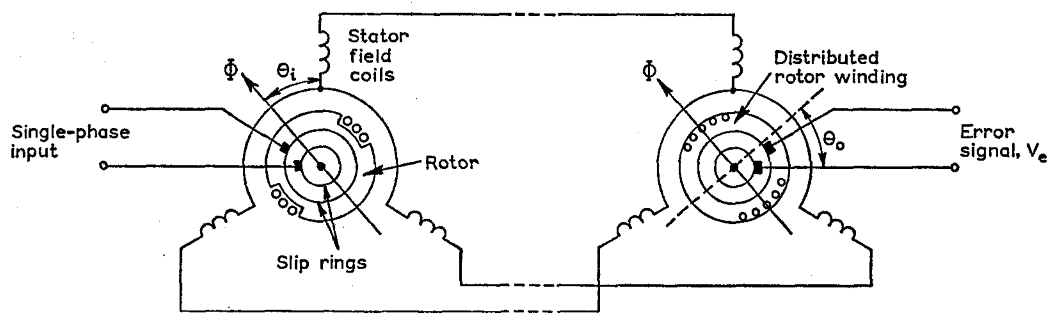


Fig. 18.4 USE OF SYNCHROS FOR SHAFT ERROR DETECTION

so that the induced e.m.f.s will set up circulating currents, which in turn cause a flux in the air-gap of the receiver. This receiver flux will have the same direction relative to the receiver stator as the transmitter air-gap flux has relative to the transmitter stator.

When the centre-line of the receiver rotor is at 90° to this flux, the e.m.f. induced in the rotor winding will be zero. For any deviation θ from the 90° position, the alternating e.m.f. induced in the receiver rotor will be

$$V_e = V \sin \theta = V \sin (\theta_i - \theta_o) \tag{18.6}$$

where V is the r.m.s. voltage induced in the receiver rotor when it links all the stator flux. For small deviations from exact quadrature between receiver flux and rotor centre-line, $\sin(\theta_i - \theta_o) \approx \theta_i - \theta_o$ and

$$V_e \approx V(\theta_i - \theta_o) \tag{18.7}$$

i.e. the error voltage is linearly related to the difference between input and output shaft positions (after allowing for the initial 90° displacement). The phase of this error voltage will be 180° different for the case when $\theta_o < \theta_i$ than it is when $\theta_o > \theta_i$, thus giving a measure of the sense of the error.

Shaft error detection is only one of several applications of these devices. As error detectors they have the advantage over voltage dividers of negligible wear, greater accuracy, and error detection over a full 360° rotation.

18.6 Small Servo Motors and Motor Drives

Electrical servos may be (*a*) entirely d.c. operated, using voltage dividers, d.c. amplifiers, and d.c. driving motors, (*b*) entirely a.c. operated using synchros, a.c. amplifiers and 2-phase a.c. driving motors, or (*c*) a.c./d.c. operated using a.c. error detection, phase-sensitive rectification and d.c. driving motors.

SPLIT-FIELD MOTOR

Small d.c. servo motors are generally of the split-field type illustrated in Fig. 18.5. Neglecting saturation and assuming a constant armature current, the output torque will be proportional to the net field current, and will reverse when this net field current reverses. The field may be fed from a push-pull amplifier stage. If the armature is not fed from a constant-current source, the build-up of armature e.m.f. with speed causes a falling torque/speed characteristic, which is equivalent to viscous-friction damping in the servo system. Approximately constant armature current can be achieved by feeding the armature through a high resistance from a high-voltage d.c. supply.

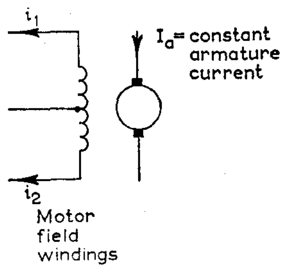


Fig. 18.5 THE D.C. SERVO MOTOR

PHASE-SENSITIVE RECTIFIER

An a.c. error signal may be used with a *phase-sensitive rectifier* (p.s.r.) to produce a d.c. error voltage. The basic operation of a p.s.r. is shown in Fig. 18.6. When there is no error voltage, each diode conducts during positive half-cycles of the reference voltage and there is no net voltage between A and B. The peak error signal,

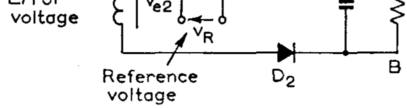


Fig. 18.6 PHASE-SENSITIVE RECTIFICATION

v_e, is arranged to be smaller than the reference voltage, v_R. When the error signal is in phase with the reference voltage, the voltage applied to D_1 during the positive half-cycles of v_R, i.e. $(v_R + v_{e1})$, is greater than that applied to D_2 $(v_R - v_{e2})$, and hence A is positive with respect to B. During the negative half-cycles of v_R both diodes remain non-conducting (since $v_R > v_e$). If the error voltage is now changed in phase by 180°, D_2 has a larger voltage applied to it during the positive half-cycles than D_1 and B is positive with respect to A. Hence the voltage between A and B gives the magnitude and sense of the error, and may be applied direct to the bases of a long-tailed pair (Section 22.11). The capacitors provide smoothing of the output signal.

TWO-PHASE SERVO MOTOR

In an a.c. servo the error signal from a synchro is fed to an a.c. amplifier, whose output feeds one phase of 2-phase motor. The phase of the voltage applied to this winding is arranged to be in quadrature with that applied from a constant reference source to the second winding of the motor (the reference winding), as shown in Fig. 18.7(*a*).

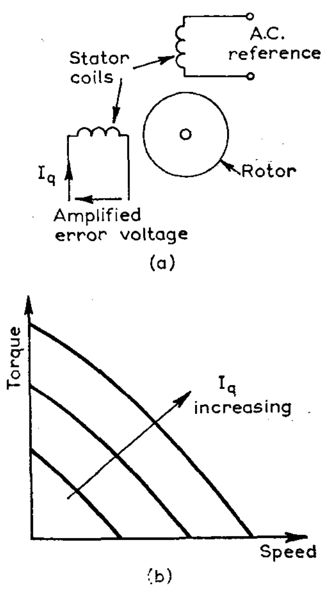


Fig. 18.7 THE 2-PHASE SERVO MOTOR

When there is no error voltage, only the reference winding is energized and the rotor is locked in position. The motor is designed to give maximum torque at standstill. The size of the output torque will depend on the magnitude of the error signal, and the direction of the torque will depend upon whether the error signal lags or leads the reference voltage by 90°. Typical characteristics are shown in Fig. 18.7(*b*).

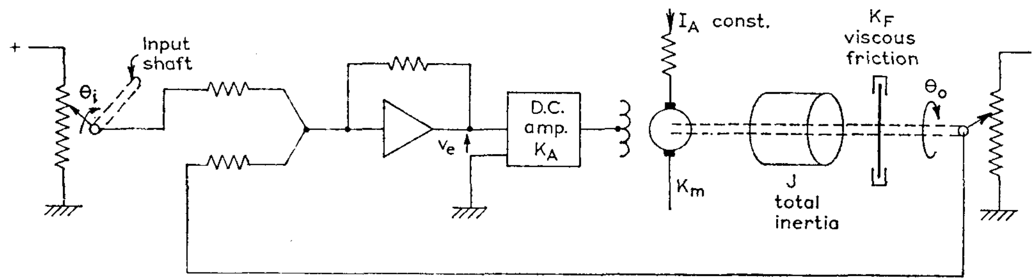


Fig. 18.8 SIMPLE TORQUE-CONTROLLED R.P.C. SERVO

The a.c. servo is not used where large power is required, because the electronic amplifier has a limited power output. In such cases rotating d.c. amplifiers (amplidyne or metadyne type) or magnetic amplifiers are used with d.c. driving motors. High-power servos will not be dealt with in this text.

Note that in both d.c. and a.c. servo motors the armatures are usually long and of small diameter in order to give a high torque/inertia ratio.

18.7 Simple Torque-controlled R.P.C. Servo

The main components of an electrical r.p.c. servo system having been briefly discussed, a simple control system can now be considered. Such a closed-loop system is shown schematically in Fig. 18.8, where the servo motor drives a load shaft (the inertia of the load and the motor being J kg-m²), and where there is viscous-friction damping (i.e. a frictional force proportional to the angular velocity of the output shaft). The driving torque produced by the motor is directly proportional to the error voltage, v_e, and is given by

$$T_D = K_A K_m v_e \quad \text{newton-metres}$$

where K_A is the amplifier *transconductance* in amperes per volt, and K_m is the motor torque constant in newton-metres per ampere of field current.

The error voltage is related to the shaft error by the equation

$$v_e = K_S(\theta_i - \theta_o) \tag{18.12}$$

where K_S is the voltage divider and summing junction constant in volts per radian of error.

Neglecting any load torque, the motor driving torque must be sufficient to overcome the inertia torque, $J(d^2\theta_o/dt^2)$, and the viscous friction torque, $K_F(d\theta_o/dt)$ (where K_F is the friction torque per unit of angular velocity), so that

$$J\frac{d^2\theta_o}{dt^2} + K_F\frac{d\theta_o}{dt} = K_A K_m v_e = K_A K_m K_S(\theta_i - \theta_o) = K(\theta_i - \theta_o) \tag{18.13}$$

where $K = K_A K_m K_S$ newton-metres per radian. Rearranging,

$$\frac{d^2\theta_o}{dt^2} + \frac{K_F}{J}\frac{d\theta_o}{dt} + \frac{K}{J}\theta_o = \frac{K}{J}\theta_i \tag{18.14}$$

Obviously, at standstill, when $d^2\theta_o/dt^2 = d\theta_o/dt = 0$, then $\theta_o = \theta_i$ and there is no error.

The response of the system (which is known as a *second-order system*) to a step change in input, θ_i, may be computed in exactly the same way as was used for the double-energy transient in Chapter 7. Rewriting eqn. (18.14) in standard form,

$$\frac{d^2\theta_o}{dt^2} + 2\zeta\omega_n\frac{d\theta_o}{dt} + \omega_n{}^2\theta_o = \omega_n{}^2\theta_i \tag{18.14a}$$

where

$$\omega_n = \sqrt{(K/J)} = \text{undamped natural angular frequency} \tag{18.15}$$

and ζ (zeta) is the *damping ratio given by*

$$\zeta = \frac{\text{Actual damping constant}}{\text{Damping constant for critical damping}} = \frac{K_F}{2\sqrt{(JK)}} \tag{18.16}$$

As shown in Fig. 18.9, there are four solutions to eqn. (18.14*a*).

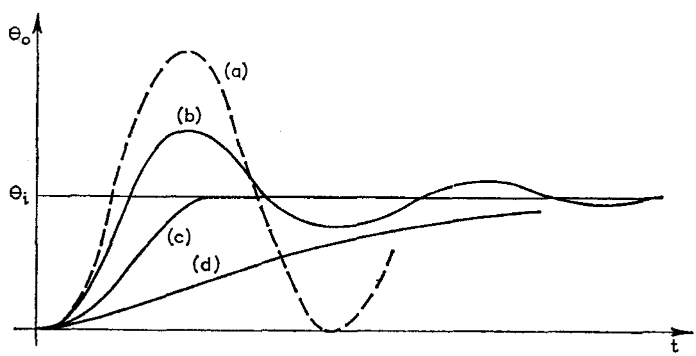


Fig. 18.9 RESPONSE OF A SECOND-ORDER SERVO TO A STEP CHANGE OF INPUT

(*a*) *Zero damping* ($\zeta = 0$; i.e. $K_F = 0$)
The response to a step input change is given by

$$\theta_o = \theta_i(1 - \cos\omega_n t) \tag{18.17}$$

The output oscillates continuously, the amplitude of the oscillations corresponding to the input step change. This condition should obviously be avoided.

(*b*) *Underdamping* ($\zeta < 1$; i.e. $K_F < 2\sqrt{(JK)}$)
The output oscillates but finally settles down to a steady value of $\theta_o = \theta_i$. The response is characterized by an overshoot (or several

overshoots) but gives a fast rise to around the value of θ_i. The actual equation of the response is

$$\theta_o = \theta_i \left\{1 - e^{-\zeta\omega_n t}\left(\cos \omega t + \frac{\zeta}{\sqrt{(1 - \zeta^2)}} \sin \omega t\right)\right\} \tag{18.18}$$

where

$$\omega = \omega_n \sqrt{(1 - \zeta^2)} \tag{18.19}$$

The time constant of the decaying oscillation is $\tau = 1/\zeta\omega_n$. The slightly underdamped response is the type usually employed for fast-acting servos, damping ratios of the order of 0·6 being common.

(*c*) *Critical damping* ($\zeta = 1$; i.e. $K_F = 2\sqrt{(JK)}$)
This condition marks the transition between the oscillatory and the overdamped solution, and the response to a step input change is

$$\theta_o = \theta_i(1 - \zeta\omega_n t\, e^{-\zeta\omega_n t} - e^{-\zeta\omega_n t}) \tag{18.20}$$

(*d*) *Overdamping* ($\zeta > 1$; i.e. $K_F > 2\sqrt{(JK)}$)
This represents a condition of slow response and is normally avoided in practice. The mathematical expression for the response is

$$\theta_o = \theta_i \left\{1 - e^{-\zeta\omega_n t}\left(\cosh \beta t + \frac{\zeta}{\sqrt{(1 - \zeta^2)}} \sinh \beta t\right)\right\} \tag{18.21}$$

where

$$\beta = \omega_n \sqrt{(\zeta^2 - 1)} \tag{18.22}$$

It should be noted that exactly the same results are obtained with and a.c. servo, or a d.c. servo with synchro error detection and a phase-sensitive rectifier.

In all servos there is a lower limit to the size of error signal which will just cause a correcting action to take place. If the system is made too sensitive, spurious operation may result from random noise inputs to the amplifier. The *dead zone* of an r.p.c. servo is the range of shaft errors over which no correcting action will take place, since the motor torque will not be sufficient to overcome stiction (static friction). This dead zone will give the limit of accuracy of the system.

18.8 Gearing

It is usually economic to design servo motors which run at a much higher speed than that required for the output shaft. A reduction gear is then used to connect the motor to the load shaft (Fig. 18.10).

If the gear ratio is $n:1$, the shaft output angular rotation, velocity and acceleration will each be $1/n$ of the corresponding input quantities. Also, assuming that there is no power loss in the gearing,

$$\text{Input power} = \text{Output power} \qquad \text{i.e.} \qquad T_1\omega_1 = T_2\omega_2$$

or

$$T_2 = nT_1 \tag{18.23}$$

where T represents torque, ω is the angular velocity, and the subscripts 1 and 2 refer to the input and output sides of the gearbox.

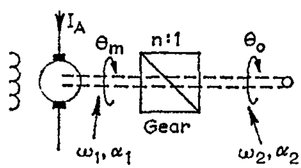


Fig. 18.10 SERVO MOTOR WITH GEAR TRAIN

The torque at the motor shaft required to overcome the motor inertia (J_m) is $J_m\alpha_1$ where α_1 is the angular acceleration at the motor. This torque at the motor shaft gives a torque of $nJ_m\alpha_1$ at the output shaft.

Hence

$$\text{Output shaft torque} = nJ_m\alpha_1 = n^2J_m\alpha_2$$

where α_2 is the output shaft acceleration; i.e. the motor inertia is equivalent to an inertia *at the output shaft* of

$$J_m' = n^2J_m \tag{18.24}$$

If the load inertia is J_L, the load acceleration, α_2, due to a motor torque T_m will be

$$\alpha_2 = \frac{nT_m}{J_L + n^2J_m}$$

where the total inertia referred to the load shaft is

$$J = J_L + n^2J_m \tag{18.25}$$

α_2 is a maximum as n varies when $n = \sqrt{(J_L/J_m)}$, and this gear ratio is said to match the inertias.

18.9 Velocity-feedback Damping

Normal viscous friction of the mechanical system is usually insufficient to provide enough damping for the satisfactory operation

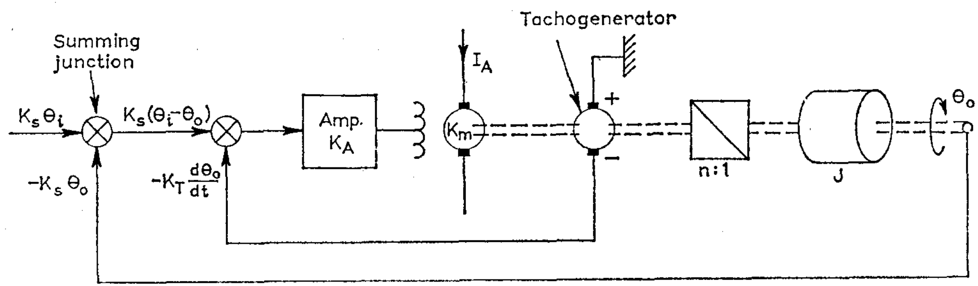


Fig. 18.11 D.C. SERVO WITH VELOCITY-FEEDBACK DAMPING

of an r.p.c. servo, and any increase in the mechanical viscous friction would involve additional power loss. The damping term in eqn. (18.14) can be readily increased, however, by *velocity feedback*, in which a second feedback loop is provided to give a negative signal at the amplifier input proportional to the shaft velocity. This signal can be obtained from a d.c. tachogenerator connected to the motor shaft, usually before the gearing, as shown in Fig. 18.11. In a.c. servos an a.c. tachogenerator can be used in the same way.

SYNCHRO ERROR DETECTION

The dynamic equation of the system is obtained as before by equating motor driving torque to the opposing torques. Neglecting mechanical viscous friction and load torque, and assuming a total inertia at the output shaft of J,

$$nK_AK_m\left\{K_S(\theta_i - \theta_o) - K_Tn\frac{d\theta_o}{dt}\right\} = J\frac{d^2\theta_o}{dt^2} \tag{18.26}$$

where K_T is the tachogenerator constant in volts per rad/s *at the motor shaft*, K_S is the error constant in volt/rad error, and it is assumed that the summing junction adds the input voltages (algebraically) directly. Rearranging,

$$\frac{d^2\theta_o}{dt^2} + \frac{n^2K_AK_mK_T}{J}\frac{d\theta_o}{dt} + \frac{nK_AK_mK_S}{J}\theta_o = \frac{nK_AK_mK_S}{J}\theta_i \tag{18.27}$$

This equation is of exactly the same form as eqn. (18.14) and will thus yield similar results. The damping can be readily varied by including a voltage divider in the tachogenerator feedback path.

EXAMPLE 18.1 An r.p.c. servo uses voltage dividers with a 300° travel and a total voltage of 30 V across them for error detection. Damping is provided by a d.c. tachogenerator, whose output is added to the shaft error voltage in an operational summing amplifier. The amplifier has a transconductance of 250 mA/V, and the motor has a torque constant of 4×10^{-4} N-m/mA and an inertia of 50×10^{-6} kg-m². The motor is coupled to the load, whose moment of inertia is 40×10^{-2} kg-m², through a 100:1 reduction gear. Calculate (*a*) the undamped natural frequency of the system, and (*b*) the tachogenerator constant in volts per 1,000 rev/min to give a damping ratio of 0·8. Neglect viscous friction.

From eqn. (18.25),

$$\text{Total inertia referred to load shaft} = (40 \times 10^{-2}) + (10^4 \times 50 \times 10^{-6})$$
$$= 90 \times 10^{-2}\,\text{kg-m}^2$$

$$\text{Voltage divider constant} = \frac{30}{300} \times \frac{360}{2\pi} = 5{\cdot}73\,\text{V/rad}$$

Hence eqn. (18.27) can be written

$$\frac{d^2\theta_0}{dt^2} + 1{,}110 K_T \frac{d\theta_0}{dt} + 72\theta_0 = 72\theta_i$$

Thus from eqn. (18.15), $\omega_n = \sqrt{72} = 8{\cdot}5$, so that

$$f_n = \frac{\omega_n}{2\pi} = \underline{\underline{1{\cdot}35\,\text{Hz}}}$$

Comparison with eqn. (18.14*a*) yields

$$2\zeta\omega_n = 1{,}110 K_T$$

Hence

$$K_T = \frac{2 \times 0{\cdot}8 \times 8{\cdot}5}{1{,}110} = 0{\cdot}0123\,\text{V per rad/s}$$

$$= \underline{\underline{1{\cdot}28\,\text{V per }1{,}000\,\text{rev/min}}}$$

18.10 Velocity Lag

In a second-order r.p.c. servo with viscous friction damping, suppose that the input shaft is rotated at a constant angular velocity $\omega_i = d\theta_i/dt$. The output shaft will continue to accelerate until it is rotating at the same angular velocity as the input. Then, since $d^2\theta_o/dt^2$ will be zero under these conditions, the steady-state equation of motion will be, from eqn. (18.13),

$$K_F \frac{d\theta_o}{dt} = K_A K_m K_S(\theta_i - \theta_o) = K_A K_m K_S \varepsilon$$

where ε is the angular difference between input and output shafts. Hence

$$\varepsilon = \frac{K_F \omega_i}{K_A K_m K_S} \tag{18.28}$$

This constant error is required to give the motor torque needed to drive the output shaft against the viscous friction loading, and is called the *velocity lag*.

For an r.p.c. servo which is stabilized by negative velocity feedback in addition to viscous friction damping, the dynamic equation is

$$J\frac{d^2\theta_o}{dt^2} + K_F \frac{d\theta_o}{dt} = nK_A K_m \left\{K_S(\theta_i - \theta_o) - K_T n \frac{d\theta_o}{dt}\right\}$$

(from eqn. (18.26)), where J is the total inertia referred to the output shaft. Under conditions of steady velocity input ($d^2\theta_o/dt^2 = 0$, $d\theta_o/dt = \omega_i$), the velocity lag is

$$\varepsilon = \theta_i - \theta_o = \frac{(K_F + n^2 K_T K_A K_m)\omega_i}{nK_A K_m K_S} \approx \frac{nK_T\omega_i}{K_S} \tag{18.29}$$

if (as is usually the case) $K_F \ll n^2 K_T K_A K_m$.

It is possible to eliminate the velocity lag in an r.p.c. servo with negative velocity feedback by arranging that the velocity feedback is removed when steady-state conditions are achieved. This is done by connecting a large capacitor in series in the velocity feedback loop. Essentially the capacitor passes any changing voltage conditions, but acts as a d.c. block to steady voltages. Thus the velocity feedback is effective only when the velocity of the output shaft is changing. This is termed *transient velocity feedback.*

EXAMPLE 18.2. An r.p.c. servo with velocity-feedback damping uses synchros as error detectors. The output of the phase-sensitive rectifier is 1·5 V/deg error, and is fed to the summing junction of an operational amplifier through a 1 MΩ resistor, the feedback resistor being also 1 MΩ. The amplifier transductance is 400 mA/V, and the motor torque at standstill is 5×10^{-2} N-m when the full field current of 81 mA flows in one half of the split field and zero in the other half. The tachogenerator output is 0·3 V per rev/s and is fed through a 2 MΩ resistor to the summing amplifier, the tacho generator being on the motor shaft. The output shaft is coupled to the motor through an 80:1 reduction gear, the total inertia at this shaft being 50×10^{-2} kg-m². Determine the system damping ratio, the magnitude of the first overshoot when the input shaft is given a sudden displacement of 10°, and the velocity lag when the input shaft is rotated at 5 rev/min.

The motor torque constant is $(5 \times 10^{-2})/81 = 6{\cdot}24 \times 10^{-4}$ N-m/mA; K_T is $0{\cdot}3/2\pi$ V per rad/s; and K_S is $1{\cdot}5 \times 360/2\pi$ V/rad.

The dynamic equation of the system is

$$J\frac{d^2\theta_0}{dt^2} = nK_AK_m\left\{K_S(\theta_i - \theta_0) - \tfrac{1}{2}nK_T\frac{d\theta_0}{dt}\right\}$$

The factor of one-half is present since the tachogenerator output is fed to the summing junction through a 2 MΩ resistor. This gives

$$\frac{d^2\theta_0}{dt^2} + 76{\cdot}2\frac{d\theta_0}{dt} + 3{,}440\theta_0 = 3{,}440\theta_i$$

Comparing with eqn. (18.14*a*),

$$\omega_n = \sqrt{3{,}440} = 58{\cdot}5 \quad \text{and} \quad 2\zeta\omega_n = 76{\cdot}2$$

Hence the damping ratio, ζ, is $\underline{\underline{0{\cdot}65}}$.

The system is thus underdamped and the response to a step input of 10° is given by eqn. (18.18) as

$$\theta_0 = 10\{1 - e^{-38\cdot1t}(\cos\omega t + 0{\cdot}86\sin\omega t)\}$$

$$= 10\left\{1 - \frac{e^{-38\cdot1t}}{\sqrt{(1 + 0{\cdot}86^2)}}\cos(\omega t - \tan^{-1}0{\cdot}86)\right\}$$

where $\omega = \omega_n\sqrt{(1 - \zeta^2)} = 58{\cdot}5 \times 0{\cdot}76 = 44{\cdot}5$.

Thus θ_0 has its first maximum when $\cos(\omega t - 40{\cdot}5°) = -1$, i.e. when

$$\omega t - \frac{40{\cdot}5 \times 2\pi}{360} = \pi \quad \text{or} \quad t = \frac{1{\cdot}23\pi}{44{\cdot}5} = \underline{\underline{0{\cdot}087\,\text{s}}}$$

Hence

$$\theta_{o_{max}} = 10\{1 + e^{-3\cdot 31}1\cdot 32\} = \underline{\underline{10\cdot 5^\circ}}$$

so that the magnitude of the first overshoot is $\underline{\underline{0\cdot 5^\circ}}$.

The velocity lag is given by eqn. (18.29) as

$$\varepsilon = \frac{nK_T\omega_i}{K_S}\text{ rad} = \frac{80 \times 0\cdot 3 \times 5}{1\cdot 5 \times 60} = \underline{\underline{1\cdot 3^\circ}}$$

18.11 Effect of Load Torque on a Simple R.P.C. Servo

So far, position control systems where the load torque is negligible have been considered. If there is a constant load torque, the motor must supply this even at standstill, and must therefore have an input. This in turn presumes that there is an error between input and output shafts. The dynamic equation for the system of Fig. 18.11 will be

$$J\frac{d^2\theta_o}{dt^2} + T_L = nK_AK_m\left\{K_S(\theta_i - \theta_o) - nK_T\frac{d\theta_o}{dt}\right\} \tag{18.30}$$

where T_L is the constant load torque.

Under steady-state conditions when the input is set at some fixed value (θ_i = costant), $d^2\theta_o/dt^2 = d\theta_o/dt = 0$, and the expression for the error is

$$\varepsilon = \theta_i - \theta_o = \frac{T_L}{nK_AK_mK_S}\quad\text{radians} \tag{18.31}$$

This error is often called the *offset*. Offset may be eliminated by altering the input signal to the servo amplifier so that it corresponds, not only to the error, but also to the error plus the time integral of the error.

18.12 Simple Speed Control—the Velodyne

The speed of the output shaft of a small servo motor can be controlled by a closed loop system similar to that used for position control. Such a system is shown in Fig. 18.12 and is commonly known as a

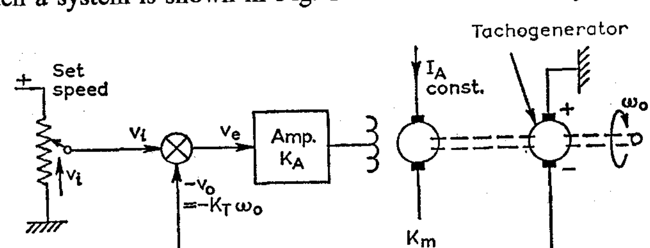


Fig. 18.12 VELODYNE SPEED CONTROL

velodyne. The output shaft drives a d.c. tachogenerator, and the voltage which it produces is compared with a set value derived from a reference voltage-divider, which can be calibrated in terms of output speed. With the polarities shown, the error voltage v_e is

$$v_e = v_i - K_T\omega_o \tag{18.32}$$

where K_T is the tachometer constant in volts per rad/s and ω_o $(= d\theta_o/dt)$ is the angular velocity of the output shaft.

The motor torque (assuming a constant armature current, an amplifier transconductance K_A, and a motor torque constant K_m) is $K_AK_mv_e$, so that, neglecting friction and loading, the dynamic equation for the system is

$$J\frac{d\omega_o}{dt} = K_AK_mv_e = K_AK_m(v_i - K_T\omega_o) \tag{18.33}$$

where J is the total inertia at the motor shaft.

Under conditions of steady output speed, $d\omega_o/dt = 0$, and this condition is fulfilled only when $v_i = K_T\omega_o$, so that the output speed is

$$\omega_o = \frac{v_i}{K_T} \tag{18.34}$$

If there is a constant load torque, T_L, coupled through an n:1 reduction gear, with the tachogenerator direct on the motor shaft, then the tachogenerator output will be $nK_T\omega_o$ (where ω_o is the output velocity). If the total inertia referred to the output shaft is J', the dynamic equation becomes

$$J'\frac{d\omega_o}{dt} + T_L = K_AK_m(v_i - nK_T\omega_o) \tag{18.35}$$

In this case the condition for steady output-shaft angular velocity (i.e. when $d\omega_o/dt = 0$) is

$$T_L = K_AK_m(v_i - nK_T\omega_o) = K_AK_m(nK_T\omega_i - nK_T\omega_o) \tag{18.36}$$

since the reference input voltage, v_i, can be calibrated in terms of the desired speed, ω_i, by eqn. (18.34), taking the gear ratio into account in this case. The difference between the desired speed, ω_i, and the actual speed, ω_o, is given by

$$\varepsilon = \omega_i - \omega_o = \frac{T_L}{nK_TK_AK_m} \tag{18.37}$$

This error is called the *droop*, and it should be noted that it is independent of the output speed. Thus at high shaft speeds it will represent a smaller percentage error than at low shaft speeds. It may be eliminated by the use of *integral-of-error compensation* as in the r.p.c. servo.

EXAMPLE 18.3 In the velodyne speed control shown in Fig. 18.12 the constants are as follows: amplifier transconductance, 200 mA/V; motor torque constant, 5×10^{-3} N-m/mA; tachogenerator constant 10 V per 1,000 rev/min. Determine the input voltage to give a speed of 2,000 rev/min. If the input setting is at half this value, find the droop when a load torque of 6×10^{-2} N-m is applied.

The tachogenerator constant in volts per rad/s is

$$K_T = \frac{10 \times 60}{1{,}000 \times 2\pi}$$

Hence, since there is no gearing, eqn. (18.34) gives

$$v_i = \omega_0 K_T = \underline{\underline{20\,\text{V}}}$$

When the input is set at 10 V, the no-load shaft speed will be 1,000 rev/min. If the load torque is now applied there must be an amplifier input voltage given by

$$v_e = \frac{T_L}{K_A K_m} = \frac{6 \times 10^{-2}}{200 \times 5 \times 10^{-3}} = 6 \times 10^{-2}\,\text{V}$$

Hence the tachogenerator output is $10 - (6 \times 10^{-2})$ V, and the actual shaft speed is $(10 - 6 \times 10^{-2} \times 1{,}000/10)$ rev/min, i.e. the droop is

$$6 \times 10^{-2} \times 100 = \underline{\underline{6\,\text{rev/min}}}$$

Note that this result can also be obtained by applying eqn. (18.37).
The speed regulation is $6/1{,}000 \times 100 = \underline{\underline{0{\cdot}6\ \text{per cent.}}}$

18.13 Some Limitations of the Simple Theory

In the simple theory developed in the preceding sections no account has been taken of any non-linearities in the system, such as the saturation of the amplifier, backlash in gearing, and stiction. Nor have the effects of actual servo-motor characteristics on the system damping, the mechanical limit of system acceleration (to keep acceleration stresses within reason), or the effect on dynamic response of the introduction of integral compensation been considered. Only velocity-feedback stabilization has been dealt with, and other forms of stabilization have been omitted. Such subjects, together with the important questions of system stability and harmonic response are the concern of full courses on servo-mechanism, as are considerations of large power systems using magnetic or rotating amplifiers.

PROBLEMS

18.1 A summing junction has three inputs A, B and C connected through resistors of $1\,M\Omega$, $2\,M\Omega$ and $250\,k\Omega$ to a star point, G. A resistor of $1\,k\Omega$ is connected between the star point and earth. The voltages of A, B and C with respect to earth are 20V, −19V and −2V respectively. Find the voltage between G and earth.

Ans. 2·49 mV.

18.2 An operational amplifier has a feedback resistor of $1\,M\Omega$ and three inputs A, B and C with input resistors of $1\,M\Omega$, $1\,M\Omega$ and $2\,M\Omega$ respectively. If the input voltages on A, B and C are 3·5V, −3·7V and −1·2V what is the output voltage?

Ans. −0·8V.

18.3 A shaft-error-measuring system employs voltage dividers with 330° of travel, having 20V d.c across them. What is the voltage per radian error?

Ans. 3·48.

18.4 The transient action of an r.p.c. servo can be described by the differential equation

$$J\frac{d^2\theta_0}{dt^2} + F\frac{d\theta_0}{dt} + K\theta_0 = 0$$

Define the terms used, and use the equation to find expressions for (*a*) the undamped natural frequency, and (*b*) the critical viscous friction coefficient, F_{crit}. Verify that these expressions are dimensionally correct. (It may be assumed that the solution to the differential equation is of the form $\theta_0 = Ae^{\alpha_1 t} + Be^{\alpha_2 t}$).

An r.p.c. servo has a moment of inertia of 0·024 kg-m² referred to the output shaft. The error-measuring element gives a signal of 1V per degree of misalignment. The motor torque constant referred to the output shaft is 10^{-2} N-m/mA, and the amplifier transconductance is 50 mA/V. Determine: (*a*) the undamped natural frequency of the system, and (*b*) the viscous friction coefficient in newton-metre-seconds [i.e. N-m per rad/s] as measured at the output shaft, which will give a damping ratio of 0·6. *(H.N.C.)*

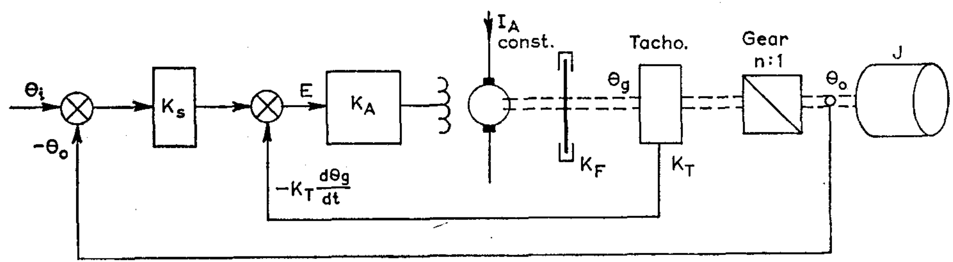
Ans. 5·5 Hz; 0·99 N-m-s.

18.5 An r.p.c. servo uses synchros for error detection, giving an output of 1V per degree error. The motor is coupled through a 100:1 reduction gear to the load. Given that amplifier transductance = 400 mA/V; motor torque constant = 5×10^{-4} N-m/mA; motor inertia = 30×10^{-6} kg-m²; load inertia = 0·5 kg-m²; calculate (*a*) the viscous friction coefficient in N-m-s at the output shaft to give a damping ratio of 0·6, (*b*) the undamped natural frequency, (*c*) the steady-state error in degrees when the input shaft is rotated at 12 rev/min. *(H.N.C.)*

Ans. 36·3 N-m-sec; 6 Hz; 2·3°.

18.6 What do you understand by the term "dead zone" as applied to a servomechanism?

A servo system has a motor with inertia 10^{-6} kg-m² matched through a gearbox to a load of inertia 0·01 kg-m². The viscous friction coefficient measured at the load is 0·64 N-m-s. The motor is excited from an amplifier and produces a torque of 10^{-4} N-m/mA of amplifier output. The error-measuring element gives 1V per degree of misalignment of input and output shafts.



θi
-θo
Ks
E
KA
$-K_T\frac{d\theta_g}{dt}$
IA
const.
KF
θg
Tacho.
KT
Gear
n:1
θo
J

Fig. 18.13

If the amplifier transconductance is 56 mA/V calculate (*a*) the undamped natural frequency of the system, (*b*) the damping factor, and (*c*) the steady-state error in degrees for an input speed of 10 rev/min. (*H.N.C.*)

Ans. 6·4 Hz; 0·4; 1·2°.

18.7 The following data relates to an r.p.c. servo with the output connected direct to the motor shaft:

Error transducer constant, 57·4 V/rad
Motor torque constant, 10^{-2} N-m/mA of amplifier output
Moment of inertia at motor shaft, 0·143 kg-m^2
Viscous friction coefficient at motor shaft, 0·13 N-m-s
Damping ratio, 1·03

Determine (*a*) the amplifier transductance, and (*b*) the undamped natural frequency of oscillation. (*H.N.C.*)

Ans. 60·7 mA/V; 2·49 Hz.

18.8 A servo motor of moment of inertia J_M drives a load of moment of inertia J_L through a reduction gear-box of ratio $n:1$. Deduce an expression for acceleration in terms of these quantities and the motor torque. Hence, determine the gear ratio for maximum acceleration of the load.

A motor is required to drive a load whose moment of inertia is 5 kg-m^2 through a gear-box of ratio 10:1. Two motors, A and B having moments of inertia 5×10^{-2} kg-m^2 and 3×10^{-2} kg-m^2 respectively are available. If both motors have the same torque-to-inertia ratio of 0·06 N-m/kg-m^2, which will give the maximum angular acceleration of the load, and what is the value of this acceleration? (*H.N.C.*)

Ans. A; 0·3 rad/s^2.

18.9 Fig. 18.13 is a block diagram of a simple remote-position servo-mechanism which drives a load through a gear-box of ratio $n:1$. The moment of inertia of the moving parts is J; K_F is the viscous friction coefficient; $K_A K_m$ is the motor torque per volt; and K_S is the error constant in volts per radian of misalignment. Establish the differential equation of the system when it is stabilized by velocity feedback, the tachogenerator having a constant of K_T volts per rad/s.

In this control system the total moment if inertia referred to the output shaft is 10 kg-m^2, the motor torque is 2 N-m/V and the error constant is 25 V/rad. If the gear-box ratio is 20 : 1 and the tachogenerator constant is 0·03 V per rad/s referred to the output shaft, calculate the natural angular frequency. Find the coefficient of viscous damping if the system is critically damped by combined viscous friction and velocity feedback and the steady-state velocity error if the input velocity is 0·2 rad/s.

Ans. 10 rad/s; 176 N-m per rad/s; 0·04 rad

18.10 A velodyne speed control uses an amplifier with a transconductance of 300 mA/V, a motor with a torque constant of 2×10^{-4} N-m/mA, and a tachogenerator with a constant of 2 V per 1,000 rev/min. The reference input and the tachogenerator feedback voltage are fed into the amplifier through equal resistors. Find (*a*) the reference input voltage for a shaft speed of 2,500 rev/min on no load, and (*b*) the droop when a constant load torque of 8×10^{-3} N-m is applied.

Ans. 5 V; 70 rev/min.

Chapter 19

ELECTRON DYNAMICS

In the following chapters it is intended to cover some of the basic concepts of the subject of electronics but it is assumed that the student is already familiar with the more elementary ideas. For this reason only a brief summary of the characteristics of electronic devices will be given.

In considering the dynamics of electron motion in a vacuum the electron will be visualized as a particle which has a negative electric charge, $-e$ coulomb. The electron is associated with a mass m, which increases as the velocity of the electron increases towards the velocity of light, but may be assumed constant at $9{\cdot}1 \times 10^{-31}$ kg up to one-third of the velocity of light. Since the negative charge is $-1{\cdot}6 \times 10^{-19}$ C, the ratio of charge to mass for a single electron is $-1{\cdot}76 \times 10^{11}$ C/kg. This huge ratio accounts for the fact that when an electron is placed in an electrostatic field the gravitational forces on it may be neglected compared with the electrostatic forces.

19.1 Acceleration of an Electron

Newton's second law of motion gives the force F required to produce an acceleration of a metres per second per second on a mass of m kilogrammes as

$$F = ma \quad \text{newtons} \tag{19.1}$$

If an electron is placed in an electrostatic field of strength E volts per metre (or newtons per coulomb), it will be acted upon by a force given by

$$F = -eE \quad \text{newtons} \tag{19.2}$$

the minus sign indicating that the force acts in the opposite direction to that of the field. The electron will, if free to move, be accelerated towards the positive electrode (Fig. 19.1). Combining eqns. (19.1)

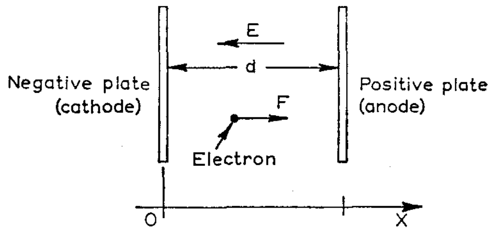


Fig. 19.1 FORCE ON AN ELECTRON BETWEEN TWO PARALLEL PLATES
E = Electric field strength = $-Fe$ = $-Vd$ F = Force on electron

and (19.2) gives the acceleration produced as

$$a = -\frac{eE}{m} \quad \text{metres/second}^2 \tag{19.3}$$

19.2 Motion of an Electron in a Plane Electrostatic Field

In considering the motion of an electron between the large plane electrodes shown in Fig. 19.1 it is assumed (i) that the space between the plates is completely evacuated, (ii) that the electrodes have a large area compared with their distance ($d$metres) apart so that the field strength at all points between the plates is V/d volts per metre, and (iii) that an electron leaves the negative plate (cathode) with zero initial velocity. If the positive X-direction is taken as shown from the cathode to the anode, then the following relations exist.

$$\text{Electric field strength} = -\frac{V}{d} \quad \text{volts/metre}$$

(negative since the cathode is more negative than the anode).

$$\text{Force on electron} = (-e) \times \left(-\frac{V}{d}\right) = e\frac{V}{d} \quad \text{newtons}$$

$$\text{Electron acceleration} = \frac{e}{m}\frac{V}{d} \quad \text{metres/second}^2 \tag{19.4}$$

Considering that the electron travels from the cathode with this uniform acceleration for a time t seconds, then

$$\text{Final velocity attained} = \frac{e}{m}\frac{V}{d}t \quad \text{metres/second} \tag{19.5}$$

$$\text{Average velocity} = \frac{1}{2}\frac{e}{m}\frac{V}{d}t \quad \text{metres/second} \tag{19.6}$$

and

$$\text{Distance moved} = \frac{1}{2}\frac{e}{m}\frac{V}{d}t^2 \quad \text{metres} \tag{19.7}$$

When the electron reaches the positive plate it will have moved a distance d metres. Let t_1 be the time taken to traverse the whole distance (transit time). Then

$$\frac{1}{2}\frac{e}{m}\frac{V}{d}t_1^2 = d \qquad \text{(from 19.7)}$$

Therefore

$$t_1 = \sqrt{\frac{d^2}{\frac{1}{2}\frac{e}{m}V}}$$

and

$$\text{Final velocity} = \frac{e}{m}\frac{V}{d}t_1 = \frac{e}{m}\frac{V}{d}\sqrt{\frac{d^2}{\frac{1}{2}\frac{e}{m}V}} = \sqrt{2\frac{e}{m}V} \tag{19.8}$$

This is seen to depend only on the potential difference V, through which the electron has moved, and on the ratio e/m.

19.3 Energy Relationships for an Electron in an Electrostatic Field

At a point in an electrostatic field an electron will possess (*a*) kinetic energy by virtue of its motion (i.e. $\frac{1}{2}mv^2$ joules), and (*b*) potential energy by virtue of its position in the field and the electrostatic forces acting upon it. When the electron is at the negative plate, or cathode, its potential energy is a maximum. The potential energy declines as the electron moves toward the positive plate, or anode.

If the electrostatic potential at a point is V volts, then the work done in bringing a positive charge of 1 coulomb to this point from a point of zero potential is V joules. The work done in bringing an electron (of charge $-e$ coulomb) up to the point will be $-eV$ joules. This is the energy stored by virtue of the electron's position.

$$\text{Potential energy of electron} = -eV \quad \text{joules} \tag{19.9}$$

If the electron moves under the action of the electrostatic forces only, it will accelerate towards the anode, losing potential energy and gaining kinetic energy but having a constant total energy, provided there are no collisions with stray gas particles.

If at one point an electron has velocity v_1 metres per second, and the point potential is V_1 volts, then

$$\text{Total electron energy} = \tfrac{1}{2}mv_1^2 - eV_1 \quad \text{joules}$$

If, at a second point, the electron velocity is v_2, and the potential is V_2, then the total energy at this point is $\frac{1}{2}mv_2^2 - eV_2$. But since the total energy must remain constant,

$$\tfrac{1}{2}mv_1^2 - eV_1 = \tfrac{1}{2}mv_2^2 - eV_2$$

Therefore

$$v_1^2 - v_2^2 = \frac{2e}{m}(V_1 - V_2) \tag{19.10}$$

Often an electron starts with effectively zero velocity from a cathode with effectively zero potential, i.e. $v_1 = 0$, $V_1 = 0$. Then

$$v_2 = \sqrt{\frac{2e}{m}V_2} \quad \text{metres/second} \tag{19.8a}$$

This gives the same result as was derived in the previous section.

These particular conditions are so often met with that it is important to remember the above equation and also to realize that, in the "zero-initial-velocity earthed-cathode" case, the magnitude of the electron velocity at any point depends on the potential of that point only, e and m being assumed constant.

19.4 The Electron-volt

Compared with the energy usually associated with a single electron, the joule is inconveniently large. The more common unit in electronics is the *electron-volt* (eV).

One electron-volt is defined as the kinetic energy gained or lost by an electron when it passes, without collision, through a p.d. of 1 V.

Now, the work done when a charge of 1 C moves through a p.d. of 1 V is 1 J. Thus

$$\text{Work done when 1 electron moves through a p.d. of 1 V} = eV \text{ joules} = 1{\cdot}6 \times 10^{-19} \times 1 \text{ joule}$$

Therefore

$$1\,\text{eV} = 1{\cdot}6 \times 10^{-19}\,\text{J}$$

If an electron "falls" through a p.d. of one million volts the work done is 1 million electron-volts (1 MeV).

19.5 Electronic Conduction through a Vacuum

The dynamics of electron motion which have been considered so far has assumed that the movement of the electrons depends only on the electrode potentials and on the properties of individual electrons. This is true only so long as the number of electrons which are involved is small (e.g. the beam current of a cathode-ray tube). For most valves the simple theory is inadequate as it takes no account of the forces which the electrons in the valve exert on each other.

Consider the simple diode circuit shown in Fig. 19.2(*a*). With the

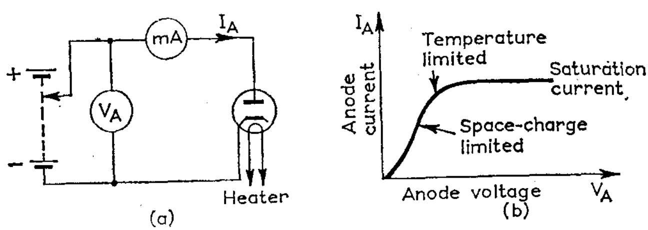


Fig. 19.2 CONDUCTION IN A THERMIONIC DIODE

heater current at a constant value the cathode will attain a constant temperature and electrons will be emitted from it at a constant rate. If the anode voltage is gradually increased, the anode current does not immediately increase to the saturation value corresponding to the cathode temperature, but rises to this value in the manner indicated in Fig. 19.2(*b*). Once saturation has been reached any alteration in the cathode temperature will alter the valve current. The operation is said to be *temperature limited.*

The more important region of the characteristic is the initial part where the anode current increases with the anode voltage. The explanation for this lies in the effect on each other of the electrons in the space between the cathode and the anode. These electrons are said to form a *space charge*, a sample space charge being shown in Fig. 19.3(*a*). For an electron on the cathode side of the space charge the electrostatic force will be reduced since the repulsion of the negative space charge counters the attraction of the positive anode. On the other hand, the force (and therefore the field strength) towards the anode will be increased for the same reason. This gives rise to the voltage distribution between cathode and anode as shown in Fig. 19.3(*b*). Unless the field strength at the cathode is reduced to zero by the space charge, all the emitted electrons must be attracted to the anode. Hence the initial portion of the characteristic is due to the field strength at the cathode being reduced to zero, and the operation in this region is called *space-charge-limited* operation. The action in the space-charge-limited region may be viewed as follows. When the cathode is heated, electrons are emitted and form a "cloud" or negative space charge round the cathode so that any further emitted electrons are repelled and fall back into the cathode. Some electrons may, however, be emitted with velocities

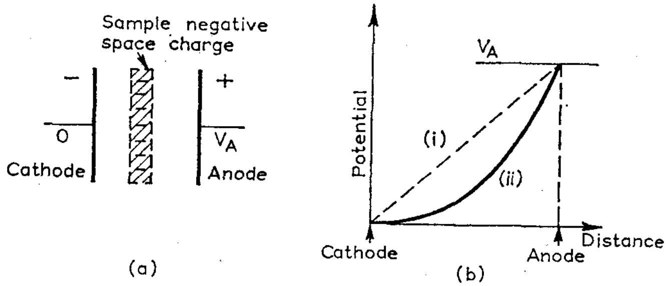


Fig. 19.3 EFFECT OF SPACE CHARGE ON VOLTAGE DISTRIBUTION IN A DIODE

(*a*) Voltage distribution without space charge
(*b*) Actual voltage distribution

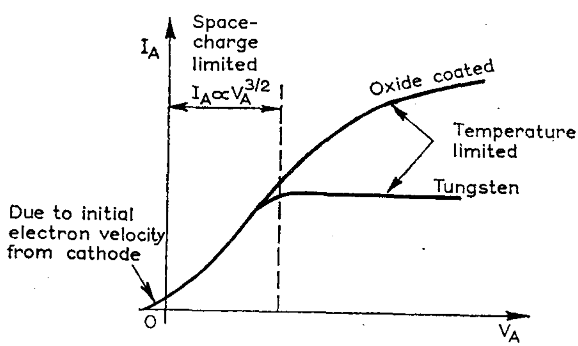


Fig. 19.4 DIODE CHARACTERISTICS

sufficient to overcome the repulsion of the space charge, and these will reach the anode, giving the small anode current for zero anode voltage shown in Fig. 19.4. When a positive potential is applied to the anode some of the electrons from the space charge are attracted to the anode, so that temporarily the space charge is reduced and further electrons will leave the cathode to join the space charge. It can be shown that in this region the anode current is proportional to the three-halves power of the anode voltage, i.e.

$$I_A \propto V_A^{3/2} \tag{19.11}$$

If the anode voltage is sufficient to draw the electrons from the space charge at the maximum rate at which they are emitted from the cathode, then the space charge disappears and the current reaches its temperature limited saturation value.

The space charge may be assumed to collect very close to the cathode.

19.6 Cathode-ray Tubes

In the cathode-ray tube a pencil-like beam of electrons is directed at a fluorescent screen, where it produces a visible spot. The beam, and hence the spot, may be directed at any part of the screen by a beam deflecting system. The construction of a typical tube with electrostatic deflexion and focusing is shown in Fig. 19.5. It con-

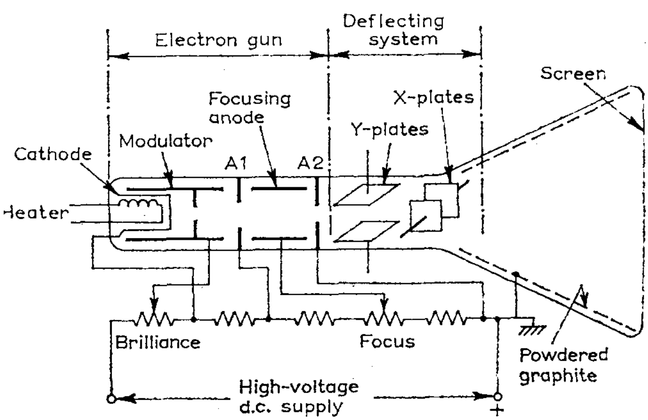


Fig. 19.5 CONSTRUCTION OF A CATHODE-RAY TUBE WITH ELECTROSTATIC FOCUSING AND DEFLEXION

sists of (*a*) an *electron gun* to produce the beam of electrons, (*b*) the *deflecting system,* and (*c*) the *fluorescent screen.* The whole is mounted in an evacuated glass envelope.

The source of electrons for the electron gun is an indirectly heated cathode in the form of a nickel cylinder, the end cap of which is oxide coated. Surrounding and extending beyond the cathode is a second cylinder (*modulator*) which is kept at a negative potential with respect to the cathode. A constriction at the end of this cylinder serves to concentrate the electrons into a rough beam before they pass through the first or accelerating anode (A1). If the modulator is sufficiently negative the beam will be cut off, so that by varying the modulator potential a brilliancy control is achieved.

The anodes A1 and A2 are in the form of discs with small central holes through which the electrons pass. Between them is a third anode in the form of a cylinder, the whole arrangement serving to focus the electrons into a narrow beam and being termed an *electron lens.*

It is usual to have the final anode A2 at earth potential so that the deflecting plates and the screen may also be at earth potential. This in turn means that the cathode is considerably negative with respect to earth, and explains why the filament supply must be well insulated from earth.

After leaving the electron gun the beam passes between two pairs of parallel deflecting plates which are mutually perpendicular. If no potential is applied between the plates, then the beam strikes the centre of the fluorescent screen. A potential applied between the X-plates produces a horizontal deflexion of the spot, and a potential between the Y-plates produces a vertical deflexion. Both deflexions may take place at the same time.

The funnel-shaped part of the tube leading to the screen is usually coated with graphite and earthed to form a shield. The screen itself is coated with a fluorescent powder such as zinc sulphide (blue glow) or zinc orthoscilicate (blue-green glow). Special powders have been developed to give black-and-white and coloured pictures for television.

For television tubes, magnetic focusing and deflexion have been widely used; electrostatic focusing and magnetic deflexion are now common; and electrostatic focusing and deflexion are coming into favour. For oscilloscopes, electrostatic focusing and deflexion are universal (apart from large-screen display osilloscopes).

19.7 Electrostatic Deflexion of an Electron Beam

Consider a beam from an electron gun passing between two charged parallel plates, d metres apart and impinging on a screen which is D metres from the centre of the plates, as shown in Fig. 19.6. If the potential of the final anode in the electron gun is V_A volts relative to the cathode, then the velocity of the electrons emerging from the gun will be, by eqn. (19.8),

$$v_Z = \sqrt{\frac{2e}{m} V_A} \quad \text{metres/second}$$

This velocity will be called the z-velocity.

If the p.d. between the parallel plates is V volts, then the field strength between them is V/d volts/metre, directed at right angles to the original direction of the electron beam. The field is assumed to be uniform between the plates; hence while the electrons are travelling between the plates they are acted on by a force of eV/d newtons (from Section 19.2), which causes an acceleration in the y-direction of $\frac{e}{m}\frac{V}{d}$ metres per second per second (eqn. (19.9)). The velocity in the z-direction is unaffected, so that the transit time (i.e. the time an electron remains between the plates) will be

$$t_1 = \frac{l}{v_z} \quad \text{seconds}$$

where l = length of plates in metres.

Hence the final velocity in the y-direction will be

$$v_y = (\text{acceleration in } y\text{-direction}) \times (\text{transit time})$$

$$= \frac{e}{m}\frac{V}{d}\frac{l}{v_z} \qquad (19.12)$$

The resultant velocity, v, after deflexion will be obtained by the vector addition of v_y and v_z:

$$v = \sqrt{(v_y^2 + v_z^2)} \qquad (19.13)$$

This velocity will make an angle $\theta = \tan^{-1}(v_y/v_z)$ with v_z. The path of the beam between the plates will be parabolic (shown dotted in Fig. 19.6), but there is no loss of accuracy in assuming that the

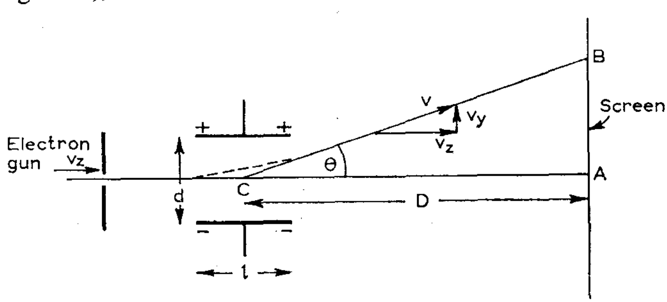


Fig. 19.6 ELECTROSTATIC DEFLEXION OF AN ELECTRON BEAM

beam suddenly deflects by an angle θ at the centre of the plates.

Assuming that the screen is flat, the resultant deflexion of the spot is AB, where

$$AB = D\tan\theta = D\frac{v_y}{v_z} = D\frac{e}{m}\frac{d}{V}\frac{l}{v_z^2} = \tfrac{1}{2}D\frac{l}{d}\frac{V}{V_A} \qquad (19.14)$$

EXAMPLE 19.1 In a given cathode-ray tube the anode voltage is 2kV. The y-deflecting plates are 3cm long and 0·5cm apart, the distance from the centre of the plates to the screen being 20cm. Determine the deflexion sensitivity in volts per centimetre.

From eqn. (19.14) the deflexion δ for a voltage V between the plates is

$$\delta = \tfrac{1}{2} \times 20 \times \frac{3}{0\cdot 5} \times \frac{V}{2{,}000} = \frac{V}{33\cdot 3}\text{cm}$$

Thus

$$\text{Sensitivity} = \frac{V}{\delta} = \underline{\underline{33\cdot 3\,\text{V/cm}}}$$

Note that in this example the transit time for an electron between the y-plates is

$$t_1 = \frac{l}{v_z} = \frac{0\cdot 03}{\sqrt{2\dfrac{e}{m}V_A}} = 1\cdot 13 \times 10^{-9}\,\text{s}$$

Thus the electron passes between the deflecting plates in an exceedingly short time. If an alternating voltage with a frequency of less than, say 1MHz is applied to the plates, then each electron will pass through the plates in such a small fraction of a cycle that during the passage of the electron the deflecting voltage may be regarded as constant at its instantaneous value.

19.8 Magnetic Deflexion of an Electron

When an electron moves across a magnetic field it is found that the path of the electron is affected by the field. The deflexion produced depends on the velocity of the electron and may be calculated by assuming that the moving electron is equivalent to a current-carrying element. Consider an electron of charge $-e$ coulombs moving with velocity v along an element of length Δl:

$$\text{Time to traverse element} = \Delta l/v \quad \text{seconds}$$

Therefore

$$\text{Rate of passage of charge} = -\frac{e}{\Delta l/v} = -\frac{ev}{\Delta l} \quad \text{coulombs/second}$$

Hence

$$\text{Equivalent current, } i = -\frac{ev}{\Delta l} \quad \text{amperes}$$

Therefore

$$i\Delta l = -ev$$

The force acting on the electron is the same as the force on the current element $i\Delta l$. If the electron is moving at right angles to the field, which has a flux density of B teslas, then

$$\text{Force on electron} = Bi\Delta l = -Bev \quad \text{newtons} \tag{19.15}$$

The minus sign is included since the direction of positive current is opposite to the direction of electron motion. The force will act in a direction which is mutually perpendicular to the electron path and to the magnetic field. The relative directions are illustrated in Fig. 19.7.

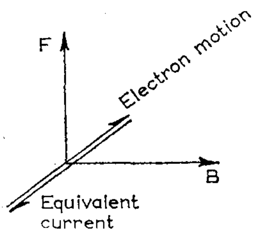


Fig. 19.7 FORCE ON A MOVING ELECTRON IN A MAGNETIC FIELD
F. Force *B.* Magnetic flux density

19.9 Motion of an Electron in a Uniform Magnetic Field

Two cases will be considered.

(*a*) *Electron passing completely through the field.* Consider an electron entering a magnetic field with a linear velocity v metres per second. Let the path of the electron be directly across the magnetic field. From eqn. (19.15) the force on the electron is Bev newtons. Since this force acts perpendicular to the direction of motion of the electron, the linear velocity of the electron will be unaffected, but the electron will describe a circular path as shown in Fig. 19.8. If the radius of the path is R metres, then since the inward magnetic force is equal to the outward centrifugal force,

$$Bev = \frac{mv^2}{R}$$

Thus

$$R = \frac{mv}{Be} \quad \text{metres} \tag{19.16}$$

By geometry the angle θ through which the electrons are deflected is the angle subtended at the centre of curvature. Hence, from Fig. 19.8,

$$\text{Deflexion of beam, } \theta = \sin^{-1}\frac{l}{R} \tag{19.17}$$

where l is the width of the magnetic field.

(*b*) *Electron moving in a large field.* Since an electron projected into a magnetic field moves with constant linear velocity along a circular path, it follows that if the extent of the field is large enough the electron will describe a complete circle. This is the basis of operation of some types of particle accelerator such as the cyclotron

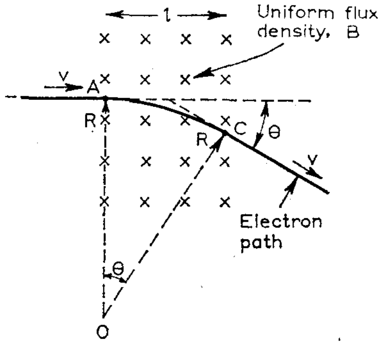


Fig. 19.8 PATH OF AN ELECTRON IN A MAGNETIC FIELD
B. Uniform magnetic field O. Centre of curvature

and the betatron, and also of some types of ultra-high-frequency valves, particularly the magnetron. It should be noted that, if the electron enters the magnetic field at an angle other than 90°, its component of velocity in the direction of the field remains unaltered, and it will describe a spiral path.

PROBLEMS

19.1 If an electron is accelerated from rest through a potential difference of 4 kV, determine the final energy of the electron and the time required for the electron to pass along an evacuated tube 10 in. long after it has attained its final velocity.

Ans. $6{\cdot}4 \times 10^{-16}$ J; $6{\cdot}73 \times 10^{-9}$ s.

19.2 Two large plane parallel electrodes are placed 1 cm apart in a vacuum and a p.d. of 100 V is applied between them. If an electron is emitted from the negative plate with zero initial velocity, calculate the time required for this electron to reach the positive plate.

Ans. $3{\cdot}37 \times 10^{-9}$ s.

19.3 Define the electron-volt.

400 V electrons are introduced at A (Fig. 19.9) into a uniform electric field of intensity 150 V/cm. If the electrons emerge at B $5{\cdot}0 \times 10^{-9}$ s later, determine (*a*) the distance AB; (*b*) the angle θ. (*H.N.C.*)

Ans. 4·95 cm; 33·8°.

19.4 Find the path of an electron moving perpendicular to a uniform magnetic field.

A 15 cm cathode-ray tube is to be magnetically deflected. The final anode voltage is 2 kV, the deflecting coil is $2\frac{1}{2}$ cm long and the distance from its centre to the screen is 15 in. Find the flux density required to deflect the spot to the edge of the screen. (*H.N.C.*)

Ans. $1\cdot06 \times 10^{-3}$T.

19.5 A cathode-ray tube has an accelerating potential of 2 kV. It is fitted with two pairs of deflecting plates, the pairs being mutually perpendicular. Each plate is 1·5 cm long, the spacing between the plates being 1 cm and the distance between

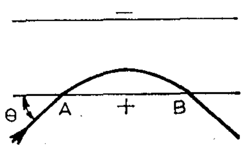


Fig. 19.9

centres of pairs being 2·5 cm. Neglecting fringing effects, calculate the length of the line produced on a fluorescent screen 30 cm distant from the centre of the nearer pair of plates when (*a*) a 50 V (r.m.s.) voltage is applied between the further pair of plates, (*b*) a 50 V (r.m.s.) voltage is applied between the nearer pair of plates, (*c*) both voltages are applied simultaneously.

Ans. 1·72 cm; 1·59 cm; 2·34 cm.

19.6 Electrons are projected with a velocity v centimetres per second into a magnetic field of strength B teslas and are observed to travel in a circular path of 10 cm radius. An electrostatic field, existing between a pair of plates 3 cm apart is now superimposed on the magnetic field at right angles to both the magnetic field and to the initial direction of motion of the electrons. If the voltage between the plates is adjusted until the electron beam passes through the combined fields without deviation, and is found to be 22·5 V, calculate the strength of the magnetic field and the velocity of the electrons at their time of entry into the field. (*H.N.C.*)

Ans. $2\cdot06 \times 10^{-4}$T; $3\cdot64 \times 10^{6}$ m/s.

Chapter 20

ELECTRONIC DEVICE CHARACTERISTICS

In order to understand electronic circuits it is necessary to know something about the characteristics of the devices which are employed in these circuits. In this chapter the operating characteristics of valves and transistors will be described. The physical basis of operation of these devices will be discussed very briefly. There will be no attempt to develop the theory of physical electronics in great detail.

Transistors and valves are called active devices, because their operation can be described in terms of equivalent constant-voltage or constant-current sources whose output is controlled by input voltages or currents. The essential basic feature of these active devices is that the controlled current or voltage can be much larger than the input current or voltage. Thus the output current or voltage variation can be an amplified version of the input signal variation. The output power is derived from an external supply and not from the controlling input signal.

20.1 The Triode

In the vacuum *triode* a heated *cathode* is surrounded by a metal *anode*, A, and a fine wire spiral, G, called the *grid* is placed between the cathode, K, and the anode, to act as the control element. The assembly is housed in an evacuated envelope. The graphical symbol for a triode is shown in Fig. 20.1(*a*). For a given positive anode voltage, the more negative the grid with respect to the cathode, the smaller will be the anode current.

When the grid–cathode voltage V_{GK} is zero, the static anode characteristic (Fig. 20.1(*b*)) is similar to that of a space-charge-limited diode. The greater the negative value of V_{GK} (i.e. the larger the negative grid bias) the larger is the anode–cathode voltage

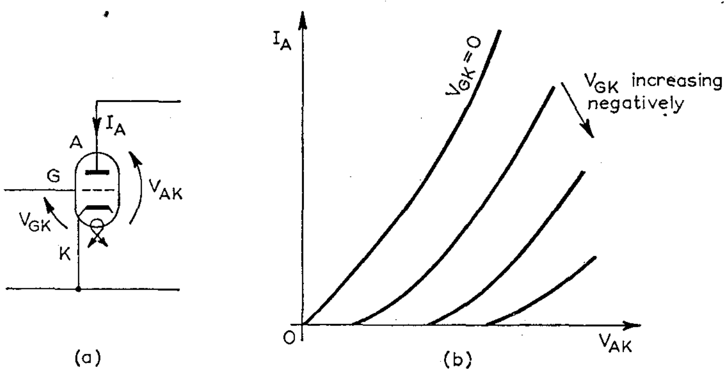


Fig. 20.1 ANODE CHARACTERISTICS OF A TRIODE

V_{AK} required to give the same anode current. Hence the static curves for various values of negative grid–cathode voltage are displaced to the right as shown in the diagram.

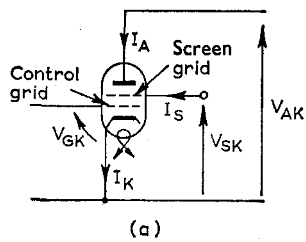
So long as the grid is maintained at a negative potential with respect to the cathode, relatively few electrons will hit it—i.e. the grid current will be negligibly small. Most of the electrons moving towards the anode will pass through the spaces in the grid wire helix.

20.2 The Tetrode—Secondary Emission

The vacuum *tetrode* was originally developed to reduce the capacitive coupling which exists between the grid and anode of a triode. To do this a second grid called the *screen grid* is introduced between the *control grid* and the anode. This grid also consists of a wire helix, through which most of the electrons comprising the cathode current can pass. It is maintained at a fixed potential and so "screens" the electrostatic field at the anode from that of the rest of the valve.

The screen potential must be positive with respect to the cathode, since otherwise there would be no electric field to draw electrons from the space charge, and hence no anode current. The graphical symbol for a tetrode is shown in Fig. 20.2(*a*).

The static anode characteristics are shown in Fig. 20.2(*b*). Because of the constant screen voltage and the screening effect on the anode, the cathode current, I_K, will be nearly constant for any negative value of grid–cathode voltage, V_{GK}. (I_K is shown for $V_{GK} = 0$ at (*b*).) This current divides between screen and anode in a manner



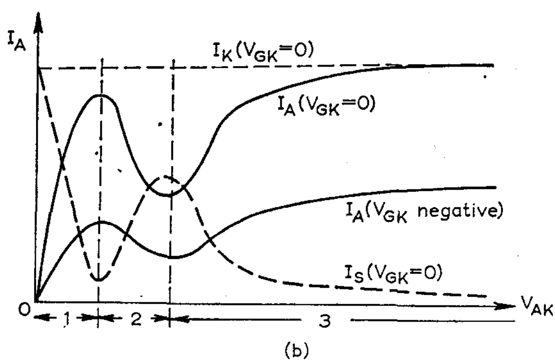


Fig. 20.2 TETRODE CHARACTERISTICS

which depends on their relative potentials. As the anode–cathode voltage is increased, three distinct stages of anode current, I_A, can be distinguished.

Stage 1. When $V_{AK} = 0$, nearly all of the cathode current passes to the screen grid so that I_A is almost zero. Any electrons passing through the screen helix are attracted back to it by the positive screen potential. As V_{AK} is increased, however, more electrons which pass through the screen helix will reach the anode, so that I_A increases and the screen current, I_S, falls by a corresponding amount.

Stage 2. When V_{AK} exceeds about 20V (the exact voltage depends on the anode material), the electrons which reach the anode bombard it with sufficient energy to cause electrons to be emitted from the anode surface. These *secondary electrons* are relatively slow moving, and will tend to be attracted to whichever electrode is more positive. So long as the screen is more positive than the anode the secondary electrons will tend to be attracted to the screen. Since the bombarding electrons can each cause more than one electron to be emitted from the anode, the anode current will fall and the screen current will correspondingly rise. As V_{AK} rises the secondary emission increases and the anode current falls further.

Stage 3. When V_{AK} exceeds V_{SK} the secondary electrons emitted from the anode will generally return to the anode, so that the anode current rises towards the value of the almost constant cathode current, and the screen current falls to a relatively low value (most of the electron stream will pass through the screen and will reach the anode). The anode current is then almost independent of the anode–cathode voltage.

Note that during stage 2 the anode exhibits a negative resistance characteristic—*an increase* in voltage causes a *decrease* in current. This is an unstable operating region. In practice, simple tetrodes were never popular owing to the negative-slope-resistance "kink" in the characteristics, but it should be noted that the curve in stage 3 gives the possibility of high amplification, since anode current in this region depends on negative grid voltage only and is almost independent of anode voltage (unlike the triode). This is discussed further in Section 20.4.

20.3 The Pentode

The *pentode* gives the advantages of the tetrode (high amplification and low electrostatic coupling between control grid and anode circuits) without the disadvantage of the negative-resistance kink in the anode characteristics. A third open-wire helical grid (the *suppressor grid*) is inserted between the screen grid and the anode. This suppressor grid is connected either internally or externally to the cathode. Hence any secondary electrons from the anode find themselves in an electric field which forces them back to the anode, so that they have no net effect on either the anode or the screen currents. The cathode current is determined almost entirely by the constant screen potential, V_{SK}, and the control grid voltage, V_{GK}. Hence the anode current is largely independent of the anode voltage above some critical value which can lie between 20V and 100V. The graphical symbol for a pentode and typical anode characteristics are shown in Fig. 20.3.

Similar "kink-free" characteristics are obtained in the *beam tetrode*, by aligning the control and screen grid helices and using the electron beam itself as a suppressor of secondary emission effects. The alignment of the grids also has the effect of reducing the screen-grid current below the value obtained in the equivalent pentode, which can be very advantageous. Beam tetrodes are often used in

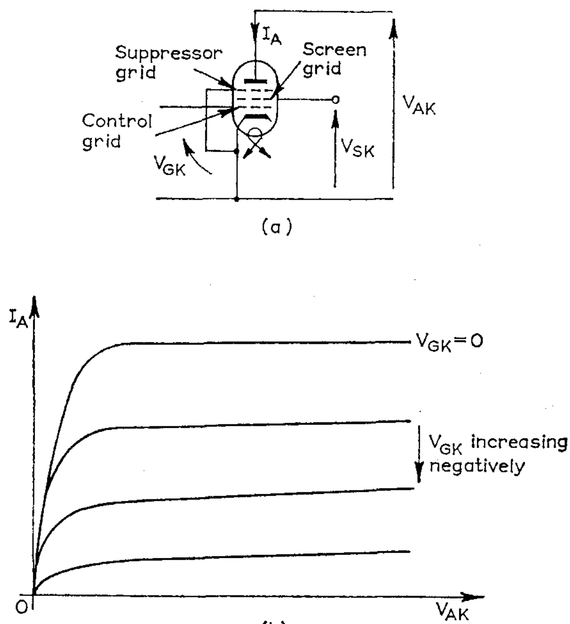


Fig. 20.3 ANODE CHARACTERISTICS OF A PENTODE

the audio-frequency power-output stages of amplifiers and in radio-frequency power amplifiers. Note that the screen–cathode voltage must be kept constant in order to obtain normal pentode or tetrode characteristics.

20.4 Simple Valve Voltage Amplifier—Load Line

The circuit of a simple amplifier using a pentode is shown in Fig. 20.4(*a*). A steady negative voltage (the grid-bias voltage V_{GB}) is applied to the grid through the grid-leak resistor, R_g. The grid carries no current, and hence if the input voltage, V_i, is zero there will be no voltage drop across R_g and $V_{gk} = -V_{GB}$. Note that this form of negative grid-bias supply from a separate source is not now used (see Section 20.5). The anode is supplied from a constant-voltage supply, V_B, through the anode load resistor R_L.

When a signal, V_i, is applied across R_g the grid voltage will vary, causing a variation, in the anode current I_A. The voltage drop across R_L will also vary, and hence the output voltage V_o will vary, since

$$V_o = V_{ak} = V_B - I_a R_L \tag{20.1}$$

If the change in V_{ak} is greater than the voltage, V_i, which caused it, the circuit is a *voltage amplifier*. Note that when V_i becomes more

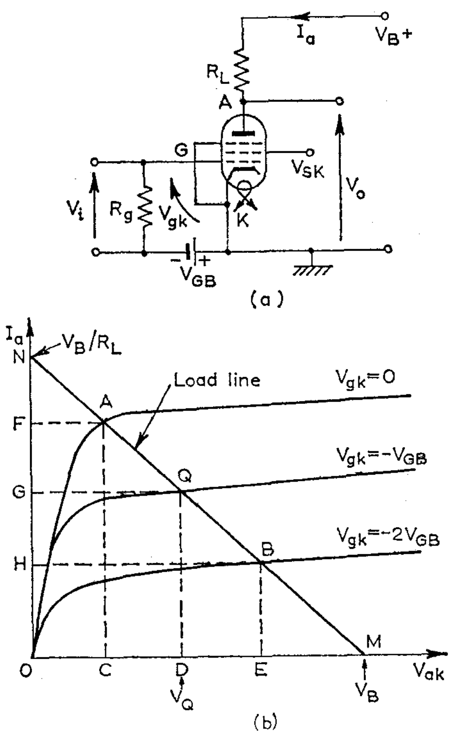


Fig. 20.4 SIMPLE VOLTAGE AMPLIFIER

positive, I_a will increase, and from eqn. (20.1), V_o will decrease, i.e. the output voltage variation is inverted with respect to the input voltage variation.

The relationship between V_{ak} and I_a must satisfy two conditions simultaneously, namely those imposed by the anode characteristics of the valve and those expressed by eqn. (20.1). If $(V_B - I_a R_L)$ is plotted on the anode characteristics, as shown in Fig. 20.4(*b*), it gives the straight line MN running from M (when $I_a = 0$, and $V_{ak} = V_B$) to N (when $I_a = V_B/R_L$, and $V_{ak} = 0$). This is called the *load line;* its slope is $-1/R_L$ amperes per volt (siemens).

For any value of grid voltage, the operating current and anode-cathode voltage must be represented by a point on that particular grid-voltage curve and also by a point on the load line, i.e. the intersection of these two curves represents the *operating point.*

Thus for $V_{gk} = 0$ the intersection is at point A, and the anode current is given by OF. The voltage across the valve is OC and the voltage across R_L is CM. With no input signal, the grid voltage is $-V_{GB}$ and so the operating point is Q, which gives the quiescent anode current (OG) and anode voltage (OD).

Suppose that an alternating input voltage of peak value $V_{i\,max}$ is applied to the grid (this maximum value is so chosen that the grid will never be positive to the cathode). The operating point moves up and down the load line between points A and B. Hence the output voltage varies between OC and OE and the voltage gain A_v is given by

$$A_v = \frac{\text{Change in output voltage}}{\text{Change in input voltage}} = -\frac{\text{CE}}{2V_{i\,max}} \tag{20.2}$$

The minus sign indicates the inversion which takes place between input and output voltage variations.

If the static characteristics are equally spaced for equal changes in V_{gk}, the output-voltage waveform will be an exact copy of the alternating input-voltage waveform. If the static characteristics are not equispaced, the output voltage will be a distorted version of the input. The grid-bias voltage is chosen so that a minimum of distortion is produced.

EXAMPLE 20.1 The anode characteristics of a pentode (Mullard EF37A) are given by the curves of Fig. 20.5(*a*) for a screen grid voltage of 100 V.

If the supply voltage is 400 V plot the load lines for anode resistances of (*a*) 50 kΩ and (*b*) 200 kΩ. (i) Choose a suitable grid bias voltage for operation with each of the above resistances. (ii) Estimate the power dissipation at the anode of the valve for each of the grid-bias/anode-resistance conbinations when there is no input grid voltage. (iii) Estimate the voltage amplification obtained with each resistance.

The load lines are drawn on the characteristic curves of Fig. 20.5, as shown. Suitable grid-bias voltages would be, say, −2 V for the 50 kΩ line (Q_1), and −3·3 V for the 200 kΩ line (Q_2).

With these grid bias values and no alternating grid voltage:
(*a*) For 50 kΩ,

Steady anode current = 3 mA
Steady anode voltage = 245 V

Therefore

Power dissipation at anode = 0·003 × 245 = 0·735 W

(*b*) For 200 kΩ,

Steady anode current = 1·2 mA
Steady anode voltage = 160 V

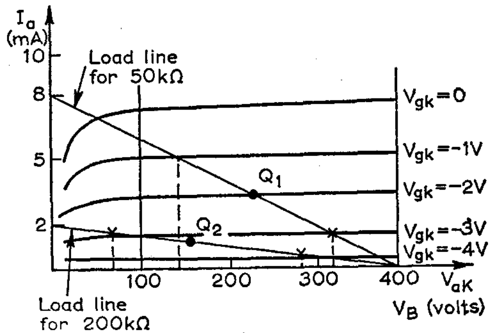


Fig. 20.5

Therefore

Power dissipation at anode = 0·192 W

To determine the voltage amplification for 50 kΩ suppose the grid voltage varies by ±1 V about the −2 V bias.

$$\text{Voltage amplification} = -\frac{325 - 150}{2} = -87{\cdot}5$$

To determine the voltage amplification for 200 kΩ suppose the grid voltage varies by ±0·5 V about the −3·3 V bias.

$$\text{Voltage amplification} = -\frac{275 - 70}{1} = -205$$

Similar conditions apply for a triode amplifier. The construction of the load line is the same as for a pentode. In this case, however, the anode current is not independent of the anode voltage, but rises and falls as the anode voltage rises and falls. Hence, if the grid voltage falls so that the anode current decreases, the anode voltage will rise, and this will tend to counteract the decrease in anode current. The result is a smaller change in anode voltage compared with a pentode, where the anode voltage hardly affects the anode current, i.e. the voltage amplification is less in the triode circuit.

20.5 Cathode Bias

The grid-cathode bias voltage (i.e. the steady or quiescent value of V_{gk}) is commonly obtained by connecting a resistor in the cathode lead as shown in Fig. 20.6.

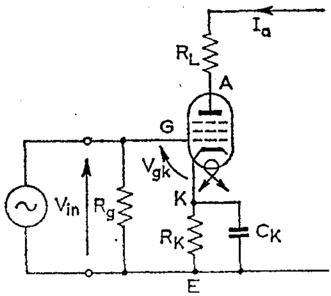


Fig. 20.6 AUTOMATIC CATHODE BIAS

Then, if the quiescent cathode current is I_Q, a direct voltage given by

$$V_{KE} = I_Q R_K \tag{20.3}$$

will be developed across R_K, and hence the grid will be more negative than the cathode by this voltage (G is connected to E through the grid-leak resistor R_g). The capacitor C_K connected across R_K serves to short-circuit the cathode to earth for alternating components of current, and so smooths out voltage variations across R_K due to any alternating component of cathode current. The calculation of the value of C_K is deferred until the next chapter, but as a rule of thumb the reactance $1/\omega C_K$ at the lowest signal frequency is normally less than $0{\cdot}1/g_m$, where g_m is a constant of the valve (see Section 20.6). Note that, if R_K is returned to the earth line, then R_K should be approximately equal to $1/g_m$.

20.6 Linear Valve Parameters

When the a.c. signals which are being dealt with by an amplifier are small, it is usually possible to operate the valve over a range of current and voltage where the static characteristics are linear and equally spaced. In this case the operation of the valve may be represented by constants (or parameters) which are assumed to remain unchanged in value over the whole operating range. This will not be true if large signals are dealt with (see Chapter 24).

Consider a section of the static anode characteristics of a valve as shown in Fig. 20.7. From these the following expressions are obtained.

1. The *anode slope resistance, r_a,* of the valve is defined as the change in anode voltage per unit change in anode current when the grid voltage is held constant. Hence from Fig. 20.7,

$$r_a = \left(\frac{\delta V_{ak}}{\delta I_a}\right)_{V_{gk}\,\text{const.}} = \frac{\text{AB(volts)}}{\text{BC(amperes)}} \tag{20.4}$$

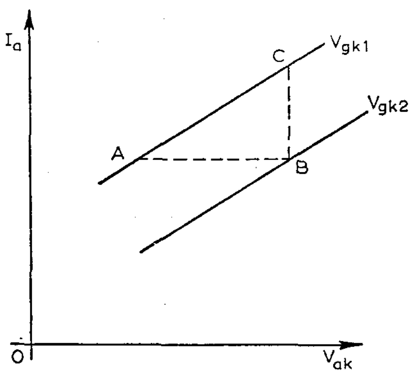


Fig. 20.7 CALCULATION OF LINEAR VALVE PARAMETERS

2. The *mutual conductance, g_m,* is defined as the change in anode current per unit change in control grid voltage when the anode voltage is held constant; i.e.

$$g_m = \left(\frac{\delta I_a}{\delta V_{gk}}\right)_{V_{ak}\,\text{const.}} = \frac{\text{BC(amperes)}}{V_{gk1} - V_{gk2}} \tag{20.5}$$

Since the anode current is normally measured in milliamperes, g_m is usually expressed in milliamperes per volt.

3. The *amplification factor, μ,* is defined as the change in anode voltage per unit change in grid voltage with the anode current held constant, i.e.

$$\mu = \left(\frac{\delta V_{ak}}{\delta V_{gk}}\right)_{I_a\,\text{const.}} = \frac{\text{AB}}{V_{gk1} - V_{gk2}} \tag{20.6}$$

From eqns. (20.4)–(20.6), the following important relation can be deduced:

$$\mu = \left(\frac{\delta V_{ak}}{\delta V_{gk}}\right)_{I_a\,\text{const.}} = \left(\frac{\delta V_{ak}}{\delta I_a}\right)_{V_{gk}\,\text{const.}} \times \left(\frac{\delta I_a}{\delta V_{gk}}\right)_{V_{ak}\,\text{const.}}$$

i.e.

$$\mu = r_a g_m \tag{20.7}$$

The mutual conductance depends mainly on the spacing and construction of the control grid and cathode. It has thus approximately the same value for triodes and pentodes, and is usually within the range 1–10 mA/V.

The anode slope resistance is the reciprocal of the slope of the static anode characteristics, and will therefore be much larger for pentodes than for triodes. Typical values range from 1 kΩ–50 kΩ for triodes and 100 kΩ–2 MΩ for pentodes.

From this and by considering eqn. (20.7) it follows that the amplification factor of a triode is much lower than that of a pentode, typical values being 20 for triodes and 1,000 for pentodes.

20.7 Anode Resistors for Voltage Amplifiers

Fig. 20.8(*a*) shows the anode characteristics of a triode with two load lines drawn for the same supply voltage V_B. Consideration of the diagram will show that the load line for the high resistance, R_H, gives a larger voltage amplification than the load line for the lower resistance, R_L. Though the amplification is not markedly different it appears that the larger the load resistance the better. However,

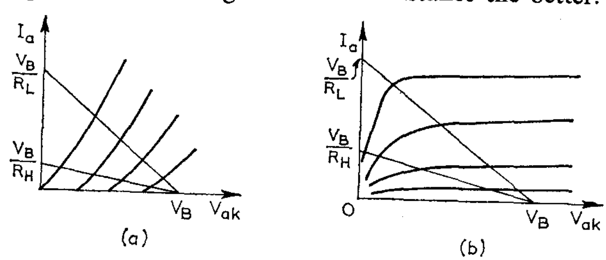


Fig. 20.8 CHOICE OF ANODE RESISTOR FROM VALVE CHARACTERISTICS
(*a*) Triode (*b*) Pentode

the load line for R_H runs through the curved and less regular parts of the characteristics so that the output voltage variations will be a less faithful reproduction of the input voltage variations than if the load line for R_L were traversed. An intermediate value, usually two or three times r_a, is chosen.

The same arguments for and against a high load resistance apply with even more force for a pentode. Load lines for both high and low resistance in this case are shown in Fig. 20.8(*b*). Since r_a for a pentode valve is very high, a load resistance of the order of r_a would give marked distortion of the output. A suitable value is very often one for which the load line runs through the knee of the $V_{gk} = 0$ characteristic. This is usually about two or three times the value which would be used with a roughly equivalent triode.

The increased amplification obtainable with a high anode resistance is often, in any case, illusory, since this gives a large internal resistance to the amplifier and hence the output voltage falls considerably when the next stage is connected to the amplifier terminals. Also, with a high anode resistance, stray capacitances have a relatively greater shunting effect, giving reduced amplification, particularly at high frequencies.

EXAMPLE 20.2 For the triode, whose anode characteristics are given in Fig. 20.9(*a*), determine the valve parameters in the linear region. Design an

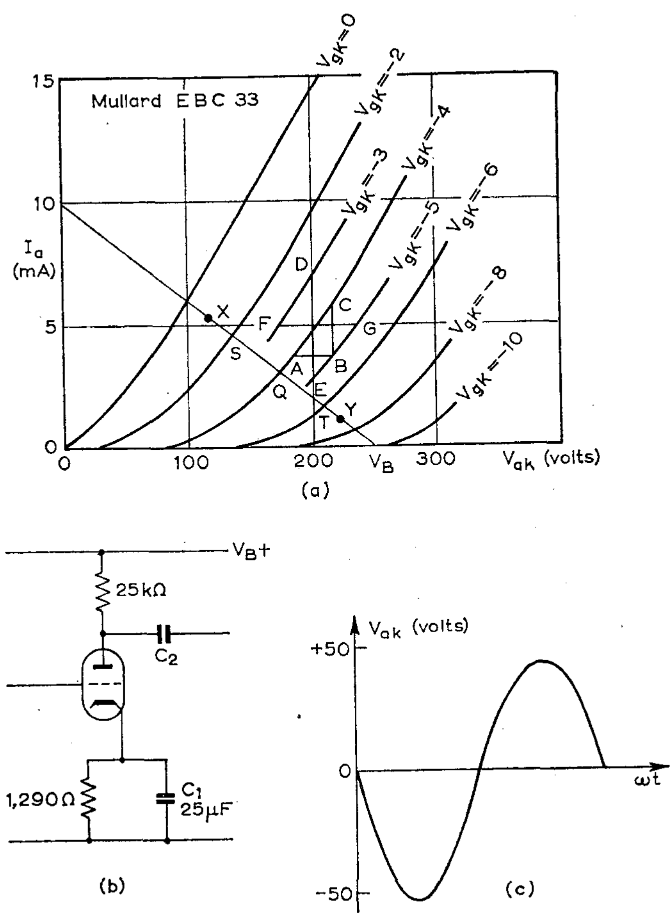


Fig. 20.9

amplifier using this valve and a supply of 250 V to give uniform amplification to audio frequencies above 50 Hz. Include a cathode bias circuit. Estimate the voltage amplification obtainable and draw the output voltage waveform corresponding to a sinusoidal input of 2 V (r.m.s.).

The parameters are taken for the region around the point $I_a = 5\,\text{mA}$, $V_{ak} = 200\,\text{V}$ in Fig. 20.9(*a*).

$$r_a = \frac{\text{AB}}{\text{CB}} = \frac{33}{2{\cdot}3} \times 1{,}000 = \underline{\underline{14{\cdot}4\,\text{k}\Omega}}$$

$$g_m = \frac{\text{DE}}{5-3} = \frac{4}{2} = \underline{\underline{2\,\text{mA/V}}}$$

$$\mu = \frac{\text{FG}}{5-3} = \frac{58}{2} = \underline{\underline{29}}$$

The -3 V and -5 V characteristics were interpolated.

A suitable anode resistance might be 25 kΩ for a 250 V supply, i.e. approximately $2r_a$ (up to about 50 kΩ may be used without severe distortion).

The load line for 25 kΩ at 250 V is shown, i.e. the line from

$$(V_{ak} = 250\,\text{V},\ I_a = 0) \quad \text{to} \quad \left(V_{ak} = 0,\ I_a = \frac{250}{25{,}000} = 10\,\text{mA}\right)$$

For an input of 2 V r.m.s., i.e. 2·83 V peak, the grid-bias voltage should be at least -3 V. Suppose -4 V is chosen for convenience; the static operating point is then Q in Fig. 20.9. Thus

Mean anode current = 3·1 mA

And by eqn. (20.3),

$$\text{Cathode bias resistance, } R_1 = \frac{4}{3{\cdot}1} \times 1{,}000 = 1{,}290\,\Omega$$

When the internal impedance of the valve is taken into account it can be shown that the effective cathode-to-earth resistance is 1,290 Ω in parallel with $1/g_m$. The capacitor C_1 is chosen so that its reactance at the lowest frequency to be amplified is very much less than the effective cathode-to-earth resistance; $C_1 = 25\,\mu\text{F}$ (say). The circuit is shown in Fig. 20.9(*b*).

To estimate the voltage amplification assume that the input voltage is 2 V peak so that the grid voltage varies from -2 V to -6 V, i.e. a variation of 4 V from S to T in Fig. 20.9(*a*).

At S, anode voltage = 135 V
At T, anode voltage = 207 V

Therefore

$$\text{Voltage amplification} = -\frac{207 - 135}{4} = \underline{\underline{-18}}$$

It will be appreciated that the voltage amplification varies somewhat with the amplitude of the input voltage.

For a sinusoidal input voltage of 2 V r.m.s., the anode voltage wave may be derived from the load line in Fig. 20.9(*a*).

ωt	0	45°	90°	135°	180°	225°	270°	315°	360°
Input voltage	0	2	2·83	2	0	−2	−2·83	−2	0
Operating point	*Q*	*S*	*X*	*S*	*Q*	*T*	*Y*	*T*	*Q*
Anode voltage	173	135	120	135	173	207	218	207	173
Variation of V_{ak}	0	−38	−53	−38	0	34	45	34	0

The anode voltage variation, i.e. the output waveform, is shown in Fig. 20.9(*c*). The waveform will, of course, be nearer to a true sinusoid if the amplitude of the input voltage is reduced.

20.8 Estimating Voltage Gain from Valve Parameters

For a valve amplifier the anode current, I_a, is some function of the grid–cathode and anode–cathode voltages, i.e.

$$I_a = f(V_{gk}, V_{ak})$$

Hence any change in anode current, δI_a, due to changes δV_{gk} and δV_{ak} in the grid–cathode and anode–cathode voltages can be expressed as

$$\delta I_a = \delta V_{gk}\left(\frac{\delta I_a}{\delta V_{gk}}\right)_{V_{ak}\,\text{const.}} + \delta V_{ak}\left(\frac{\delta I_a}{\delta V_{ak}}\right)_{V_{gk}\,\text{const.}}$$

or

$$\delta I_a = g_m \delta V_{gk} + \frac{1}{r_a}\delta V_{ak} \qquad (20.8)$$

Also

$$\delta V_{ak} = -R_L \delta I_a \qquad (20.9)$$

where R_L is the anode circuit resistance, and the minus sign is used to indicate that if I_a increases V_{ak} decreases.

From eqns. (20.8) and (20.9),

$$-\frac{\delta V_{ak}}{R_L} = g_m \delta V_{gk} + \frac{1}{r_a}\delta V_{ak}$$

or

$$\delta V_{ak} = \frac{-g_m r_a R_L \delta V_{gk}}{r_a + R_L} = \frac{-\mu R_L \delta V_{gk}}{r_a + R_L} \qquad (20.10)$$

Let V_{gk} represent the r.m.s. alternating grid–cathode voltage, and V_{ak} the r.m.s. alternating anode-cathode voltage; then, in terms of complexors, eqn. (20.10) can be rewritten as

$$V_{ak} = \frac{-\mu R_L V_{gk}}{r_a + R_L} \tag{20.11}$$

$$= \frac{\mu R_L V_{gk} \underline{/180^\circ}}{r_a + R_L} \quad (\text{since } -1 \equiv 180^\circ) \tag{20.11a}$$

The complex stage gain, A_v, is thus

$$A_v = \frac{V_{ak}}{V_{gk}} = \frac{-\mu R_L}{r_a + R_L} = \frac{\mu R_L \underline{/180^\circ}}{r_a + R_L} \tag{20.12}$$

Eqns. (20.11) and (20.12) thus represent the operation of the amplifier under small-signal linear conditions. They refer only to the alternating components of voltage.

20.9 Coupling of Multistage Valve Amplifiers

In order to increase the overall gain of an amplifier several stages may be connected in cascade. A two-stage amplifier is shown in Fig. 20.10, in which resistance-capacitance coupling is used to isolate

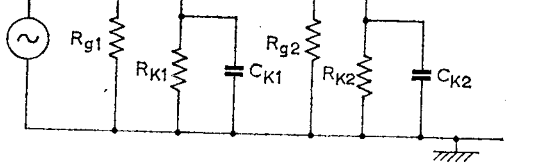


Fig. 20.10 CR-COUPLED TWO-STAGE AMPLIFIER

each stage from the d.c. levels of the previous stage. Thus capacitor C_1 acts as a d.c. block for the input signal source, and C_3 prevents the direct anode voltage of valve V_1 appearing at the grid of valve V_2. The values of C_1 and C_2 must be chosen so that at the lowest frequency of the input signal they are virtually short-circuits to alternating currents. Usually it is sufficient to make the reactance of the capacitors less than one-tenth of the resistance of the following grid resistors, at the lowest frequency that is to be amplified; i.e.

$$\frac{1}{\omega C_1} < 0{\cdot}1 R_{g1} \quad \text{and} \quad \frac{1}{\omega C_3} < 0{\cdot}1 R_{g2}$$

Coupling circuits will be discussed in greater detail in the next chapter.

Note that in a two-stage amplifier the effective resistance, R_P, in the anode circuit at signal frequencies is the parallel combination of the anode resistor, R_{L1}, and the following grid resistor, R_{g2}. In estimating the gain from the characteristics an a.c. load line must be used. This has a slope of $-1/R_P$ and passes through the d.c. operating point, Q.

20.10 The Semiconductor Junction Diode

It is assumed that the reader is aware that in an *intrinsic* semiconducting material (such as pure monocrystalline germanium, silicon, gallium arsenide, indium antimonide, etc.) conduction takes place by the movement of two types of *charge carriers, electrons* and *holes*. These are produced in equal numbers in the intrinsic material by thermal generation or by radiation. The number of *electron-hole pairs* increases rapidly as the temperature rises. Holes may be regarded as positive charge carriers, while electrons are, of course, negative charge carriers. In an *extrinsic* or *doped* semiconductor, impurities are added to the intrinsic material to give a predominance of either electrons (in *n*-type material) or holes (in *p*-type material) as charge carriers. The conductivity of a doped semiconductor is greater than and is less temperature sensitive than that of the intrinsic material up to fairly high temperatures. The charge carriers which predominate in a doped semiconductor are called the *majority carriers*.

Elemental semiconductors (e.g. silicon and germanium, in group IV of the periodic table of elements) have four electrons in the valence band, and form crystals of the diamond lattice structure where each atom shares its valence electrons with four adjacent atoms. This is called covalent bonding. If impurities in group V of the periodic table (e.g. phosphorus or arsenic) are added to the intrinsic semiconductor, these impurity atoms fit into the crystal lattice quite happily, and share four of their five valence electrons with four adjacent elemental semiconductor atoms, the fifth electron moving into the conduction band to contribute to the conductivity of the material. The material is then said to be *n*-type. Impurities which produce this effect are called *donor* impurities.

Intrinsic semiconductors are also formed when elements of group III (e.g. gallium) and group V (e.g. arsenic) form compounds (e.g. gallium arsenide) which have an average of four valence electrons per atom. Similarly some group II–group VI compounds form semiconductor crystals (e.g. cadmium sulphide), and silicon carbide is an example of two elements in Group IV forming a semiconductor crystal. Doping to give an *n*-type material can be achieved in a similar manner to that for elemental semiconductors. Thus if a group VI impurity (e.g. tellurium or sulphur) is added to a III–V compound, the impurity atom can replace one of the group V elements, its extra valence electron then being free to contribute to conduction. The same effect can be achieved by adding an excess of the group V element to the compound. Compound semiconductors are more difficult to purify and dope than elemental semiconductors.

In all *n*-type semiconductors, each added impurity atom gives one extra conduction electron. This increased number of conduction electrons results in a reduction in the mean density of thermally generated holes. This follows since if there are more conduction electrons there is a greater statistical probability of their recombining with holes, so that the mean lifetime of the holes falls. Note that since the donor atom loses one of its valence electrons it becomes a positive ion. These ions do not contribute to conduction since they are fixed in the crystal lattice. The net result of adding an *n*-type impurity is to increase the number of mobile electrons and to reduce the number of holes. Conduction in *n*-type semiconductors is thus mainly by electrons.

Elemental semiconductors become extrinsic *p*-type material when *acceptor* impurities (e.g. indium or boron in group III of the periodic table) are added to the intrinsic material. The impurity atoms fit into the crystal lattice, but since they have only three valence electrons instead of the four of the semiconductor atom that they replace, they each contribute a mobile hole (or space in the valence band that can be filled by an electron from an adjacent atom). Compound semiconductors can be similarly doped (e.g. zinc, in group II is often used as an acceptor to dope Group III–V compounds).

By analogy with the conditions in *n*-type semiconductors the increase in the number of holes in a *p*-type material due to the acceptor impurities gives rise to a reduction in the mean density of thermally generated electrons, and results in a material in which conduction is mainly by the movement of holes. These act as positive charge carriers.

In general, doped semiconductors are of higher conductivity than intrinsic semiconductors, the conductivity being a function of the number of impurity atoms added. The conductivity is also less temperature dependent than that of intrinsic material.

The conductivity of a material can be expressed in terms of the number of mobile charge carriers per unit volume, the size of the charge, and a constant, μ, called the *mobility* of the charge carriers. The mobility simply expresses the mean velocity attained by the charge carriers per unit of applied electric field. Mathematically the conductivity, σ, of a semiconductor may be written as

$$\sigma = en\mu_n + ep\mu_p$$

where e = electronic charge ($1{\cdot}6 \times 10^{-19}$C); n = electron density, i.e. the number of mobile electrons per cubic metre; μ_n = electron mobility (m/s per V/m); p = hole density; μ_p = hole mobility.

In intrinsic material $n = p = n_i$ (the intrinsic carrier density), and it may be shown that in doped semiconductors the product np is a constant, equal to n_i^2.

Generally hole mobility is smaller than electron mobility. Table 20.1 gives some idea of mobilities in different materials at room temperature (mobility varies with temperature—at around room temperature it falls as the temperature rises).

TABLE 20.1

Material	Electron mobility, μ_n	Hole mobility, μ_p	Intrinsic conductivity
	m/s per V/m	m/s per V/m	S/m
Germanium	0·38	0·18	2
Silicon	0·19	0·05	2×10^{-3}
Gallium arsenide	0·5	0·03	
Indium antimonide	10·0	0·3	
Silicon carbide	<0·01	<0·01	

Consider a junction formed between p-type and n-type materials as shown in Fig. 20.11(a). The alloy junction shown can be achieved by fusing an indium pellet into a small n-type wafer. This alloying process produces a very heavily doped p-type region to the left of the junction as the indium fuses into the n-type wafer. The n-type region may be lightly doped. Three conditions will be discussed.

1. *No external bias and open-circuit connexion.* At the junction shown schematically at (b), holes from the p-type material diffuse across into the n-type region, where there is an excess of conduction

electrons. The holes recombine with electrons in the n-type region so that a positive charge of captured (bound) holes is built up near the junction. Similarly, electrons diffusing into the p-type material from the n-type combine with the majority carriers (holes) there, to form a bound negative charge just on the p-side of the junction. The bound charges build up until the repulsion between them is just sufficient to prevent any further net diffusion of holes from p to n or electrons from n to p.

However, the bound charges attract across the junction the *minority carriers* (electrons in p and holes in n) which are generated thermally

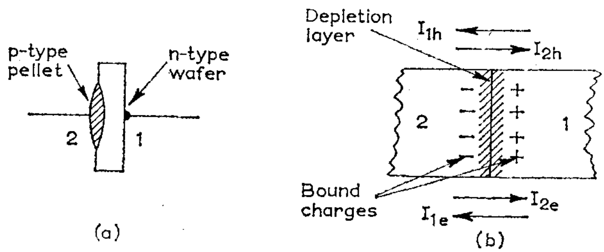


Fig. 20.11 ALLOY JUNCTION DIODE
(*a*) Construction of diode (*b*) The p–n junction

on the opposite side of the junction; i.e. minority holes in n are attracted across to the bound negative charge in p, and minority electrons in p are attracted to the bound positive charge in n. Let these currents be

I_{1h} = Minority hole current from n to p
I_{2e} = Minority electron current from p to n (see Fig. 20.11(*b*))

Since the p-region is assumed to be much more heavily doped than the n-region, there will be a *lower concentration* of thermally generated electrons in p than of holes in n; i.e.

$I_{1h} \gg I_{2e}$ (a ratio of 100:1 can easily be achieved)

A junction of this nature is referred to as a p^+–n junction, to indicate that the p-region is more heavily doped than the n-region.

In equilibrium these minority-carrier currents must be exactly balanced by the majority-carrier currents crossing the opposite way to maintain the bound-charge distribution; there is no net junction current.

In the region in which bound charges exist there are virtually no free carriers, and this is called the *depletion layer*. Mobile carriers are moved out of this region because of the electrostatic field due to the bound charges. Since the diffusing carriers must penetrate further into a lightly doped region before finding an opposite carrier with which to recombine, the depletion layer extends further into the lightly doped side of the junction than into the heavily doped side. Indeed if one side is much more heavily doped than the other, the depletion layer exists almost entirely on the lightly doped side of the junction.

If the junction is short-circuited by an external connexion, an alternative path for the return of charge carriers exists and there

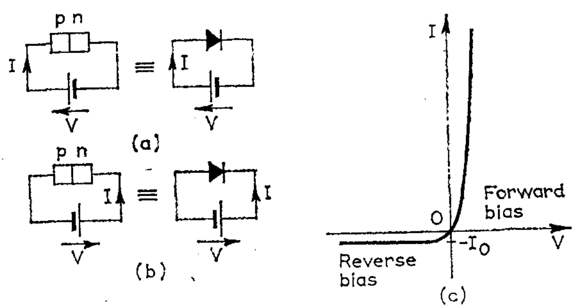


Fig. 20.12 EFFECT OF BIAS ON A *p–n* JUNCTION

(*a*) Forward bias
(*b*) Reverse bias
(*c*) Forward and reverse *I*/*V* characteristics

can be no net build-up of charge and hence the current will normally be negligible.

It is worth noting that the heavier the doping of the two regions the greater is the charge barrier which is produced to maintain equilibrium.

2. *Forward bias.* If an external d.c. supply is connected across a *p–n* junction as shown in Fig. 20.12(*a*), electrons in the *p*-type material will drift away from the junction towards the positive connexion to the supply, while holes in the *n*-type region will drift towards the negative connexion (where they recombine with electrons from the external circuit). This drift reduces the bound charge density at the junction, and allows majority carriers to flood both ways across the junction, giving rise to a large current. This is the foward bias connexion.

Since the *p*-region is assumed to be more heavily doped than the *n*-region, the junction current will consist mainly of holes crossing from *p* to *n*, i.e. majority carriers from the heavily doped region. The current in the connecting wires is, of course, an electron current. Above about 0·2 to 0·3V forward bias for germanium and about

0·6 to 0·7V for silicon, the current is limited only by the bulk resistance of the semiconductor material and the external circuit resistance.

3. *Reverse bias.* If the polarity of the supply is reversed (to give the reverse bias connexion of Fig. 20.12 (*b*)),the potential barrier at the junction is increased so that the majority carrier current falls, and eventually (above about 0·1V reverse bias) becomes negligibly small. However, the minority carriers continue to be attracted strongly across the junction. The number of these minority carriers is limited by the thermal generation rate of electron-hole pairs on each side of the junction to the equilibrium values in the unbiased junction i.e. to I_{1h} for holes crossing from n to p and I_{2e} for electrons crossing from p to n. Hence the reverse bias current becomes constant at a value

$$I_0 = I_{1h} + I_{2e}$$

It has been seen that $I_{1h} \gg I_{2e}$ for a p^+–n junction so that the reverse bias current depends mainly on the flow of minority carriers from the lightly doped region.

The complete characteristic is shown in Fig. 20.12(*c*). The reverse-bias saturation current is typically of the order of a few microamperes for germanium diodes, and at least two orders of magnitude less for similar silicon diodes at room temperature, depending upon fabrication techniques.

Since the reverse bias current depends on the minority carrier concentration in the lightly doped region it increases rapidly with temperature, approximately doubling for every 10°C rise in temperature in germanium and 6°C rise in silicon. The equation for junction current can be shown to be

$$I = I_0(e^{eV/kT} - 1) \tag{20.13}$$

where e is the magnitude of the charge of an electron or a hole ($1{\cdot}6 \times 10^{-19}$C), V is the bias applied to the depletion layer, k is Boltzmann's constant ($1{\cdot}38 \times 10^{-23}$J/K), and T is the absolute temperature of the junction in kelvins.

Similar considerations apply to a junction in which the n-region is more heavily doped than the p-region (a p–n^- junction). In this case the forward current across the junction is mainly electrons from the n-region (majority carriers from the heavily doped region), while the reverse bias current is mainly electrons from the p-region (minority carriers from the lightly doped region).

20.11 The Junction Transistor

There are two main types of transistor—*junction* also known as *bipolar* because carriers of both polarities, electrons and holes, take part in the action, and *field-effect*, also known as *unipolar* because carriers of only one polarity are involved. Field-effect transistors are considered in Chapter 26.

The construction of a *p–n–p* germanium alloy junction transistor is shown in Fig. 20.13 together with its graphical symbol. A small indium pellet is fused into one side of a thin, lightly doped *n*-type wafer of germanium (called the *base*) to give a heavily doped *p*-region, called the *emitter*. A p^+–*n* junction is thus formed. On the other side of the base wafer a slightly larger indium pellet is fused, to give a second *p–n* junction. This second *p*-region is called the *collector*.

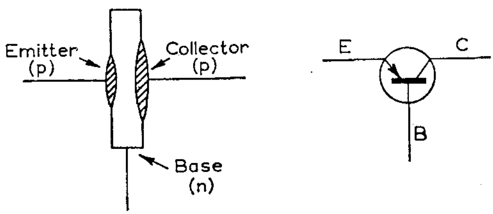


Fig. 20.13 ALLOY JUNCTION *p–n–p* TRANSISTOR

The *p–n–p* alloy junction germanium transistor was one of the earliest types to be produced. It is still widely used for low-frequency applications. Owing to difficulties with the materials used, *n–p–n* transistors (in which the base layer is of *p*-type material and the collector and emitter regions are of *n*-type) are not made by the alloy process; nor are satisfactory devices made of silicon manufactured in this way.

Since silicon has considerable advantages electrically compared to germanium when used to make transistors, considerable effort has been put into the technology of silicon device manufacture, despite the fact that silicon is a more difficult material to handle than germanium. It has been found that *p*- and *n*-regions can be introduced into semiconductor crystals by a diffusion process, and that both silicon and germanium devices can be made using this process. With diffusion methods it is relatively easy to obtain silicon *n–p–n* transistors with higher gain, lower leakage current and much better high-frequency response than is possible with germanium *p–n–p* alloy junction types. In addition it is possible to use diffusion in order to produce complete circuits on a single silicon crystal chip. Such integrated circuits are becoming widely used in present-day technology.

The construction of a planar *n–p–n* transistor is shown in Fig. 20.14(*a*), together with its graphical symbol. A lightly doped wafer of *n*-type silicon is used as the collector (the light doping enables an increased collector voltage to be used compared to the alloy junction type). A thin *p*-type layer is formed by exposing an etched area of the surface to hot gases containing acceptor impurities. The *p*-type impurity diffuses into the *n*-type wafer to form the base of the transistor. A small *n*-type layer is then diffused into a further etched area in the centre of the base layer to give the emitter; this

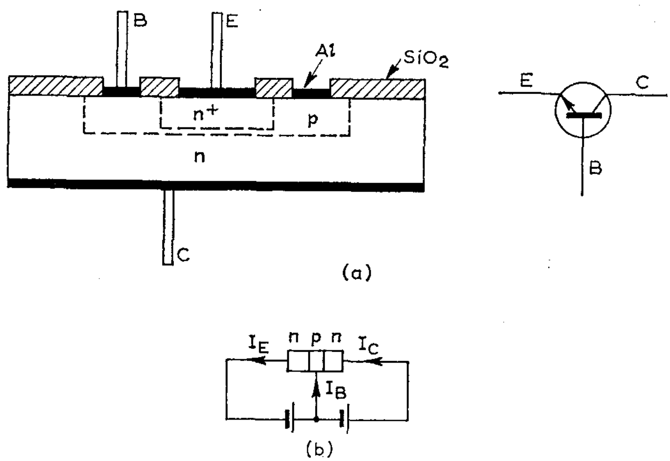


Fig. 20.14 *n–p–n* SILICON PLANAR TRANSISTOR
(*a*) Construction and graphical symbol
(*b*) Method of biasing

is a heavily doped region. Aluminium electrodes are evaporated on, and the terminations of the junctions are protected by a layer of silicon dioxide, which prevents the ingress of moisture and reduces surface leakage current.

Since the diffusion process can be accurately controlled (as distinct from the alloying process), very thin base widths (about 1 μm) are possible, and this increases the gain and improves the high-frequency response of the device.

In operation the emitter–base junction is forward biased, and the collector–base junction is reverse biased as shown for an *n–p–n* transistor in Fig. 20.14(*b*). Owing to the forward bias on the emitter–base junction, a large current, I_E, flows. Since the emitter is heavily doped, I_E consists mainly of electrons injected into the base from the emitter. Transistor action depends on these injected majority carriers, and this is why it is essential to have a heavily doped emitter and a lightly doped base (the holes injected from the base into the emitter play no part in transistor action). A ratio of electron current to hole current of more than 100:1 can readily be achieved. The ratio of electron current to hole current is called the *emitter injection ratio*.

The injected electrons diffuse across the base towards the reverse-biased collector–base junction. As soon as they reach the collector-junction depletion layer the electrons are pulled into the collector (or collected) so that the collector current increases by almost the same amount as the injected electron current from the emitter. (Remember that the current across a reverse-biased junction is mainly formed by minority carriers from the lightly doped side, and electrons are the minority carriers in the very lightly doped *p*-type base). Since the base is very thin and the collector area large, almost all the injected electrons are collected. However, some recombination of injected electrons with the majority holes in the base does take place. To maintain charge neutrality a current of electrons must flow out of the base. This current, I_B, will normally be only a very small fraction of the collector current I_C. Obviously,

$$I_E = I_C + I_B \tag{20.14}$$

Note that the transistor (unlike a valve which requires a heated filament in order to produce an electron stream) does not require a heater supply. This is one of its main advantages over valves.

The *p–n–p* transistor operates in a similar manner, but the injection is of holes at the emitter junction, and it is the hole current which is collected by the collector. The polarities for forward and reverse bias are reversed.

20.12 Static Characteristics—Common Base Connexion

The connexion of the *n–p–n* transistor illustrated in Fig. 20.15(*a*) in which the input is applied between emitter and base and the output is taken between collector and base is called *common base* (C.B.) connexion. The input characteristic relating emitter current and base–emitter voltage shown at (*b*) exhibits the curvature of a forward-biased junction.

The static collector, or output, characteristics are shown at (*c*). In the first quadrant they have been drawn as horizontal lines, but they should really have positive slopes of less than 1 μA/V, and should therefore have counterparts in a family of current-transfer curves, each for a particular value of V_{CB}. However, corresponding to the horizontal lines is the single current-transfer characteristic at (*d*). When the extremely small slope of the output characteristics is taken into account, the equation for the collector bias current at a particular quiescent point is

$$I_C = I_{CBO} + \bar{\alpha}_B I_E \qquad (V_{CB} \text{ constant}) \qquad (20.15)$$

I_{CBO} is the collector leakage current, i.e. the current through the reverse-biased collector–base junction when $I_E = 0$, and $\bar{\alpha}_B$ is the mean static current amplification factor, or the inherent static

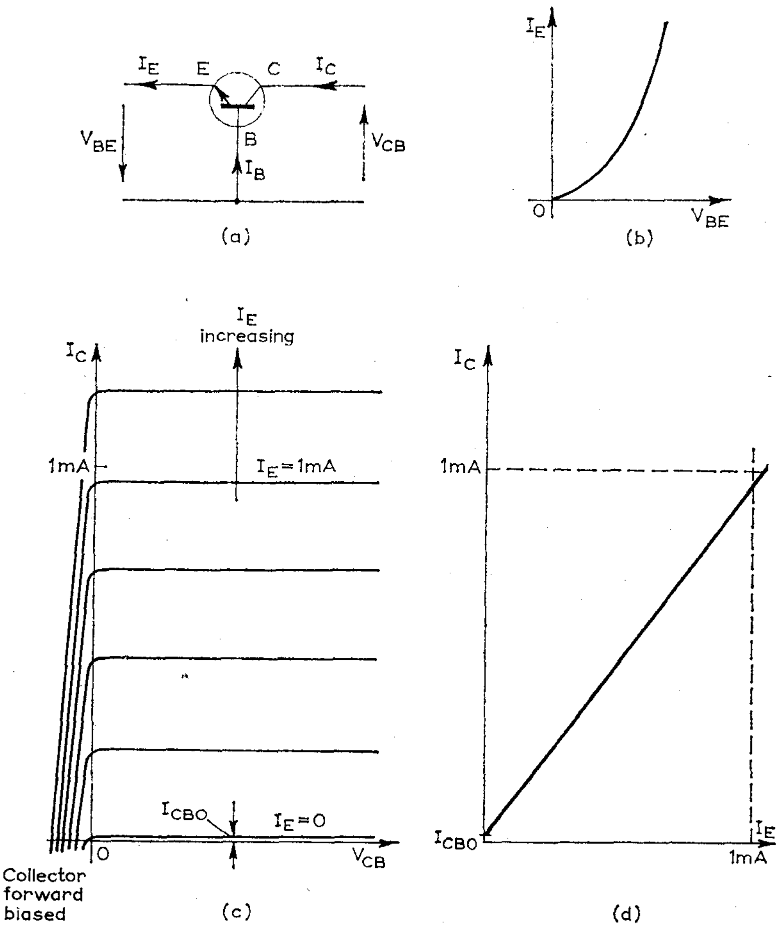


Fig. 20.15 COMMON BASE CONNEXION

(*a*) Circuit
(*b*) Input characteristic
(*c*) Output characteristics

forward current-transfer ratio, for the C.B. connexion. I_{CBO} is small, especially with silicon, and $\bar{\alpha}_B$ may lie between 0·95 and almost unity, so that I_C is very nearly equal to I_E. The ratio I_C/I_E is given the symbol $-h_{FB}$, which, with silicon, means much the same as $\bar{\alpha}_B$.

EXAMPLE 20.3 A silicon transistor has a collector leakage current, I_{CBO}, of 0·02μA at 300K. The leakage current doubles for every 6K rise in temperature. Calculate the base current at (*a*) 300K, (*b*) 330K when the emitter current is 1 mA, given that $\bar{\alpha}_B = 0{\cdot}99$.

(*a*) At 300K, eqn. (20.15) gives

$$I_C = 0{\cdot}00002 + 0{\cdot}99 = 0{\cdot}99\,\text{mA}$$

Hence, since $I_E = I_C + I_B$,

$$I_B = 1{\cdot}0 - 0{\cdot}99 = 0{\cdot}01\,\text{mA} = \underline{\underline{10\mu\text{A}}} \text{ directed towards the base}$$

(*b*) At 330K, I_{CBO} will be 0·64. Hence

$$I_C = 0{\cdot}00064 + 0{\cdot}99 \approx 0{\cdot}9906\,\text{mA} \qquad \text{so that} \qquad I_B = \underline{\underline{9{\cdot}4\mu\text{A}}}$$

If the temperature were to increase sufficiently the base current would reverse in direction.

At any given temperature, germanium transistors have a much larger leakage current than silicon transistors. Hence they are much more affected by temperature variations, despite the fact that the per-unit change in I_{CBO} with temperature is smaller in germanium than in silicon.

Note that when $V_{CB} = 0$ the collector still collects almost all the injected emitter current, since its barrier charge pulls the minority carriers across. This is particularly true of transistors with thin bases. It is only when the collector junction is forward biased that the collector current becomes zero, due to the forward collector–base current cancelling the reverse current. The larger the emitter current the greater must be the collector forward current and hence the greater must be the forward bias voltage, in order to give $I_C = 0$. This explains why the knee of the characteristics is usually in the forward-biased collector-voltage region.

20.13 Static Characteristics—Common Emitter Connexion

The common emitter (C.E.) connexion of a transistor is illustrated in Fig. 20.16(*a*). This corresponds to a normal common-cathode valve configuration. The input is applied between base and emitter, and the output is taken between collector and emitter. The static collector characteristics give the relation between collector current, I_C, and collector–emitter voltage, V_{CE}, for various values of base current, I_B, as shown at (*b*).

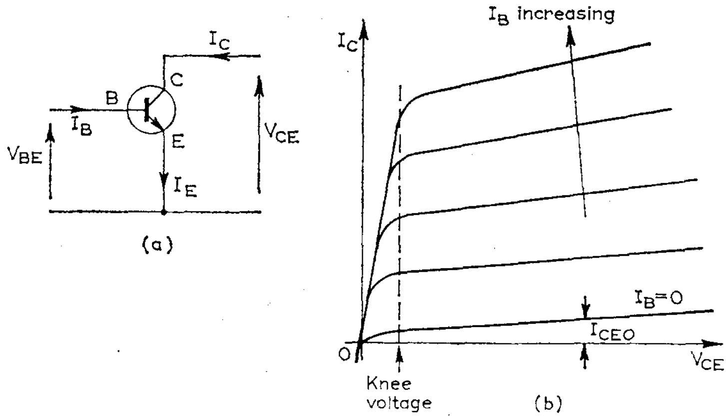


Fig. 20.16 COMMON EMITTER CONNEXION

(*a*) Circuit (*b*) Output characteristics

The collector current, I_C, is obtained from eqns. (20.14) and (20.15):

$$I_E = I_C + I_B \quad \text{and} \quad I_C = I_{CBO} + \bar{\alpha}_B I_E$$

whence

$$I_C = I_{CBO} + \bar{\alpha}_B(I_C + I_B)$$

or

$$I_C = \frac{I_{CBO}}{1 - \bar{\alpha}_B} + \frac{\bar{\alpha}_B I_B}{1 - \bar{\alpha}_B} \tag{20.16}$$

This equation is of the same form as eqn. (20.15), so that, for the C.E. connexion we may write

$$I_C = I_{CEO} + \bar{\alpha}_E I_B \qquad (V_{CE}\text{ constant}) \tag{20.17}$$

I_{CEO} $(=I_{CBO}/(1 - \bar{\alpha}_B))$ is the collector leakage current, i.e. the current through the collector–base junction when $I_B = 0$, and $\bar{\alpha}_E$ $(=\bar{\alpha}_B/(1 - \bar{\alpha}_B))$ is the static current amplification factor, for the C.E. connexion.

Since $\bar{\alpha}_B$ is nearly unity $(1 - \bar{\alpha}_B)$ is small, so that I_{CEO} is very much larger than I_{CBO}. Thus any temperature effect which increases I_{CBO} will cause a much greater increase in I_{CEO}, so that if germanium transistors are used it is particularly important to stabilize against temperature changes. Again because $(1 - \bar{\alpha}_B)$ is small, $\bar{\alpha}_E$ can have values ranging from about 20 to 100 or more. The ratio I_C/I_B is given the symbol h_{FE}, which, with silicon, means much the same as $\bar{\alpha}_E$.

The C.E. static characteristics shown at (*b*) have a greater slope than those for the C.B. connexion, mainly because $\bar{\alpha}_B$ increases slightly with increased collector–base voltage and the small increase in I_{CBO} with collector voltage is magnified $1/(1 - \bar{\alpha}_B)$ times—indeed the slope in the C.E. connexion will be greater by about this factor than that in the C.B. connexion. The knee voltage of a C.E. characteristic occurs at a voltage which is much lower than that of a pentode and in small power transistors is less than a volt. The knee occurs at a positive value of collector–emitter voltage (for *n–p–n* transistors) since the voltage drop of the forward biased emitter junction must be added to V_{CB} in order to obtain the corresponding value of V_{CE}.

In the C.E. connexion the input current/voltage relation is a curve which resembles that relating emitter current to base–emitter voltage in the C.B. connexion. Eqn. (20.16) shows that the change in collector current is a linear function of base current, and it therefore follows that the change in collector current is *not* a linear function of base–emitter voltage. This is why transistors are normally operated with a current drive, as distinct from valves, which have a voltage drive. It should be noted that in fact $\bar{\alpha}_B$ and $\bar{\alpha}_E$ both vary slightly with collector current so that the linear relations expressed by eqns. (20.16) and (20.15) are themselves an approximation.

20.14 The C.E. Amplifier

In order to utilize the transistor as an amplifier in the C.E. connexion, the base–emitter junction is forward biased, and a load resistor is connected in the collector lead. The signal source is connected in series with the base connexion, and the resultant variation of base current causes a change in collector current. This in turn gives an alternating component of voltage across the load resistor. A simple circuit is shown in Fig. 20.17(*a*), using an *n–p–n* transistor.

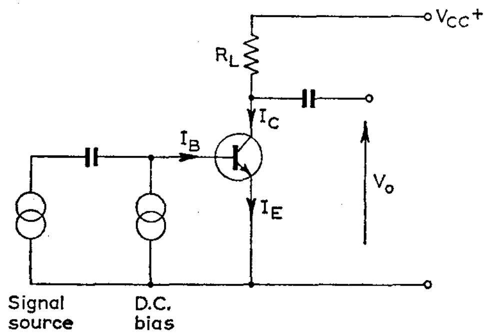
A constant-current d.c. bias source is used to give a suitable quiescent collector current. When the constant-current a.c. source is connected in parallel with the bias source, the base current will vary about its quiescent value to give a corresponding amplified change in collector current.

The operation of the circuit may be described graphically by drawing a load line on the static collector characteristics in a manner similar to that for the valve amplifier. Thus the operating point of the transistor must satisfy two requirements, namely

1. The collector current must be related to the base current by the static characteristics.
2. The collector current must satisfy the equation

$$V_{CE} = V_o = V_{CC} - I_C R_L$$

This is the equation of the load line shown in Fig. 20.17(*b*), which runs from ($I_C = V_{CC}/R_L$, $V_{CE} = 0$) to ($I_C = 0$, $V_{CE} = V_{CC}$). Then if the bias is adjusted to give a base current of I_{BO}, the quiescent operating point will be Q. If now the signal causes the base current



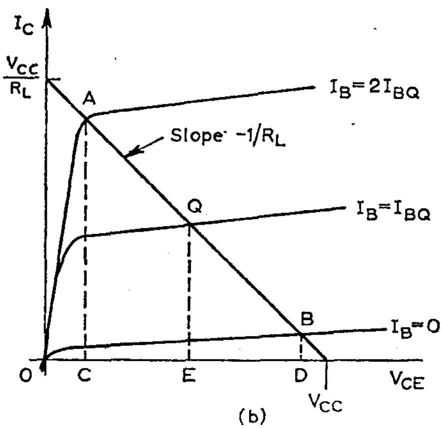


Fig. 20.17 SIMPLE C.E. AMPLIFIER

(*a*) Circuit
(*b*) Output characteristics with load line

to vary sinusoidally with a peak value of I_{BQ}, the operating point will move up and down the load line between A and B, giving an output peak-to-peak swing of CD volts. This simple circuit gives no compensation for temperature variations in I_{CEO}.

Note particularly that when I_B increases, I_C increases, so that the voltage drop across R_L increases and the output voltage (between collector and emitter) *decreases*, i.e. there is inherent phase reversal in the amplifier just as there is in the equivalent valve amplifier.

The steady bias current is chosen so that the quiescent operating point lies about the middle of the load line where the characteristics are most linear. Distortion of the signal is thus minimized.

20.15 Automatic Bias and Stability

The base-bias circuits employed in C.E. amplifiers should not only provide a suitable quiescent base current, but should also provide against changes in temperature, transistor characteristics, and supply voltage from affecting the quiescent operating point. It has been seen that, in an uncompensated transistor any increase in leakage current, ΔI_{CBO}, causes an increase in I_{CEO} of $\Delta I_{CBO}/(1 - \bar{\alpha}_B)$. In germanium transistors a rise or fall in temperature may change the quiescent collector current so much that the transistor is forced to operate over a non-linear range of its characteristics, and distortion will result. The *stability factor*, S, may be defined by the relation

$$S = \frac{\text{Change in } I_C}{\text{Given change in } I_{CBO}} = \frac{\delta I_C}{\delta I_{CBO}} \tag{20.18}$$

For the unstabilized circuit $S = 1/(1 - \bar{\alpha}_B) = \bar{\alpha}_E + 1$ (since $\bar{\alpha}_E = \bar{\alpha}_B/(1 - \bar{\alpha}_B)$).

Two other stability factors may also be defined:

$$M = \frac{\delta I_C}{\delta \bar{\alpha}_B} \tag{20.19}$$

which relates the change in I_C to changes in the C.B. current amplification factor $\bar{\alpha}_B$, and

$$N = \frac{\delta I_C}{\delta V_{CC}} \tag{20.20}$$

which relates the response of I_C to changes in the supply voltage, V_{CC}. In the ideal case S, M and N are all zero. Generally, if N turns out to be unacceptably high, the answer is to provide a supply that is better stabilized, so that $\delta V_{CC} \to 0$.

20.16 Base Resistor Bias

Perhaps the simplest way to achieve the required base-bias current is that shown in Fig. 20.18. A high-value resistor, R_1, is connected

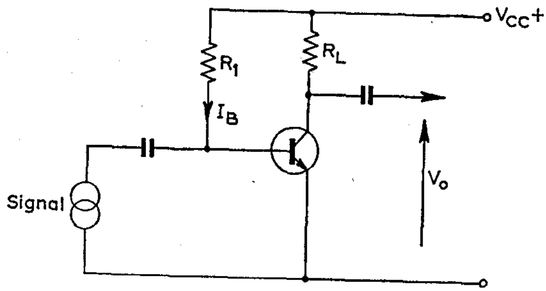


Fig. 20.18 SIMPLE BASE BIAS

from the supply direct to the base. Then if the base–emitter voltage drop is V_{BE}, it follows that the quiescent base current is given by

$$I_B = \frac{V_{CC} - V_{BE}}{R_1} \tag{20.21}$$

The quiescent collector current is

$$\begin{aligned} I_C &= I_{CBO}/(1 - \bar{\alpha}_B) + \bar{\alpha}_B I_B/(1 - \bar{\alpha}_B) = (1 + \bar{\alpha}_E)I_{CBO} + \alpha_E I_B \\ &= (1 + \bar{\alpha}_E)I_{CBO} + \bar{\alpha}_E(V_{CC} - V_{BE})/R_1 \end{aligned} \tag{20.22}$$

so that the stability factor is

$$S = \frac{\delta I_{CQ}}{\delta I_{CBO}} = 1 + \bar{\alpha}_E \tag{20.23}$$

The circuit is not stabilized at all against changes in I_{CBO} or $\bar{\alpha}_E$ caused by changes in temperature or tolerances in transistor parameters, since S has the same value as in the uncompensated circuit.

Also from eqn. (20.22),

$$N = \frac{\delta I_C}{\delta V_{CC}} = \frac{\bar{\alpha}_E}{R_1} \tag{20.24}$$

In a typical case with $\bar{\alpha}_E = 100$ and $R_1 = 100\,\mathrm{k}\Omega$, S would be 101 and N would be 1 mA/V.

20.17 Collector–Base Resistance Bias

A measure of stability can be achieved by connecting the base-bias resistor (R_1 in Fig. 20.18) to the collector instead of direct to the supply rail. This is shown in Fig. 20.19(a).

Stabilization against changes in I_{CBO} is achieved since the quiescent base current now depends on the collector potential. If I_{CBO} increases, I_C will also increase; the collector potential will therefore decrease and the base current will decrease, tending to reduce the collector current to its original value.

The emitter resistor R_3 shown at (*b*) further increases stability, since any increase in quiescent collector current, I_C, will now reduce the base–emitter voltage and again cause a reduction in I_B counteracting the increase in I_C. To prevent this effect from also reducing signal-current changes, R_3 is short-circuited for alternating currents by the capacitor C_3. As far as d.c. conditions are concerned, inserting R_3 is equivalent to increasing the collector load resistance, R_L.

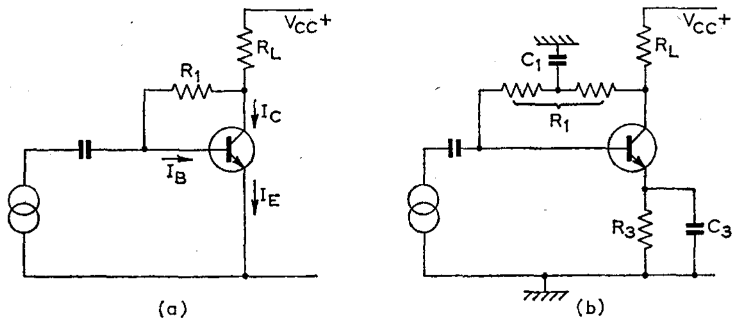


Fig. 20.19 COLLECTOR–BASE BIAS

The stability factor, S, is obtained from the following relations.

$$I_E = I_C + I_B \quad \text{(hence the current in } R_L \text{ is } I_E)$$

$$I_C = (1 + \bar{\alpha}_E)I_{CBO} + \bar{\alpha}_E I_B$$

and

$$V_{CC} = I_E R_L + I_B R_1 + V_{BE} + I_E R_3$$

Hence, after some manipulation,

$$I_C = \frac{R_1 + R_L + R_3}{\dfrac{R_1}{(1 + \bar{\alpha}_E)} + R_L + R_3} I_{CBO} + \frac{\bar{\alpha}_E(V_{CC} - V_{BE})}{R_1 + (R_L + R_3)(1 + \bar{\alpha}_E)}$$

so that

$$S = \frac{\delta I_C}{\delta I_{CBO}} = \frac{R_1 + R_L + R_3}{R_1/(1 + \bar{\alpha}_E) + R_L + R_3} \qquad (20.25)$$

Since I_B must always be positive in this circuit, the minimum collector current is equal to I_{CEO}, and this occurs when $I_B = 0$.

Note that there will be signal feedback through R_1 which will reduce the overall gain unless R_1 is divided into two sections, the junction being connected through a capacitor to earth as shown at (*b*).

It follows from the equation for I_C that

$$N = \frac{\delta I_C}{\delta V_{CC}} = \frac{\bar{\alpha}_E}{R_1 + (R_L + R_3)(1 + \bar{\alpha}_E)} \tag{20.26}$$

This factor in turn is obviously very dependent on any changes in $\bar{\alpha}_E$ (and hence in $\bar{\alpha}_B$).

20.18 Base Voltage-divider Bias

The circuit shown in Fig. 20.20 is one of the most commonly used in C.E. amplifiers for base current bias and operating point stabilization. Essentially the resistors R_1 and R_2 fix the base potential. The capacitor C_3 short-circuits the emitter for a.c. signals to prevent reduction of signal gain.

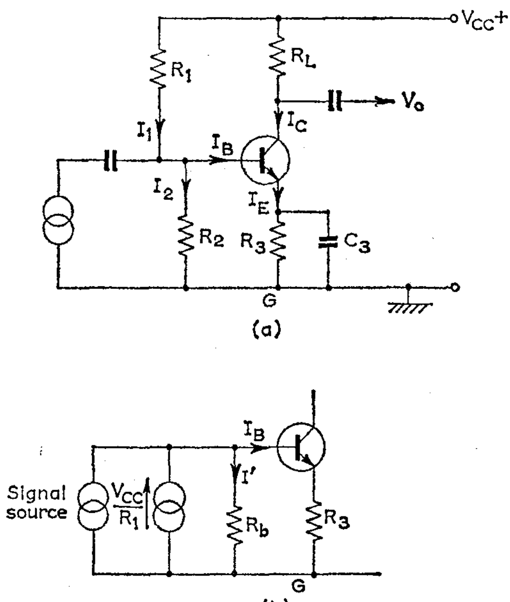


Fig. 20.20 VOLTAGE-DIVIDER BASE BIAS
(*a*) Complete circuit (*b*) A.C. circuit

The stabilizing action is achieved because the base potential is kept nearly constant, so that any increase in I_{CBO} (and hence increase in I_C and I_E) causes the emitter potential to rise, and hence the base–emitter voltage to fall. The base current is therefore reduced and the collector current falls towards its original value.

Expressions for the stability factors can be obtained by using Norton's theorem to replace the voltage divider R_1, R_2 by the constant-current source V_{CC}/R_1 in parallel with the resistance $R_b \doteq R_1R_2/(R_1 + R_2)$ as shown in Fig. 20.20(*b*). Then, for the d.c. conditions (zero signal),

$$I_C = (1 + \bar{\alpha}_E)I_{CBO} + \bar{\alpha}_E I_B \tag{20.27}$$

$$V_{BG} = V_{BE} + (I_B + I_C)R_3 = I'R_b \tag{20.28}$$

and

$$I' + I_B = \frac{V_{CC}}{R_1} = \frac{V_{BE}}{R_b} + (I_B + I_C)\frac{R_3}{R_b} + I_B$$

or

$$I_B = \frac{\dfrac{V_{CC}}{R_1} - \dfrac{V_{BE}}{R_b} - I_C\dfrac{R_3}{R_b}}{\dfrac{R_3}{R_b} + 1} \tag{20.29}$$

Substituting in eqn. (20.27) from eqn. (20.29),

$$I_C = (1 + \bar{\alpha}_E)I_{CBO} + \frac{\bar{\alpha}_E R_b}{R_3 + R_b}\left(\frac{V_{CC}}{R_1} - \frac{V_{BE}}{R_b} - I_C\frac{R_3}{R_b}\right)$$

so that

$$I_C = \frac{(1 + \bar{\alpha}_E)I_{CBO}}{1 + \dfrac{\bar{\alpha}_E R_3}{R_3 + R_b}} + \frac{\bar{\alpha}_E R_b}{(R_3 + R_b)\left(1 + \dfrac{\bar{\alpha}_E R_3}{R_3 + R_b}\right)}\left(\frac{V_{CC}}{R_1} - \frac{V_{BE}}{R_b}\right)$$

It follows that

$$S = \frac{\delta I_C}{\delta I_{CBO}} = \frac{1 + \bar{\alpha}_E}{1 + \dfrac{\bar{\alpha}_E R_3}{R_3 + R_b}}$$

$$= \frac{R_3 + R_b}{\dfrac{R_b}{(1 + \bar{\alpha}_E)} + R_3} = \frac{R_3 + R_b}{R_b(1 - \bar{\alpha}_B) + R_3} \tag{20.30}$$

Also

$$N = \frac{\delta I_C}{\delta V_{CC}} = \frac{R_b}{R_1} \frac{\bar{\alpha}_E}{R_b + R_3(1 + \bar{\alpha}_E)} \tag{20.31}$$

With this circuit the above equations indicate that it is possible to bias so that the quiescent collector current $I_C < I_{CEO}$, the base current then being in the reverse direction. Indeed if $R_b \to 0$ and $R_3 \to \infty$ the minimum collector current will tend to I_{CBO}. The quiescent base current I_B will then be negative. Note that the improvement in stability factor S compared with the unstabilized circuit is given by the factor $1/(1 + \bar{\alpha}_E R_3/(R_3 + R_b))$.

From eqn. (20.26) it will be seen that the stabilization is best when $R_b \ll R_3$. In practical circuits this cannot be achieved since if R_b is small a large part of the signal input current will be bypassed by it (so that current gain is lost), while if R_3 is large, a large value of supply voltage is required.

The stability factor S can be used in designing circuits for germanium transistors. Thus if the largest permissible change in I_C is known, and the greatest variation in I_{CBO} due to expected temperature changes is assumed, then S and hence the ratio R_3/R_b is determined.

EXAMPLE 20·4 For the circuit of Fig. 20.20 *(a)* $R_1 = 33$ kΩ, $R_2 = 5$ kΩ, $R_3 = 2{\cdot}2$ kΩ and $\bar{\alpha}_B = 0{\cdot}99$. Determine the stability factors S and N.

$$R_b = \frac{R_1 R_2}{R_1 + R_2} = 4{\cdot}5 \text{ k}\Omega$$

Hence from eqn. (20.30),

$$S = \frac{\delta I_C}{\delta I_{CBO}} = \frac{(4{\cdot}5 + 2{\cdot}2)10^3}{(4{\cdot}5 \times 0{\cdot}01 + 2{\cdot}2)10^3} = \underline{\underline{3}}$$

From eqn. (20.31),

$$N = \frac{\delta I_C}{\delta V_{CC}} = \frac{4{\cdot}5 \times 10^3}{33 \times 10^3} \times \frac{99}{(4{\cdot}5 + 2{\cdot}2 \times 100)10^3}$$
$$= 6 \times 10^{-5} \text{ A/V} = \underline{\underline{60\,\mu\text{A/V}}}$$

Although the foregoing analysis applies equally to germanium and silicon transistors, the value of I_{CBO} in silicon is so small that stabilization against any change of temperature on this account is virtually unnecessary unless the circuit is operating at unusually low values of collector current or at very high temperatures.

Another factor which affects the quiescent operating point in transistor circuits is the variation in $\bar{\alpha}_E$ from one transistor to another. This spread in values is generally greater than the production spread of valve parameters. It could give rise to large variations in the gain of transistor amplifiers.

Moreover, $\bar{\alpha}_E$, like I_{CBO}, has a positive temperature coefficient, and V_{BE} has a negative temperature coefficient. Generally, however, if S is small, the circuit will also be stabilized against changes in $\bar{\alpha}_E$ and V_{BE}.

Stabilization circuits are applications of the negative feedback principles discussed in Chapter 22.

20.19 Reverse Breakdown—the Zener Diode

The forward and reverse bias characteristics of a *p–n* junction were described in Section 20.10, where it was seen that when reverse bias is applied the current is almost constant at the very small reverse-bias

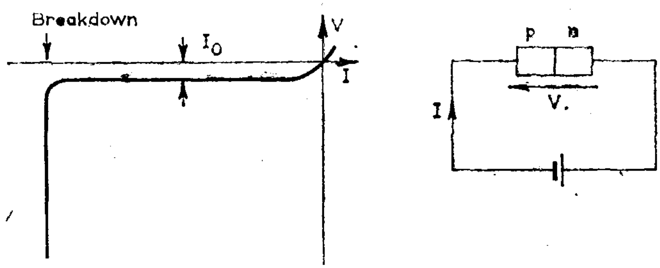


Fig. 20.21 REVERSE CHARACTERISTIC OF A *p–n* JUNCTION

saturation value. If the reverse bias voltage is increased, a point will be reached at which a large reverse current flows, and *breakdown* of the junction is said to have occurred.

The voltage across the junction then remains almost constant as shown in Fig. 20.21. Breakdown is due to two causes: *Zener effect* and *avalanche effect*, both of which arise in the depletion layer.

Avalanche Effect. The reverse bias voltage creates an electric field across the depletion layer. Carriers in the depletion layer are accelerated by the field, and if the electric stress is large enough, they may gain sufficient kinetic energy to produce electron-hole pairs by collision with some of the atoms in the crystal lattice. The electrons released by this ionization may then themselves be accelerated sufficiently to cause further electron-hole generating collisions, so that the process becomes cumulative. This is the avalanche effect. The holes produced by collision are less mobile than the electrons and are unlikely to add to the cumulative process. Avalanche effect gives rise to a very large increase in current, with little change in the reverse bias voltage. Hot spots may form at any imperfection at the junction, causing destructive breakdown.

Zener Effect. This effect occurs when the electric field in the depletion layer is large enough to produce electron-hole pairs by pulling electrons forcibly away from their parent atoms, i.e. by breaking the covalent bonds. This high field effect causes an increase in free charge carriers and hence a large increase in current for small increases in voltage, as in the avalanche effect.

Both effects depend on a high electric field, i.e. a large voltage gradient across the depletion layer. This in turn depends on (*a*) the applied voltage, and (*b*) the width of the depletion layer. Since the depletion layer is wider in a lightly doped material it follows that the *breakdown voltage* depends on the degree of doping of the lightly doped side of the junction, being higher for a more lightly doped material.

In *Zener diodes* the junction is specially made to be as uniform and free from imperfections as possible, so that breakdown will occur uniformly and will not be destructive. The voltage at which breakdown occurs can be chosen by a suitable choice of doping, and this gives a wide range of diodes suitable as voltage stabilizers. With a high level of doping the depletion layer will be narrow so that a small voltage will establish a high enough field in the depletion layer to cause breakdown. In a lightly doped junction, the depletion layer will be wide and the applied voltage will require to be large in order to give the electric field necessary for breakdown.

For breakdown voltages below about 6 V the main effect is Zener breakdown. Above this voltage the main effect is avalanche breakdown. The effects may be distinguished since the temperature coefficient of avalanche breakdown is positive because a higher temperature reduces the mean free path of the electrons and so requires a higher voltage for breakdown; whereas the temperature coefficient of Zener breakdown is negative, because increased thermal vibrations help the electric field to break the covalent bonds in the semiconductor. Diodes which break down at about 6 V show almost zero temperature coefficient—both avalanche and Zener effects are present, and their temperature coefficients almost cancel, being slightly negative for low currents and increasing through zero to become slightly positive at high currents due to the increased avalanche effect. Another distinguishing feature is that, since avalanching is a multiplicative process, the onset of breakdown is much more sudden than in the case of Zener breakdown.

20.20. Depletion Layer Capacitance

In Section 20.10 it was seen that when a *p–n* junction is formed in a semiconductor, holes from *p* diffuse across to *n* while electrons from *n* diffuse across to *p*. These diffusing minority carriers combine with the majority carriers on the other side of the junction to set up a bound charge distribution. The result of this recombination is that a layer exists on both sides of the junction which is depleted of mobile carriers. On the *p*-side of the metallurgical junction this *depletion layer* will have a net negative charge. On the *n*-side it will have an equal positive charge. Hence opposite charges exist across the junction separated by a depletion layer with few free carriers. This constitutes a capacitance called the *depletion layer capacitance*.

The extent of the junction depletion layers is dependent on the doping levels on either side. Thus for a p^+n junction holes must diffuse further into the lightly doped *n*-type region before they

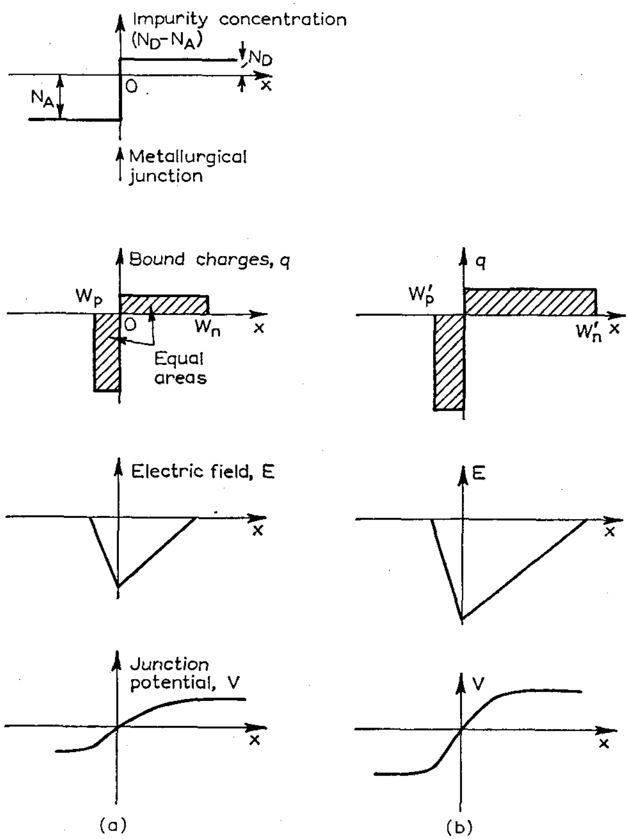


Fig. 20.22 JUNCTION CONDITIONS AT A *p–n* BOUNDARY
(*a*) Zero external bias. (*b*) Applied reverse bias

recombine to form bound charges than do the electrons that diffuse across to the p^+-side. In other words, the depletion layer is wider on the more lightly doped side of a p–n junction. Indeed if the difference in doping levels is large enough it can usually be assumed that the depletion layer exists entirely on the lightly doped side of the junction.

Idealized junction relations are shown in Fig. 20.22(a) for zero bias conditions. The top diagram shows the impurity doping density at an abrupt junction, (N_D donors per cubic metre in n, and N_A acceptors per cubic metre in p). For a p^+n junction $N_A \gg N_D$. The next diagram shows the bound charge density. The relative

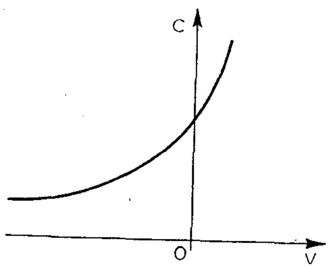


Fig. 20.23 VARIATION OF DEPLETION-LAYER CAPACITANCE WITH VOLTAGE

lengths of the depletion layers, W_p and W_n, may be estimated from the fact that the bound charges must be equal, i.e. $W_pN_A = W_nN_D$, and the shaded areas are thus equal. The charge distribution sets up an electric field across the junction, as shown in the third diagram. The electric field E, has a maximum value at the metallurgical boundary (it can be shown that E is proportional to $\int$ (bound charge density) dx). The junction potential, V, is found (since $V = -\int E dx$) by taking the negative of the integral of the curve of E, as shown in the fourth diagram. The total depletion layer width is $W_p + W_n$, which in a heavily doped p^+n junction becomes W_n. If a reverse bias is applied externally the corresponding conditions are shown at (b). The top diagram here shows the bound charge distribution. The reverse bias increases the junction potential, and this in turn increases the electric field and also the total bound charge. Since the height of the bound charge distribution is dependent on the impurity concentration, this will be constant. Hence the width of the depletion layer will increase and the depletion layer capacitance will decrease (the usual capacitor effect—if the dielectric is wider the capacitance is less).

It may be shown that the depletion layer width varies with the square root of the reverse bias voltage, and that the associated capacitance varies inversely as $\sqrt{V}$. Typical curves of capacitance variation are shown in Fig 20.23. This indicates the possibility of a voltage-sensitive capacitive element which finds applications in frequency-modulation circuits and parametric amplifiers.

PROBLEMS

20.1 A region of the anode characteristics of the Mullard EC52 triode is given in the following table. Estimate the values of r_a, g_m and μ for the given region.

$V_{GK} = -2\text{V}$	I_A (mA)	5	10	15
	V_A (V)	150	205	245
$V_{GK} = -3\text{V}$	I_A (mA)	4	10	15
	V_A (V)	200	260	305
$V_{GK} = -4\text{V}$	I_A (mA)	4	9	14
	V_A (V)	260	310	360

Ans. 9,500Ω, 6·0mA/V, 57·5.

20.2 One stage of a low-frequency amplifier employs a tetrode having static characteristics as given in the accompanying table. The anode load of the valve is a resistance of 17kΩ. Determine the maximum amplification of the stage.

V_A volts	50	100	150	200
V_{GK} volts	I_A milliamperes			
−2·25	11·6	12·4	13·0	
−4·5		7·0	7·5	7·75
−6·75		2·9	3·2	3·4

Ans. 33·5.

20.3 The accompanying table gives anode-current/anode-voltage characteristics for a beam tetrode for two different values of screen voltage. Determine:

(*a*) the constants of the equivalent generator for the valve operated normally as a tetrode, and calculate the voltage amplification obtained with a 5,000Ω pure-resistance anode load; and
(*b*) the corresponding data for the valve operated as a triode with anode and screen directly connected. The current drawn by the screen under tetrode operation may be neglected.

Assume the static operating point to be at $V_A = 200$V and $V_{GK} = -5$V.

Screen Voltage 250V

$V_{GK} = 0$V		$V_{GK} = -5$V		$V_{GK} = -10$V	
V_A	I_A	V_A	I_A	V_A	I_A
100	104mA	100	76mA	100	55mA
200	110	200	82	200	59
300	112	300	85	300	61

Screen Voltage 300V

$V_{GK} = 0$V		$V_{GK} = -5$V		$V_{GK} = -10$V	
V_A	I_A	V_A	I_A	V_A	I_A
100	125mA	100	100mA	100	76mA
200	135	200	108	200	84
300	141	300	114	300	88

(*L.U. part question*)

Ans. (*a*) 5·1 mA/V, 33kΩ, 22; (*b*) 5·1 mA/V, 1,600Ω, 6·2.

20.4 Explain the action of the three-electrode valve when used as an amplifier. In a particular case, with a load resistance of 8,000Ω the voltage amplification was 5·5 and with 12,000Ω it was 6·5. What amplification at a frequency of 800 Hz would be expected, using a choke coil of 10H inductance? (*L.U.*)
Ans. 10.

20.5 A single-stage triode voltage amplifier has a gain of 20. The gain is reduced to 15 by halving the load resistance. Find the amplification factor of the triode. If the operating voltages are 300V and −3V, calculate the change in static current caused by doubling the bias assuming that the 3/2 power law is adhered to. (*H.N.C.*)

Ans. 30; 56·7 per cent reduction.

20.6 The static collector characteristics of an *n–p–n* transistor are linear over the range indicated.

$I_B = 0$	V_{CE} (volts)	1	10
	I_C (mA)	0·04	0·08
$I_B = 20\mu$A	V_{CE} (volts)	1	10
	I_C (mA)	0·93	0·98
$I_B = 40\mu$A	V_{CE} (volts)	1	10
	I_C (mA)	1·9	1·96

The transistor is used as a common-emitter amplifier from a 10V d.c. supply. The load in the collector circuit is 5kΩ, and the quiescent base current is 20μA. Determine the current gain (ratio of change in I_C to change in I_B) when the signal gives a base current variation of ±20μA. What are the maximum and minimum values of collector voltage?

Ans. 46; 9·6V; 0·5V.

20.7 In the CE amplifier circuit shown in Fig. 20.20(*a*) a *p–n–p* germanium transistor is used, the supply voltage is −6V, and $R_1 = 25$kΩ, $R_2 = 5$kΩ, $R_3 = 0{\cdot}8$kΩ, $R_L = 2{\cdot}5$kΩ, and $I_{CBO} = 1\mu$A. Determine the quiescent collector–earth voltage if $\bar{\alpha}_B = 0{\cdot}98$ and the base–emitter voltage is 0·3V. What is the value of the stability factor S?

Ans. −3·75V; 5·7.

20.8 A silicon n–p–n transistor is used in the circuit of Fig. 20.19. If $\bar{\alpha}_B = 0{\cdot}99$, $I_{CBO} = 0{\cdot}1\,\mu\text{A}$, $R_L = 3\,\text{k}\Omega$, $R_1 = 330\,\text{k}\Omega$ and $R_3 = 0{\cdot}5\,\text{k}\Omega$, determine the quiescent collector voltage if the supply voltage is 10 V and $V_{BE} = 0{\cdot}7\,\text{V}$. Evaluate the stability factor S.

Ans. 6V; 49.

20.9 In the circuit of Fig. 20.24 show that

$$\frac{\delta I_C}{I_{CBO}} = \frac{R_1 + R_2}{R_1(1 - \bar{\alpha}_B) + R_2} \quad \text{and} \quad \frac{\delta I_C}{\delta V_{CC}} = \frac{\bar{\alpha}_E}{R_1 + (1 + \bar{\alpha}_E)R_2}$$

Neglect the base–emitter voltage drop. Would you expect this circuit to be more stable than that of Fig. 20.18? Explain your answer.

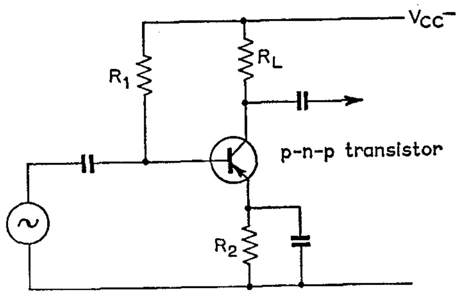


Fig. 20.24

20.10 An n–p–n silicon transistor at room temperature has its emitter disconnected. A voltage of 5 V is applied between collector and base, with the collector positive. A current of 0·2 A flows. When the base is disconnected and the same voltage is applied between collector and emitter the current is found to be 20 μA. Explain this effect and calculate the d.c. short-circuit current gain of the transistor, and the base and emitter currents when the collector current is 1 mA.

Ans. 0·99; 10μA; 11,010 μA.

20.11 In the transistor of Problem 20.10 the leakage current may be assumed to double for every 8°C rise in temperature above ambient. Determine the base and emitter currents when the collector current is 1 mA at a temperature 40°C above ambient. Assume that the d.c. short-circuit current gain remains constant.

Ans. 3·6 μA; 1,004 μA.

20.12 Find the values of R_1 and R_L for the simple base-resistor bias circuit of Fig. 20.18 assuming that a p–n–p germanium transistor is used and that $V_{CC} = -15\,\text{V}$, $V_{BE} = -0{\cdot}3\,\text{V}$, $I_{CBO} = 5\,\mu\text{A}$ and $\bar{\alpha}_E = 50$. The quiescent values of I_C and V_{CE} are to be 2.5 mA and −5 V respectively. Determine the values of the stability factors S and N.

Ans. 330 kΩ; 4kΩ; 51; 0·15 mA/V.

20.13 In the circuit of Fig. 20.18 an additional resistor R_E is connected between the emitter and earth. Determine the values of R_1, R_E and R_L assuming that the transistor used is the same as in Problem 20.12 and that the quiescent values of I_C, V_{CE} and V_E (the voltage across R_E) are 2·5 mA, −3 V and −3 V respectively. Also find the values of the stability factors S and N.

Ans. 260 kΩ; 1·2 kΩ; 3·6 kΩ; 41·5; 0·155 mA/V.

20.14 A silicon *n–p–n* transistor is used with the collector–base resistance bias circuit of Fig. 20.19. Design the circuit (i.e. find suitable values of R_1, R_L and R_3) for the following conditions: $V_{CC} = 20$ V; $\bar{\alpha}_E = 100$; $I_{CBO} = 0{\cdot}2\ \mu$A; $V_{BE} = 0{\cdot}7$ V; the quiescent values of I_C, V_{CE} and the voltage across R_3 are 1·5 mA, 8 V and 2 V respectively. Find the values of the stability factors S and N.

Ans. 480 kΩ; 6·66 kΩ; 1·33 kΩ; 38; 78 μA/V.

Chapter 21

SMALL-SIGNAL AMPLIFIERS

Small-signal amplifiers are assumed to operate over the linear range of the active device (valve or transistor) which they employ. They may be either voltage or current amplifiers, and will normally feed a further amplifier stage. The d.c. bias circuits are designed so that the active device operates over the linear parts of its characteristics. The quiescent voltages are found by considering the maximum peak-to-peak currents or voltages required for the signal (the quiescent current may have to be appreciably larger than the peak current swing to ensure linear operation). Under linear operating conditions the a.c. operation of the amplifier may be deduced by representing the active device by an equivalent circuit consisting of linear generators and circuit elements. In this chapter bipolar-transistor small-signal circuits will be considered. Field-effect transistor amplifiers will be dealt with in Chapter 26.

21.1 Transistor Equivalent Circuits

Unlike valves, which are voltage-controlled devices with almost infinite input impedance (the grid normally takes no current), bipolar junction transistors are generally current-controlled devices. The input impedance is not normally very high, and must be taken into account when circuit calculations are made.

Several transistor equivalent circuits can be drawn, and manufacturers quote many different parameters. Only two of these will be considered, namely (*a*) the equivalent-T circuit, and (*b*) the *h*-parameter equivalent circuit. It is assumed that the quiescent point has been chosen in the middle of the linear portion of the transistor characteristics, and that suitable d.c. bias circuits have been selected.

In a.c. (signal) equivalent circuits only those external components that are effective under a.c. conditions are shown. Thus the d.c. power supply is simply a short-circuit to alternating currents. Similarly, at mid-frequencies, decoupling and coupling capacitors will act as short-circuits. The effect of d.c. bias circuits on the signal is usually small so that they too can often be neglected in a.c. equivalent circuits.

EQUIVALENT-T CIRCUIT

The graphical symbol for a *p–n–p* transistor is shown in Fig. 21.1, together with the equivalent-T circuit, which applies to either type of transistor. The parameters of this circuit can be related to the

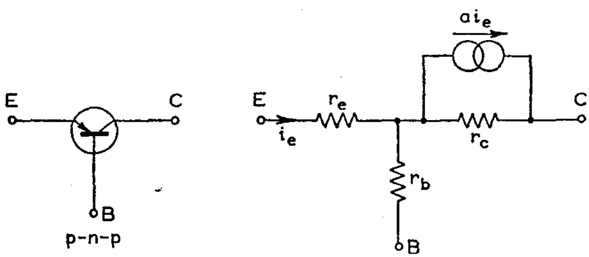


Fig. 21.1 TRANSISTOR EQUIVALENT-T CIRCUIT

physical operation of the transistor. Thus the resistance r_e represents the relatively low resistance of the forward-biased emitter–base junction, r_b represents approximately the resistance of the base layer (which will be appreciable since the base is very lightly doped, thin, and fed from its edge), and r_c represents the very high resistance of the reverse-biased base–collector junction. Physically, the emitter current is injected into the base and is transported across the base–collector junction by diffusion or drift, so that a proportion, α, actually crosses to the collector. This is represented by the constant-current generator, ai_e, across r_c, where $a \approx \alpha$ except at high frequencies.

The small-signal operation of a transistor amplifier stage may be determined by replacing the transistor by its equivalent circuit and solving the resulting circuit by conventional means. It will be noted that in common-base connexion, the resistance r_b is common to both input and output circuit loops. This means that the output load will affect input conditions—i.e. there is internal feedback in the transistor—and this causes considerable complications.

In practice the equivalent-T parameters are now seldom quoted on manufacturers' data sheets since it is rather difficult to measure the parameters r_e, r_b and r_c and the approximation made lead to inaccuracies if the common-base T-circuit is converted for common-emitter connexion. Typical values at a quiescent collector current of 1mA for a small germanium audio-frequency transistor are $r_e = 20\Omega$, $r_b = 700\Omega$, $r_c = 1\text{M}\Omega$ and $a = 0{\cdot}98$; and for a small silicon transistor, $r_e = 10\Omega$, $r_b = 1{,}000\Omega$, $r_e = 0{\cdot}5\text{M}\Omega$ and $a = 0{\cdot}99$.

THE *h*-PARAMETER EQUIVALENT CIRCUIT

In the small-signal range the transistor can be regarded as an active two-port "black box", i.e. it has two input terminals and two output terminals (Fig. 21.2(*a*)). There are four external variables, the

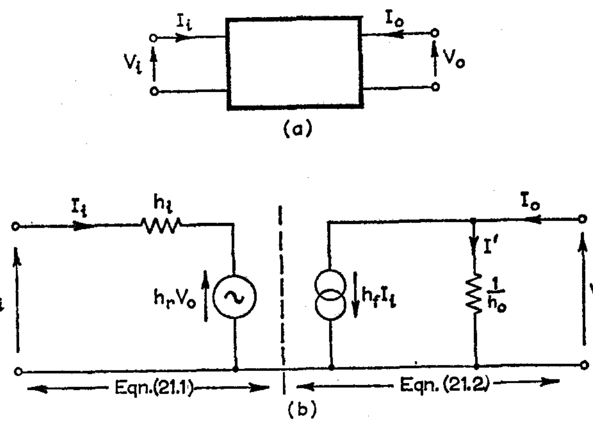


Fig. 21.2 *h*-PARAMETER EQUIVALENT CIRCUIT

input voltage and current (V_i and I_i) and the output voltage and current (V_o and I_o). Since operation takes place over the linear range of transistor characteristics, V_i, I_i, V_o and I_o can be related to each other by constant parameters. Thus V_i and I_o can be related to I_i and V_o by four constants, called the *h-parameters*, giving the following complexor equations:

$$V_i = h_iI_i + h_rV_o \tag{21.1}$$

$$I_o = h_fI_i + h_oV_o \tag{21.2}$$

The directions of the currents and the polarities of the voltage are conventionally chosen as shown in Fig. 21.2(*a*). There are several alternative ways of relating the input and output voltages and currents but these will not be discussed.

Eqns. (21.1) and (21.2) give rise to the equivalent circuit of Fig. 21.2(*b*). Thus the input voltage to the black box is made up of the voltage drop $h_i I_i$ across a resistor and the e.m.f. $h_r V_o$ of a constant-voltage generator (the left-hand half of Fig. 21.2(*b*)). It follows that h_i has the dimensions of a resistance and that h_r is a numerical constant. Also I_o is made up of the current $h_f I_i$ in a constant-current source and the current $V_o h_o$ through an admittance, h_o, which is connected across the output terminals (the right-hand half of Fig. 21.2(*b*)). The parameter h_f is a numerical constant.

The subscripts used for the *h*-parameters indicate their function in the equivalent circuit—h_i is the *input impedance* with short-circuited output (i.e. when $V_o = 0$), h_r is the *reverse voltage transfer constant*, h_o is the *output admittance* when the input is open-circuited (i.e. when $I_i = 0$), and h_f is the *forward short-circuit current-transfer constant* (i.e. $h_f = I_o/I_i$ when $V_o = 0$).

Since the transistor is a three-terminal device, it is possible to choose any one of its terminals as the one which is common to input and output circuits. This gives rise to the three practical connexions, common base, common emitter and common collector. The *h*-parameter equivalent circuit has the same form for each connexion, but the values of the parameters are different for all three. It is usual to indicate the connexion to which the parameters refer by a second subscript, *b* for common base, *e* for common emitter and *c* for common collector. Thus h_{fe} is the forward short-circuit current gain in the common-emitter connexion. Typical values of the small-signal *h*-parameters are indicated in Table 21.1, for a quiescent collector current of 1 mA.

The small-signal *h*-parameters can be evaluated by simple circuit tests. Suppose the output terminals are short-circuited ($V_o = 0$) and a voltage V_i is applied to the input. Then if a small change δV_i in V_i causes a small change δI_i in the input current I_i,

$$h_i = \left.\frac{\delta V_i}{\delta I_i}\right|_{V_o=0} \tag{21.3}$$

If the corresponding change in the output current is δI_o, then

$$h_f = \left.\frac{\delta I_o}{\delta I_i}\right|_{V_o=0} \tag{21.4}$$

h_f may be negative or positive, according to the connexion used (see Table 21.1).

TABLE 21.1

Parameter	OC71 Ge *p–n–p*	BCY30 Si *n–p–n*	2N930 Si *n–p–n*
h_{ib} (Ω)	35	30	20
h_{rb}	7×10^{-4}	$1{\cdot}5 \times 10^{-4}$	6×10^{-4}
h_{fb}	$-0{\cdot}98$	$-0{\cdot}98$	$-0{\cdot}995$
h_{ob} (μS)	1·0	0·5	0·5
h_{ie} (kΩ)	1·5	1·4	4·0
h_{re}	8×10^{-4}	6×10^{-4}	4×10^{-4}
h_{fe}	49	49	199
h_{oe} (μS)	50	25	100
h_{ic} (kΩ)	1·5	1·4	4·0
h_{rc}	1	1	1
h_{fc}	-50	-50	-200
h_{oc} (μS)	50	25	100

In order to determine h_r and h_o the input is open-circuited ($I_i = 0$) and a voltage V_o is applied at the output terminals, giving rise to a current I_o and a voltage across the input terminals of V_i. For a small change δV_o in V_o, let the corresponding changes in I_o and V_i be δI_o and δV_i. Then

$$h_r = \left.\frac{\delta V_i}{\delta V_o}\right|_{I_i=0} \tag{21.5}$$

and

$$h_o = \left.\frac{\delta V_o}{\delta I_o}\right|_{I_i=0} \tag{21.6}$$

21.2 Common-emitter Amplifier

In the common-emitter connexion, the input to the transistor is between base and emitter and the output is taken between collector and emitter. The graphical symbols for a *p–n–p* and an *n–p–n* transistor, and the common-emitter *h*-parameter equivalent circuit, are shown in Figs. 21.3(*a*) and (*b*). It is convenient in the analysis to assume that all signal currents are directed towards the transistor as shown, whether the transistor is *n-p-n* or *p-n-p*. Thus the following equation is always true

$$I_e + I_b + I_c = 0 \tag{21.7}$$

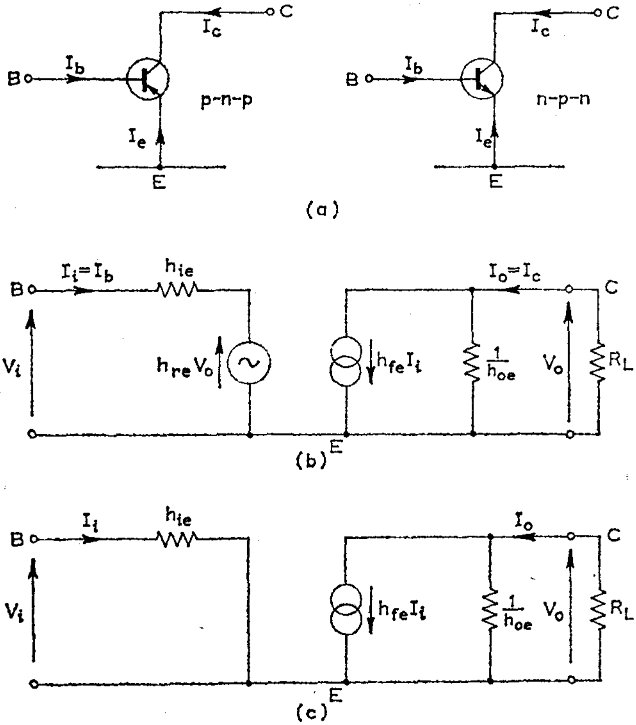


Fig. 21.3 COMMON-EMITTER *h*-PARAMETER EQUIVALENT CIRCUIT

Comparing Figs. 21.3(*a*) and (*b*) it will be seen that

$$I_b = I_i \qquad I_c = I_o \qquad V_{be} = V_i \qquad V_{ce} = V_o$$

Under normal operating conditions, and considering alternating currents and voltages only,* the value of h_{re} is small, so that there is small error in neglecting the generator $h_{re}V_o$ in the input circuit. This assumption considerably simplifies the equivalent circuit calculations.

From the simplified small-signal equivalent circuit of Fig. 21.3(*c*), it follows that

$$V_i = h_{ie}I_i \tag{21.8}$$

and

$$I_o = h_{fe}I_i + h_{oe}V_o \tag{21.9}$$

* Subscripts in small letters are used with the symbols for signal quantities, and subscripts in capital letters, with the symbols for quiescent d.c. quantities. Thus I_e = emitter signal current and I_E = emitter bias current.

Hence

$$\text{Input resistance, } R_i = \frac{V_i}{I_i} = h_{ie} \tag{21.10}$$

and

$$\text{Output resistance, } R_o = \frac{1}{h_{oe}} \tag{21.11}$$

The output impedance of an amplifier is the internal impedance as seen by any load placed upon it and may be defined as the ratio of output voltage to output current when an a.c. source is applied to the output terminals and the input is represented by the internal impedance of any input source.

$$\text{Current gain, } A_i = \frac{V_o/R_L}{I_i} = \frac{-I_o}{I_i}$$

From eqn. (21.9),

$$\begin{aligned} I_o &= h_{fe}I_i + h_{oe}V_o \\ &= h_{fe}I_i - h_{oe}R_LI_o \quad (\text{since } V_o = -I_oR_L) \end{aligned}$$

so that

$$I_o = \frac{h_{fe}I_i}{1 + h_{oe}R_L}$$

and

$$A_i = \frac{-h_{fe}}{1 + h_{oe}R_L} = \frac{h_{fe}\underline{/180^\circ}}{1 + h_{oe}R_L} \tag{21.12}$$

Since h_{fe} is positive, the minus sign indicates the 180° phase reversal between output and input.

$$\begin{aligned} \text{Voltage gain, } A_v &= \frac{V_o}{V_i} = \frac{-I_oR_L}{I_iR_i} \\ &= A_i\frac{R_L}{R_i} = \frac{-h_{fe}R_L}{R_i(1 + h_{oe}R_L)} \end{aligned} \tag{21.13}$$

Again the minus sign indicates the 180° phase reversal between output and input voltage.

$$\text{Power gain, } G = \frac{\text{Output power}}{\text{Input power}} = \frac{I_o^2R_L}{I_i^2R_i}$$

i.e.

$$G = A_i^2\frac{R_L}{R_i} = A_iA_v \tag{21.14}$$

Typical figures for the C.E. circuit are as follows:

	Germanium	*Silicon*
Input resistance	1·5 kΩ	2·5 kΩ
Output resistance	25 kΩ	70 kΩ
Current gain	−45	−80
Voltage gain	−80	−120

The actual values depend on the source and load resistances, as can be seen by carrying out an exact analysis of the circuit.

It is left as an exercise for the reader to deduce from the complete h-parameter equivalent circuit (i.e. including the feedback generator) that

$$\text{Input impedance} = h_{ie} - \frac{h_{re}h_{fe}R_L}{1 + h_{oe}R_L} \tag{21.15}$$

$$\text{Output admittance} = h_{oe} - \frac{h_{re}h_{fe}}{h_{ie} + R_g} \tag{21.16}$$

where R_g = resistance of signal source, and

$$\text{Current gain, } A_i = \frac{-I_o}{I_i} = \frac{-h_{fe}}{1 + h_{oe}R_L} \tag{21.17}$$

From these expressions it is seen that the input resistance is h_{ie} if R_L is small (this is the value of input resistance which the simplified circuit gives), and falls towards $h_{ie} - h_{re}h_{fe}/h_{oe}$ as $R_L \to \infty$ (an unusual operating condition). Also the output admittance rises towards h_{oe} as R_g rises.

EXAMPLE 21.1 For the common-emitter amplifier stage shown in Fig. 21.4(a), determine (a) the approximate d.c. conditions, (b) the small-signal mid-frequency current and power gains. The transistor is a BCY30; $R_1 = 60\,\text{k}\Omega$; $R_2 = 10\,\text{k}\Omega$; $R_3 = 1\,\text{k}\Omega$; $R_4 = 5\,\text{k}\Omega$; $R_L = 1{\cdot}4\,\text{k}\Omega$; $C_1 = C_2 = 100\,\mu\text{F}$; $C_3 = 250\,\mu\text{F}$. Neglect the leakage current.

The coupling capacitors, C_1 and C_2, are required to isolate the source and the load from the direct voltages on the transistor; and the emitter bypass capacitor, C_3, effectively connects the emitter to earth for the a.c. signal.

It is assumed that the reactances of the capacitors are negligibly small, and that the d.c. supply is also of negligible impedance. Hence the a.c. equivalent circuit is as shown in Fig. 21.4(b). The parallel combination of R_1 and R_2 shunts some of the input current, I_{in}, and so reduces the gain. Also the transistor output current, I_o, is divided between the load R_L and the collector circuit resistance R_4. On small a.c. signals the input impedance of the stage, R_{in}, is the parallel combination of R_1, R_2 and h_{ie}, i.e.

$$R_{in} = R_1 \| R_2 \| h_{ie} = 10^3/(1/60 + 1/10 + 1/1{\cdot}4 = 1{,}200\,\Omega$$

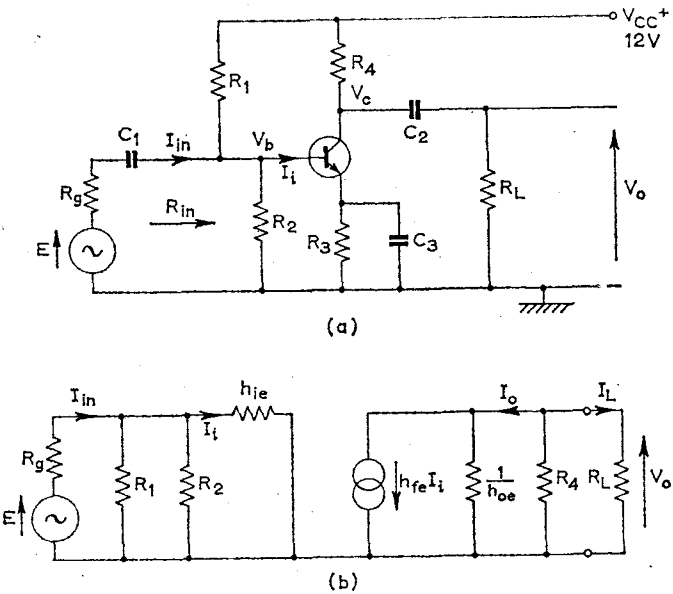


Fig. 21.4

(*a*) The d.c. operating conditions may be obtained by assuming that the base bias current, I_B, is very much smaller than the direct current through R_2, so that the direct voltage on the base, V_B, is given approximately

$$V_B = \frac{V_{CC}R_2}{R_1 + R_2} = \frac{12 \times 10{,}000}{70{,}000} = \underline{\underline{1{\cdot}7\text{V}}}$$

Since a silicon transistor is used it may be assumed that the base-emitter voltage is 0·7V, so that the emitter bias voltage is

$$V_E = V_B - 0{\cdot}7 = \underline{\underline{1\text{V}}}$$

The emitter quiescent current is therefore

$$I_E = \frac{V_E}{R_3} = \underline{\underline{1\text{mA}}}$$

This is also approximately the collector quiescent current, I_C, so that the quiescent voltage on the collector is

$$V_C = V_{CC} - I_C R_4 \approx 12 - 5 \times 10^{-3} \times 1000 = \underline{\underline{7\text{V}}}$$

Note that the quiescent base current is $I_B \approx I_C/h_{FE}$, where h_{FE} is normally approximately the same as h_{fe} (50 in this case). Hence $I_B \approx 1/50 = 0{\cdot}02$mA. The quiescent current through R_2 is $I_{R2} = V_B/R_2 = 1{\cdot}7/10^4 = 0{\cdot}17$mA. This demonstrates that the quiescent base current is indeed small compared to the current through R_2, as was originally assumed.

(*b*) From eqn. (21.12) the mid-frequency current gain of the transistor is given by

$$\frac{-I_0}{I_i} = \frac{h_{fe}\angle 180^\circ}{1 + h_{oe}R'}$$

where R' is the resistance of R_4 and R_L in parallel (i.e. $1{\cdot}1\,\text{k}\Omega$).

The actual stage current gain is $A_i = I_L/I_{in}$ and is less than the transistor gain since (i) the input current is partially shunted by R_1 and R_2, and (ii) the output current is divided between R_4 and R_L. Thus

$$I_i = I_{in}\frac{R_p}{R_p + h_{ie}} \quad \text{where } R_p = \frac{R_1R_2}{R_1 + R_2} = 8{\cdot}6\,\text{k}\Omega$$

and

$$I_L = \frac{-I_0R_4}{R_4 + R_L}$$

Hence the current gain of the stage is

$$A_i = \frac{I_L}{I_{in}} = \frac{-I_0R_4}{(R_4 + R_L)}\,\frac{R_p}{I_i(R_p + h_{ie})}$$

$$= \frac{-h_{fe}R_4R_p}{(1 + h_{oe}R')(R_4 + R_L)(R_p + h_{ie})}$$

Inserting numerical values,

$$A_i = \frac{-50 \times 5 \times 10^3 \times 8{\cdot}6 \times 10^3}{\{1 + (25 \times 10^{-6} \times 1{\cdot}1 \times 10^3)\}(5 + 1{\cdot}4) \times 10^3(8{\cdot}6 + 1{\cdot}4) \times 10^3}$$

$$= -33 = \underline{\underline{33\angle 180^\circ}}$$

The mid-frequency power gain is

$$G = A_i^2\frac{R_L}{R_{in}} = 33^2 \times \frac{1{\cdot}4}{1{\cdot}2} = \underline{\underline{1{,}250}}$$

21.3 Common-base Amplifier

The common-base amplifier was frequently used in the early days of transistors owing to its good high-frequency response. It has a very low input impedance and high output impedance combined with a current gain of less than unity, so that matching transformers are required for cascaded stages. The fractional current gain is the main disadvantage of this connexion. The modern improvement in the high-frequency performance of transistors has meant that the common-base circuit has been largely replaced by the common-emitter circuit, which can be cascaded simply. However, the common-base circuit can be used as a low-to-high impedance-matching stage. This has basically the same buffer action for a current amplifier as the emitter follower (see Chapter 22) has for a voltage amplifier. A typical stage is shown in Fig. 21.5(*a*).

The a.c. small-signal equivalent circuit is shown in Fig. 21.5(*b*), where again the feedback generator has been omitted. The expressions for input and output resistance, current and voltage gain have

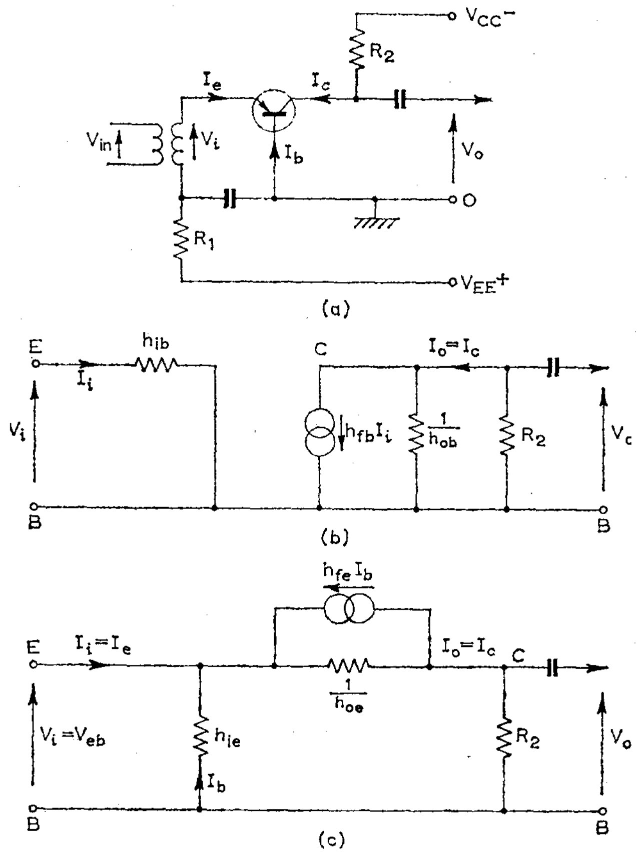


Fig. 21.5 COMMON-BASE AMPLIFIER STAGE

therefore the same form as those derived for the common-emitter stage, but the common base (C.B.) *h*-parameters replace the common-emitter (C.E.) *h*-parameters. Thus

$$V_i = V_{eb} = h_{ib}I_e \tag{21.18}$$

$$I_o = I_c = h_{fb}I_e + h_{ob}V_{cb} \tag{21.19}$$

and

$$\text{Input resistance} = h_{ib} \tag{21.20}$$

$$\text{Output resistance} = 1/h_{ob} \tag{21.21}$$

$$\text{Internal current gain} = \frac{-I_o}{I_i} = \frac{-h_{fb}}{1 + h_{ob}R_L} \tag{21.22}$$

$$\text{Voltage gain} = \frac{-h_{fb}}{1 + h_{ob}R_L}\frac{R_L}{R_{in}} \tag{21.23}$$

The fact that the current gain is positive is accounted for by describing h_{fb} as a negative constant. There is, of course, no phase reversal through the amplifier—as is apparent since the input current (the emitter current) flows across to the collector to become the output current. Expressions similar to eqns. (21.15)–(21.17) are obtained if the full h-parameter circuit is used. As with the common-emitter circuit, and because of the production spreads of transistor parameters and the small value of h_{rb} it is normally sufficiently accurate to use the simplified circuit.

Frequently only common-emitter parameters are quoted by manufacturers, so that it is often more convenient simply to rearrange the common-emitter equivalent circuit for common-base operation as shown in Fig. 21.5(c). Then, from eqn. (21.9),

$$\begin{aligned} I_c &= h_{fe}I_b + h_{oe}V_{ce} \\ &= -h_{fe}(I_c + I_e) + h_{oe}V_{ce} \end{aligned}$$

since $I_b + I_c + I_e = 0$. Hence

$$\begin{aligned} I_c &= \frac{-h_{fe}I_e}{1 + h_{fe}} + \frac{h_{oe}V_{ce}}{1 + h_{fe}} \\ &\approx \frac{-h_{fe}I_e}{1 + h_{fe}} + \frac{h_{oe}V_{cb}}{1 + h_{fe}} \end{aligned} \tag{21.24}$$

since $V_{cb} = V_{ce} + V_{eb}$; and if the voltage gain is high, $V_{cb} \gg V_{eb}$, so that $V_{cb} \approx V_{ce}$.

Comparing eqns. (21.24) and (21.19), it will be seen that

$$h_{fb} = \frac{-h_{fe}}{1 + h_{fe}} \tag{21.25}$$

$$h_{ob} = \frac{h_{oe}}{1 + h_{fe}} \tag{21.25a}$$

It is sometimes convenient to remember that $h_{ob} \approx 1/r_c$ and $h_{fb} \approx -a$ in the common-base equivalent-T circuit.

Also

$$\begin{aligned} V_{eb} &= -h_{ie}I_b = h_{ie}I_e + h_{ie}I_c \\ &= h_{ie}I_e - \frac{h_{ie}h_{fe}}{1 + h_{fe}} I_e + \frac{h_{ie}h_{oe}}{1 + h_{fe}} V_{cb} \quad \text{(from eqn. (21.24))} \\ &= \frac{h_{ie}I_e}{1 + h_{fe}} + \frac{h_{ie}h_{oe}}{1 + h_{fe}} V_{cb} \end{aligned}$$

Then, since $h_{ie}h_{oe}V_{cb} \ll h_{ie}I_e$,

$$V_{eb} \approx \frac{h_{ie}}{1 + h_{fe}} I_e \tag{21.26}$$

Comparing eqns. (21.26) and (21.18),

$$h_{ib} \approx \frac{h_{ie}}{1 + h_{fe}} \tag{21.27}$$

It will thus be seen that the input resistance, output *admittance* and current gain of the C.B. amplifier are less by a factor of $(1 + h_{fe})$ than the corresponding quantities for the C.E. amplifier. Owing to the low input and high output resistances, however, there is both voltage and power gain in a common-base stage. It is also evident that, since $h_{fb} = -h_{fe}/(1 + h_{fe})$, the value of h_{fb} will always be slightly less than unity. Note that h_{fe} corresponds to a term which is frequently found in manufacturers specifications under the symbols α' or β, and that h_{fb} corresponds to a.

21.4 Low-frequency Response of C.E. Amplifier

EFFECT OF COUPLING CAPACITORS

Consider the common-emitter stage shown in Fig. 21.6(a), which is fed from an a.c. source, I, of internal impedance R_g. If the reactance of C_3 is small enough, the impedance, R_{in}, seen looking into the transistor from the coupling capacitor C_1, consists of R_1, R_2 and h_{ie} in parallel (i.e. the impedance to the signal current), so that the equivalent a.c. input circuit is as shown in Fig. 21.6(b). At mid-frequencies, the reactance of C_1 is negligibly small, and the current, I_{in}, is given by

$$I_{in} = \frac{IR_g}{R_g + R_{in}}$$

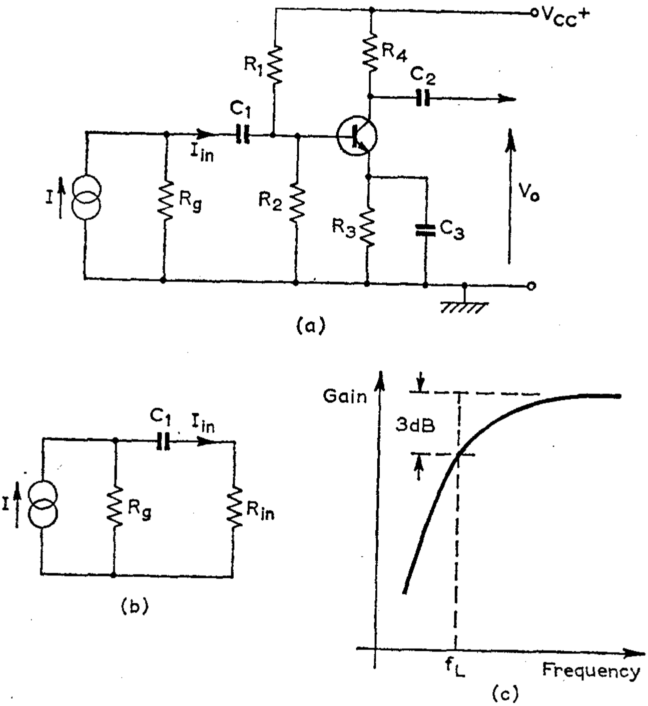


Fig. 21.6 COMMON-EMITTER AMPLIFIER STAGE

The low-frequency response is determined by the coupling and emitter bypass capacitors, C_1–C_3

As the frequency falls, however, the reactance of C_1 must be taken into account, and the current I_{in}' is then

$$I_{in}' = \frac{IR_g}{R_g + R_{in} + \dfrac{1}{j\omega C_1}}$$

As the frequency falls the term $1/j\omega C_1$ becomes comparable in size with $(R_g + R_{in})$ so that the current I'_{in} becomes smaller than I_{in}. The effective current gain I_o/I will therefore fall. At mid-frequencies I_o/I has the value

$$A_{io} = \frac{I_o}{I} = \frac{I_o}{I_{in}} \times \frac{I_{in}}{I} = \frac{-h_{fe}}{1 + h_{oe}R_L} \frac{R_g}{R_g + R_{in}}$$

at low frequencies this ratio becomes

$$A_i = \frac{I_o}{I} = \frac{I_o}{I_{in}'}\frac{I_{in}'}{I_o} = \frac{-h_{fe}}{1 + h_{oe}R_L}\frac{R_g}{R_g + R_{in} + 1/j\omega C_1}$$

The difference between the expressions for A_{io} and A_i is the inclusion of the term $1/j\omega C_1$ in the denominator. This term causes a reduction in the size of I_o/I and also an increase in its phase angle. A convenient way of describing this effect is to determine the frequency for which

$$R_g + R_{in} = \frac{1}{\omega C_1}$$

At this frequency (called the lower cut-off frequency, f_L) $R_g + R_{in} + 1/j\omega C_1$ may be written as

$$\begin{aligned} R_g + R_{in} + \frac{1}{j2\pi f_L C_1} &= (R_g + R_{in}) + \frac{1}{j}(R_g + R_{in}) \\ &= (R_g + R_{in})(1 - j1) \\ &= \sqrt{2}\,(R_g + R_{in})\,\underline{/-45^\circ} \end{aligned}$$

i.e. the current gain will be $1/\sqrt{2}$ of its mid-frequency value and the phase angle of the current gain will be *increased* by 45°. The frequency at which this occurs is given by

$$2\pi f_L C_1 = \frac{1}{R_g + R_{in}}$$

i.e.

$$f_L = \frac{1}{2\pi C_1(R_g + R_{in})} \tag{21.28}$$

The current gain at any frequency, f, can now be written as

$$\begin{aligned} A_i &= \frac{-h_{fe}}{1 + h_{oe}R_L}\frac{R_g}{R_g + R_{in} + f_L/j2\pi f f_L C_1} \\ &= \frac{-h_{fe}}{1 + h_{oe}R_L}\frac{R_g}{R_g + R_{in} + (R_g + R_{in})f_L/jf} \end{aligned}$$

Hence

$$A_i = \frac{A_{io}}{1 + f_L/jf} = \frac{A_{io}}{1 - jf_L/f} \tag{21.29}$$

When $f = f_L$, $A_i = A_{io}/(1 - j1) = A_{io}/\sqrt{2}\underline{/-45^\circ}$ as derived previously. Note that *for this condition*

$$\begin{aligned} 20 \log_{10}(A_i/A_{io}) = -20 \log_{10}\sqrt{2} &= -10 \log_{10}2 \\ &= -3 \text{ dB} \end{aligned}$$

The frequency f_L is often called the lower 3 dB cut-off frequency. The bandwidth of an amplifier is arbitrarily taken to extend between those frequencies for which the gain falls by 3 dB from the mid-frequency value (the upper 3 dB cut-off frequency is dealt with in Section 21.5).

Similar conditions apply to the voltage gain of the stage.

Very often it is sufficiently accurate when calculating the value of C_1 required for a given low-frequency response to assume that R_1 and R_2 are large enough to be neglected compared with h_{ie} so that R_{in} in eqn. (21.28) becomes simply h_{ie}. It is further obvious from eqn. (21.28) that for a given coupling capacitor, C_1, the higher the value of R_g the lower will be the cut-off frequency f_L. Hence a voltage signal source (which will normally be of low impedance) gives a poorer low-frequency performance than a current signal source (which will be of high impedance—e.g. a preceding common-base or common-emitter amplifier stage).

It is evident that the second coupling capacitor, C_2, will produce a similar effect, which can be calculated in the same way. The approximate shape of the gain/frequency characteristic for low frequencies is shown in Fig. 21.6(*c*).

EFFECT OF EMITTER BYPASS CAPACITOR

The emitter bypass capacitor, C_3, is required in order to connect the emitter to earth with respect to the signal. If it were removed there would be a signal feedback effect and a consequent fall in gain, which

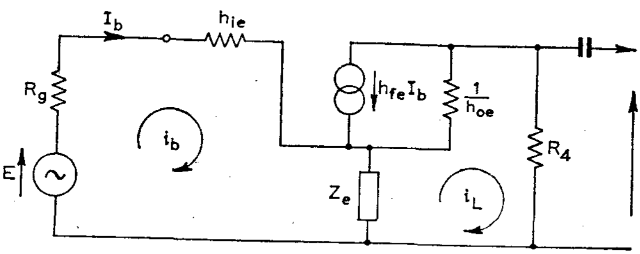


Fig. 21.7 A.C. EQUIVALENT CIRCUIT OF FIG 21.6, INCLUDING EMITTER-CIRCUIT RESISTANCE

will be discussed further in the next chapter. At low frequencies the reactance of C_3 increases, and hence the emitter may no longer be considered as an earth point for the signal. This reduces the gain of the stage by an amount which may be determined as follows.

The equivalent a.c. circuit for the stage shown in Fig. 21.6 is shown in Fig. 21.7, including a finite emitter impedance, Z_e, and neglecting the effect of the coupling capacitors C_1 and C_2. The signal source is represented by the equivalent constant-voltage generator of e.m.f. E and internal impedance R_g. For simplicity the bias resistors, R_1 and R_2, are assumed high enough in value to carry negligible signal current, and additional loading on the stage is assumed incorporated in R_4. The mesh equations can then be written (using small letters for mesh currents) as

Mesh 1
$$E = (R_g + h_{ie} + Z_e)i_b - Z_e i_L$$
$$\approx (R_g + h_{ie})i_b - Z_e i_L \tag{21.30}$$

Mesh 2
$$0 = -Z_e i_b + \left(\frac{1}{h_{oe}} + Z_e + R_4\right) i_L + \frac{h_{fe}}{h_{oe}} I_b$$
$$\approx \frac{h_{fe}}{h_{oe}} I_b + \left(\frac{1}{h_{oe}} + R_4\right) i_L \tag{21.31}$$

and $i_L \gg I_b$; $i_b = I_b$; hence $Z_e i_b \ll Z_e i_L$. Also $Z_e \ll 1/h_{oe}$.

From eqns. (21.30) and (21.31),

$$E = \frac{-(R_g + h_{ie})\left(\dfrac{1}{h_{oe}} + R_4\right) i_L}{h_{fe}/h_{oe}} - Z_e i_L$$
$$= -\frac{(R_g + h_{ie})(1 + h_{oe}R_4) + Z_e h_{fe}}{h_{fe}} i_L$$

so that

$$\frac{V_o}{E} = \frac{i_L R_4}{E} = \frac{-R_4 h_{fe}}{(R_g + h_{ie})(1 + h_{oe}R_4) + Z_e h_{fe}} \tag{21.32}$$

As $Z_e \to 0$ this ratio gives the mid-frequency value of gain:

$$\left.\frac{V_o}{E}\right|_{\text{mid-}f} = \frac{-R_4 h_{fe}}{(R_g + h_{ie})(1 + h_{oe}R_4)}$$

If it is now assumed that Z_e is the reactance of C_3 alone (i.e. $Z_e = 1/j\omega C_3$), the ratio of V_o/E, or the voltage gain, will be 3 dB down on its mid-frequency value when the reactive component of the denominator in eqn. (21.32) is equal to the resistive component, i.e. when

$$\frac{h_{fe}}{\omega C_3} = (R_g + h_{ie})(1 + h_{oe}R_4)$$

or

$$\frac{1}{\omega C_3} = \frac{(R_g + h_{ie})(1 + h_{oe}R_4)}{h_{fe}}$$

Hence the frequency for a 3dB fall due to C_3 alone is

$$f_L \approx \frac{h_{fe}}{2\pi(R_g + h_{ie})(1 + h_{oe}R_4)C_3} \tag{21.33}$$

This again shows that the higher the source impedance, R_g, the lower will be the 3dB cut-off frequency. Also note that, since $h_{oe}R_4$ is generally less than 1, the value of C_3 is generally h_{fe} times the value of the coupling capacitor for the same low-frequency cut-off.

In the actual circuit both the coupling-capacitor and emitter-capacitor effects will occur together. Generally the leakage through C_3 is less critical than that through C_1 or C_2, so that it may economically be chosen to have a high enough capacitance to make the emitter effect occur at a much lower frequency than the cut-off frequency due to C_1 or C_2.

The emitter circuit resistance, R_3, can affect the result, but if $1/\omega C_3 < 0{\cdot}1R_3$ at the low-frequency cut-off the effect is small.

EXAMPLE 21.2 The transistor used in the circuit of Fig. 21.6(*a* has $h_{oe} = 40\,\mu$S; $h_{ie} = 1{\cdot}5\,$kΩ; $h_{fe} = 100$; R_1 and R_2 are large enough to be neglected compared with h_{ie} at signal frequencies; $R_3 = 1\,$kΩ; and $R_4 = 5\,$kΩ. Find suitable values for C_1 and C_3 to give a low-frequency cut-off at 15 Hz for a source resistance of 5 kΩ. If the source resistance is reduced to 500 Ω what will be the low-frequency cut-off?

(*a*) From eqn. (21.28), the coupling capacitance is

$$C_1 = \frac{1}{2\pi f_L(R_g + R_{in})} \quad \text{where } R_{in} \text{ is taken as equal to } h_{ie}$$

$$= \frac{10^6}{2\pi \times 15(5 + 1{\cdot}5) \times 10^3} = 1{\cdot}63, \text{ or say } \underline{\underline{2\mu\text{F}}}$$

If the emitter decoupling capacitor is to have only a small effect at this frequency then, from eqn. (21.32),

$$C_3 \gg \frac{h_{fe}}{\omega(R_g + h_{ie})(1 + h_{oe}R_4)} \gg \frac{100}{2\pi \times 15 \times 6{\cdot}5 \times 10^3 \times 1{\cdot}2} \gg 150\mu\text{F}$$

A suitable value for C_3 would be $\underline{\underline{500\mu\text{F}}}$.

(*b*) For a source resistance of 500 Ω, eqn. (21.28) gives

$$f_L = \frac{1}{2\pi C_1(R_g + R_{in})} = \frac{10^6}{2\pi \times 2 \times 2 \times 10^3} = \underline{\underline{40\,\text{Hz}}}$$

21.5 Transistor Characteristics at High Frequencies

Whereas the low-frequency response of transistor amplifiers is determined mainly by the associated circuit components, the high-frequency response is largely a function of the transistor itself. By far the most important high-frequency effect is due to the time taken for charge carriers to cross from the emitter junction to the collector junction. This is called *base transit time*. In addition, the junction capacitances and stray circuit-wiring capacitances will affect the high-frequency performance of transistor amplifiers, but these effects are of less importance in general.* The following treatment represents a much simplified approach. In practical transistors, different physical constructions give different high-frequency characteristics, which are only approximated by the following derivation.

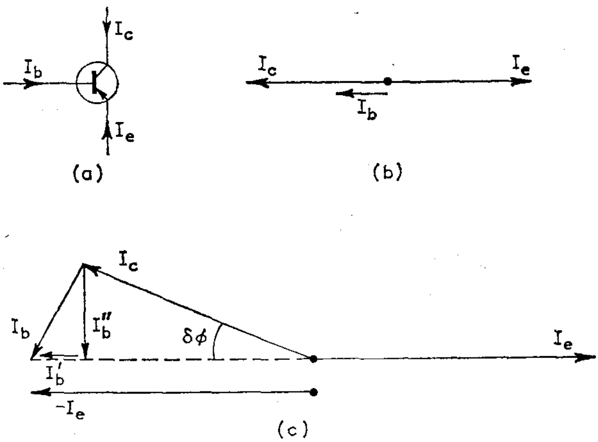


Fig. 21.8 EFFECT OF TRANSIT TIME ON THE HIGH-FREQUENCY RESPONSE OF A TRANSISTOR

In order to understand the physical basis for the loss of gain at high frequencies in a transistor, consider the device as a "black box" in which there is an input current, $\boldsymbol{I_e}$. This current divides between the collector and the base; i.e. if $\boldsymbol{I_e}$ enters the device, $\boldsymbol{I_c}$ and $\boldsymbol{I_b}$ must leave it—this corresponds to the physical operation. In the purely mathematical convention adopted earlier, all alternating currents are assumed to flow *into* the device as shown in Fig. 21.8(*a*). It follows that the currents are related both instantaneously and in complexor representation by the equation

$$\boldsymbol{I_e} = -(\boldsymbol{I_b} + \boldsymbol{I_c})$$

At low frequencies the transit time is only a small fraction of the period of the signal, and hence there will be negligible delay through the transistor and the complexor diagram will be as shown in Fig. 21.8(*b*), the 180° phase reversal being due to the conventional current directions chosen.

* The situation, however, is very dependent on current advances in production technology, particularly with silicon transistors.

Under short-circuit conditions at the output, eqn. (21.9) gives

$$I_c = h_{feo} I_b$$

where h_{feo} is the low-frequency value of h_{fe}.

In this equation, I_b is the normal recombination base current of the transistor.

As the frequency increases the time taken for the charge carriers to cross the base becomes a significant fraction of the signal period. The collector current, I_c, will therefore lag behind its low-frequency phase position by a small angle $\delta\phi = 2\pi f \tau_B$, where τ_B is the base transit time. This follows since for a signal period $T = 1/f$ second, a delay of T corresponds to a phase-angle lag of 2π radians, and hence a delay of τ_B corresponds to a lag of $2\pi\tau_B/T = 2\pi f\tau_B$ radians. But the base current, I_b, is equal to $-(I_e + I_c)$, so that I_b must be the complexor joining the end points of I_c and $-I_e$, as shown in Fig. 21.8(*c*). This current can be resolved into two components, I_b' in phase with $-I_e$, and I_b'' leading $-I_e$ by 90°. The current I_b' represents the normal low-frequency base-recombination current I_c/h_{feo}.

Assuming that $\delta\phi$ is small,

$$I_b'' \approx I_c \delta\phi = I_c 2\pi f \tau_B$$

and leads $-I_e$ by 90°. Hence

$$I_b = I_b' + I_b'' \approx \frac{I_c}{h_{feo}} + j2\pi f \tau_B I_c$$

The complex short-circuit current gain is defined as $h_{fe} = I_c/I_b$, so that

$$h_{fe} = \frac{1}{\dfrac{1}{h_{feo}} + j2\pi f \tau_B} = \frac{h_{feo}}{1 + j2\pi h_{feo} f \tau_B} \qquad (21.34)$$

Obviously if τ_B is very small so that $2\pi h_{feo} f \tau_B \ll 1$, then $h_{fe} = h_{feo}$, and further, the value of h_{fe} will fall by 3 dB and will introduce a phase *lag* of 45° when $2\pi f h_{feo} \tau_B = 1$. The frequency at which this occurs is often called the β *cut-off frequency*, f_β or alternatively f_{hfe}, and is given by

$$f_{hfe} = f_\beta = \frac{1}{2\pi h_{feo} \tau_B} \qquad (21.35)$$

It is now more usual to quote the frequency, f_1, at which h_{fe} falls to unity. This is obtained by letting f in eqn. (21.34) become so large that $2\pi f h_{feo} \tau_B \gg 1$. Then

$$h_{fe} \approx \frac{h_{feo}}{2\pi f h_{feo} \tau_B} = \frac{1}{2\pi f \tau_B}$$

This will have a magnitude of unity when

$$f = f_1 = \frac{1}{2\pi\tau_B} \tag{21.36}$$

Combining eqns. (21.35) and (21.36), the following important relation is obtained:

$$f_{hfe}h_{feo} = \frac{1}{2\pi\tau_B} = f_1$$

Also, for frequencies at which $2\pi f h_{feo}\tau_B \gg 1$, eqn. (21.34) gives

$$h_{fe}f \approx \frac{h_{feo}f}{2\pi f h_{feo}\tau_B} = \frac{1}{2\pi\tau_B} = f_1 \tag{21.37}$$

This is called the *gain-bandwidth product* of the transistor, and is an important design parameter.

Sometimes a frequency, f_T, is quoted in relation to transistors. This is the frequency at which the magnitude of $\boldsymbol{h}_{fe}$ falls to unity if the fall in gain is constant at 6dB per octave from the break frequency f_{hfe}. Normally f_T is only slightly different from f_1.

From eqns. (21.34) and (21.35), a general expression for $\boldsymbol{h}_{fe}$ can be written as

$$\boldsymbol{h}_{fe} = \frac{h_{feo}}{1 + j\dfrac{f}{f_\beta}} \tag{21.38}$$

Although this applies only so long as the frequency is below or not much above f_β, it gives some approximation to the actual characteristic up to the frequency f_1.

21.6 High-frequency Response of C.E. Amplifier

The characteristics of a common-emitter amplifier at high frequencies can be obtained by substituting the expression for $\boldsymbol{h}_{fe}$ given by eqn. (21.38) in eqns. (21.10)–(21.17). Thus the internal current gain is

$$\begin{aligned} A_i &= \frac{-\boldsymbol{h}_{fe}}{1 + h_{oe}R_L} \\ &= \frac{-h_{feo}}{\left(1 + j\dfrac{f}{f_\beta}\right)(1 + h_{oe}R_L)} \end{aligned} \tag{21.39}$$

This is 3dB down on the mid-frequency gain when $f = f_\beta\ (= f_{hfe})$

and falls at a rate of 6dB per octave for frequencies above f_β (i.e. for each doubling in frequency the gain falls by a further 6dB).* Also when the frequency is f_β there will be an added 45° phase lag. A similar situation exists for voltage gain. In both cases there is an increase of phase lag up to an additional 90° as the frequency approaches infinity.

It is obvious from Fig. 21.8(*c*) that at sufficiently high frequencies there is a component of base (input) current which leads the normal low-frequency base current by 90°—i.e. there is an added capacitive component of input impedance. This is also demonstrated by substituting the expression for h_{fe} of eqn. (21.38) in eqn. (21.15), which gives the exact expression for input impedance.

21.7 High-frequency Response of C.B. Amplifier

The response of the common-base amplifier to high-frequency signals is obtained by substituting for h_{fe} from eqn. (21.38) in the expression representing the operation of this type of amplifier. Thus from eqn. (21.22) the current gain of the common-base amplifier is

$$A_i = \frac{-h_{fb}}{1 + h_{ob}R_L} = \frac{h_{fe}}{(1 + h_{fe})(1 + h_{ob}R_L)}$$
$$= \frac{h_{feo}/(1 + jf/f_\beta)}{\left\{1 + \dfrac{h_{feo}}{1 + jf/f_\beta}\right\}(1 + h_{ob}R_L)}$$
$$= \frac{h_{feo}}{(1 + h_{feo} + jf/f_\beta)(1 + h_{ob}R_L)}$$

Thus as the frequency increases the current gain falls, owing to the term jf/f_β, and is 3dB down on its mid-frequency value of $h_{feo}/(1 + h_{feo})(1 + h_{ob}R_L)$, when

$$1 + h_{feo} = \frac{f}{f_\beta}$$

* The ratio $\dfrac{\text{Gain at a high frequency, } f}{\text{Gain at mid-frequencies}} = \dfrac{1}{1 + j\dfrac{f}{f_\beta}}$

so that

$$\text{Gain in decibels} = 20 \log_{10} \left| \frac{1}{(1 + jf/f_\beta)} \right| = -20 \log_{10} \sqrt{(1 + f^2/f_\beta^2)}$$
$$= -20 \log_{10} (f/f_\beta) \quad \text{for } f \gg f_\beta$$

Hence for a frequency of $2f$ (an octave higher than f),

$$\text{Gain} = -20 \log_{10} 2f/f_\beta = -20 \log_{10} 2 - 20 \log_{10} f/f_\beta$$
$$= (\text{Gain at frequency } f - 6) \text{ decibels}$$

The frequency at which this occurs is called the α *cut-off frequency* f_α, given by

$$f_\alpha = (1 + h_{feo})f_\beta \tag{21.40}$$

The expression for current gain can therefore be written

$$A_i = \frac{-h_{fbo}}{\left(1 + j\dfrac{f}{f_\alpha}\right)(1 + h_{ob}R_L)} \quad \text{where} \quad h_{fbo} = \frac{h_{feo}}{1 + h_{feo}}$$

At the frequency f_α there will be a phase lag of 45°.

The response characteristics of a transistor in common-emitter and common-base connexions are shown in Fig. 21.9, where gain

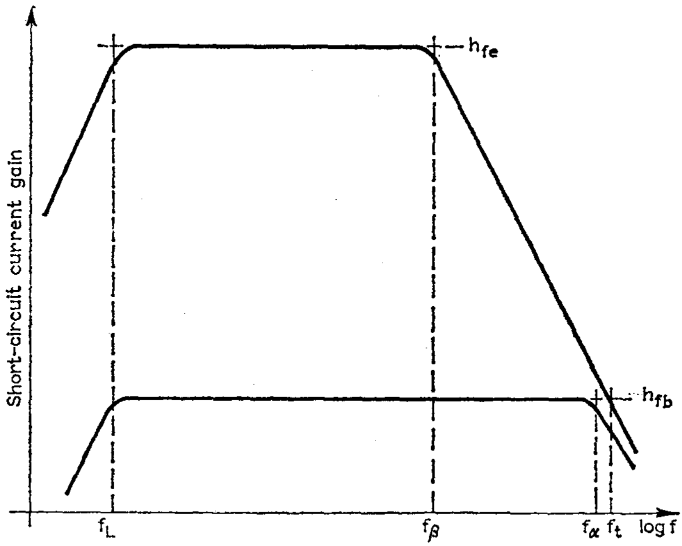


Fig. 21.9 FREQUENCY RESPONSE CHARACTERISTICS OF TRANSISTORS IN C.E. AND C.B. CONNEXIONS

in decibels is plotted against the logarithm of frequency. The bandwidth of the amplifier is the frequency range between the lower and upper cut-off frequencies. For a good high-frequency transistor the base transit time may be as short as 0·125 ns (i.e. $0{\cdot}125 \times 10^{-9}$ second), giving

$$f_1 = \frac{1}{2\pi \times 0{\cdot}125 \times 10^{-9}} \approx 1\,\text{GHz}$$

Table 21.2 summarizes the characteristics of common-base and common-emitter amplifiers at mid-frequencies. The numerical values given are to be considered as giving orders of magnitude, and depend on the values of load and generator resistances. Thus the input resistance in common-base connexion rises significantly as the load resistance, R_L, increases (above about 10kΩ) while that in

TABLE 21.2

Connexion	Input resistance	Output resistance	Short-circuit current gain	Voltage gain	Mid-frequency phase characteristic
Common emitter	medium 1·5kΩ (Ge) 2·5kΩ (Si)	medium 25kΩ (Ge) 70kΩ (Si)	high (20 – over 200)	high	180°
Common base	low 25Ω (Ge) 40Ω (Si)	high 1 MΩ (Ge) 1 MΩ (Si)	low <1	high	0°

common-emitter connexion falls slightly as R_L increases. Similarly the output resistance in common-base connexion falls appreciably as the source resistance, R_g, falls (below about 1kΩ) while that in common-emitter connexion rises slightly as R_g falls.

If the collector resistance $R_L \ll 1/h_{oe}$, then C.E. and C.B. amplifiers have approximately the same voltage gain. If R_L is very high the C.B. connexion will give the higher voltage gain.

21.8 Tuned Transistor Amplifiers

A simple tuned amplifier with transformer input is shown in Fig. 21.10(*a*), with the simplified equivalent a.c. circuit at (*b*). At resonance (angular frequency, $\omega_0 = 1/\sqrt{(LC)}$) the impedance of the tuned circuit is $R_0 = L/Cr$ and at an angular frequency ω near resonance it is

$$Z = \frac{R_0}{1 + j2Q\Delta} \tag{21.41}$$

where $Q = \omega L/r = R_0/\omega L$ and $\Delta \approx (\omega - \omega_0)/\omega_0$ is the per-unit frequency deviation.

Capacitors C_2 and C_3 provide a low-impedance path between input and emitter for signal currents. Their capacitances may be

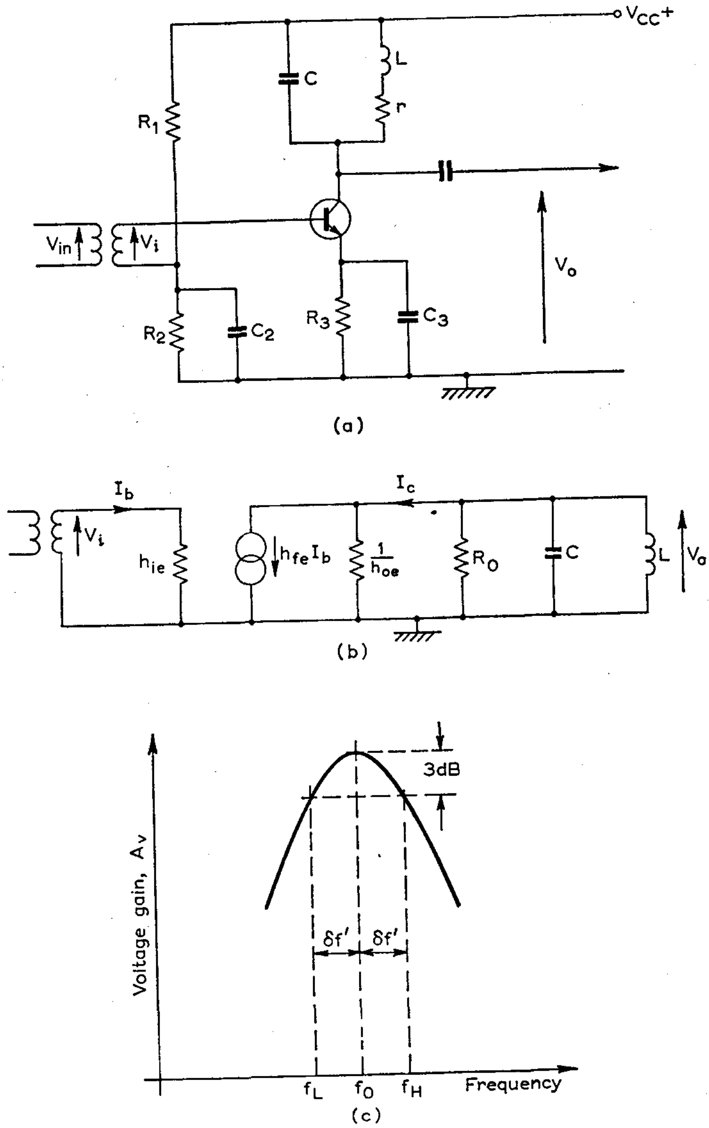


Fig. 21.10 TUNED TRANSISTOR AMPLIFIER

relatively small, since they are only required to bypass signal-frequency currents, and not low-frequency currents. Hence the input current is given by

$$I_b = \frac{V_i}{h_{ie}} \tag{21.42}$$

and the output current is

$$I_c = h_{fe}I_b + V_o h_{oe}$$

Hence the output voltage, V_o, is

$$V_o = -I_c Z = -\frac{(h_{fe}I_b + V_o h_{oe})R_0}{1 + j2Q\Delta}$$

or, from eqns. (21.41) and (21.42),

$$V_o + \frac{V_o R_0 h_{oe}}{1 + j2Q\Delta} = \frac{-h_{fe}V_i R_0}{h_{ie}(1 + j2Q\Delta)}$$

The voltage gain is thus

$$A_v = \frac{V_o}{V_i} = \frac{-h_{fe}R_0}{(1 + h_{oe}R_0 + j2Q\Delta)h_{ie}} \tag{21.43}$$

The relationship between gain and frequency around resonance is shown in Fig. 21.10(*c*). At resonance, $\Delta = 0$, and the gain has a maximum value of

$$A_{vm} = \frac{-h_{fe}R_0}{(1 + h_{oe}R_0)h_{ie}} \tag{21.44}$$

It follows that the gain falls 3dB below its resonance value when (from eqn. (21.43))

$$2Q\Delta = 1 + h_{oe}R_0$$

i.e. when

$$\Delta = \frac{\delta f'}{f_0} = \frac{1 + h_{oe}R_0}{2Q} = \frac{1}{2Q'} \tag{21.45}$$

where $\delta f'$ is the half bandwidth at the 3dB points, and Q' is the effective Q-factor of the circuit when it is loaded by the output resistance $(1/h_{oe})$ of the transistor; i.e.

$$Q' = \frac{R_{eq}}{\omega L}$$

where R_{eq} is R_0 in parallel with $1/h_{oe}$:

$$R_{eq} = \frac{R_0(1/h_{oe})}{R_0 + 1/h_{oe}} = \frac{R_0}{1 + h_{oe}R_0}$$

Hence

$$Q' = \frac{R_0}{(1 + h_{oe}R_0)\omega L} = \frac{Q}{1 + h_{oe}R_0}$$

The voltage gain (eqn. (21.43)) can thus be expressed in terms of R_{eq} and Q' as

$$A_v = \frac{-h_{fe}R_{eq}}{(1 + j2Q'\Delta h_{ie})} \tag{21.46}$$

If the lower and upper frequencies at which the gain falls 3 dB below the resonance value are f_L and f_H, then

$$f_L = f_0 - \delta f' = f_0 - \frac{f_0}{2Q'} \tag{21.47}$$

and

$$f_H = f_0 + \delta f' = f_0 + \frac{f_0}{2Q'} \tag{21.48}$$

while the overall bandwidth in hertz is

$$B = f_H - f_L = \frac{f_0}{Q'} \tag{21.49}$$

Note that the bandwidth between half-power frequencies of the unloaded tuned circuit is given by $B = f_0/Q$.

It is thus apparent that the active device lowers the effective Q-factor and increases the bandwidth between half-power points. In a C.E. transistor amplifier $1/h_{oe}$ may be appreciably smaller than R_0, and the bandwidth may thus be undesirably wide (the whole object of a tuned amplifier usually being to give selective amplification over a narrow bandwidth around the resonant frequency). In order to overcome this the Q-factor of the circuit must be increased. This may be done by increasing the C/L ratio while keeping the product LC constant (for a given resonant frequency).* However this can lead to unrealistically large values for C (if it is a tuning capacitor).

* $Q = \dfrac{\omega L}{r} = \dfrac{\omega L}{L/CR_0}$

since $R_0 = L/Cr$, where r is the series resistance of L; i.e.

$$Q = \omega C R_0 = R_0\sqrt{\frac{C}{L}}$$

since at resonance $\omega = 1/\sqrt{(LC)}$.

Hence for a high Q, C/L must be large. Note also that if L is reduced, then r will also fall; hence, since L is proportional to the square of the number of turns, and r, to the number of turns, Q will increase.

An alternative approach is to connect the transistor across a tapping of the tuning inductor as shown in Fig. 21.11(*a*). This gives the simplified equivalent circuit shown at (*b*), which may be further reduced to the circuit shown at (*c*), assuming that there is perfect coupling between the primary and secondary windings. The tapping is taken at a fraction, k_t, of the total number of turns. It is left as

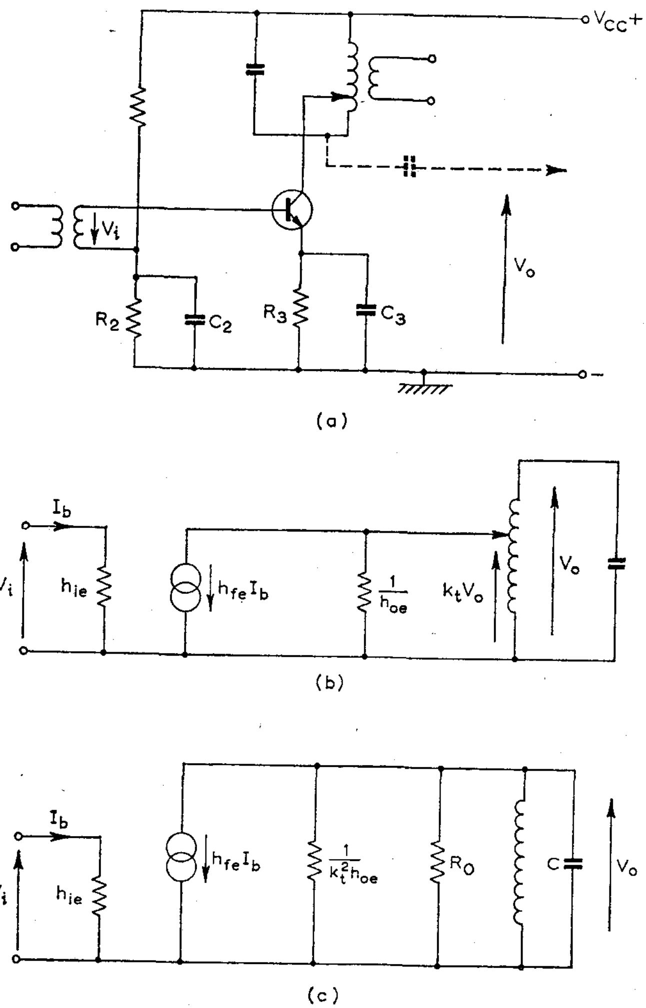


Fig. 21.11 USE OF A TAPPED COIL TO INCREASE THE EFFECTIVE Q-FACTOR OF A TUNED TRANSISTOR AMPLIFIER

an exercise for the reader to verify that circuits (*b*) and (*c*) are in fact equivalent.

The resultant Q-factor is then $Q' = R_{eq}'/\omega L$ where R_{eq}' is the resistance of R_0 in parallel with $1/(k_t^2 h_{oe})$. Without the tapping the value of R_{eq} has already been shown to be R_0 in parallel with $(1/h_{oe})$, so that the tapping has increased R_{eq} ($k_t < 1$). The voltage gain becomes

$$A_v = \frac{-k_t h_{fe} R_{eq}'}{(1 + j2Q'\Delta)h_{ie}} \qquad (21.50)$$

where

$$R_{eq}' = \frac{R_0}{(1 + k_t^2 h_{oe} R_0)} \qquad (21.51)$$

A further complication arises in transistor tuned amplifiers, since the input resistance, R_{in}, of the following stage may be low, and its input capacitance, C_{in}, high. This can be partially compensated by taking the output from a secondary winding as shown in Fig. 21.11. Then if the ratio of primary to secondary turns is k_{t2}, the input will reflect a parallel resistance $k_{t2}^2 R_{in}$ into the tuned circuit (where $k_{t2} > 1$). It will also reflect a parallel capacitance C_{in}/k_{t2}^2 into the tuned circuit, so that the effects of R_{in} and C_{in} on bandwidth and tuning are very much reduced.

Tuned amplifiers are designed on a *power gain* basis. Hence $1/h_{oe}$ is matched to the load by suitable transformer tapping. The tuned circuit acts as a negligible power-loss shunt at resonance. Off resonance it absorbs more signal power, and so reduces the gain.

21.9 Neutralization and Unilaterization

The capacitance, C_{bc}, which exists between the base and collector of a transistor can result in a part of the output voltage of a tuned amplifier being fed back to the input base connexion, as shown in Fig. 21.12(*a*) where only the a.c. components of the circuit have been shown. As will be seen in Chapter 22, this feedback can have undesirable effects, and may result in the amplifier becoming an oscillator. It can be eliminated by extending the tuning inductance, L_1, by a few turns, and connecting a *neutralizing capacitor*, C_n, between this extension and the base input. Then L_1, L_2, C_{bc} and C_n form a simple bridge circuit, in which at balance

$$\frac{V_{o1}}{V_{on}} = \frac{1/C_{bc}}{1/C_n} \qquad \text{or} \qquad C_n = \frac{V_{o1}}{V_{on}} C_{bc}$$

(i.e. the neutralizing current is equal to the feedback current, so that the alternating voltage between B and E is zero, and no base current can flow). Assuming perfect coupling the ratio of voltages is the same as the turns ratio, N_1/N_2, so that

$$C_n = \frac{N_1}{N_2} C_{bc} \qquad (21.52)$$

A practical alternative is to use the secondary winding to provide a connexion for the neutralizing capacitor as shown at (*b*).

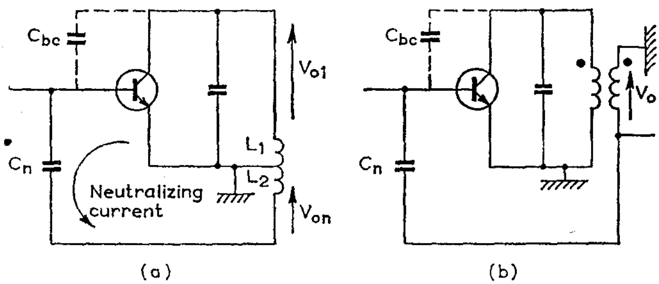


Fig. 21.12 NEUTRALIZATION OF A TUNED AMPLIFIER TO PREVENT OSCILLATION

In some instances the internal feedback inherent in the transistor can provide the equivalent of a resistance connected across C_{bc}. In this case a resistance must also be connected in parallel with C_n, and the process is then called *unilaterialization.* In practice the parallel combination of C_n with a resistor would give a d.c. path, and it is usual to employ the equivalent series combination. This means that unilateralization is achieved at one particular frequency only.

Both these processes are nowadays important only at the highest frequencies, due to the cheapness and improved quality of high-frequency transistors, which means incidentally that they can be used for low-frequency amplification.

21.10 Transistor Amplifiers in Cascade

The gain of a single-stage small-signal transistor amplifier is normally between 30 and 50dB. To increase the gain, amplifiers are connected in cascade and a typical two-stage amplifier is shown in Fig. 21.13.

Equivalent circuits may be used to determine the overall gain of such amplifiers. The load on the first stage is then the input resistance of the second stage, comprising the parallel combination of the bias resistors and the input resistance of the second transistor. The effective input current to the second transistor is only that fraction

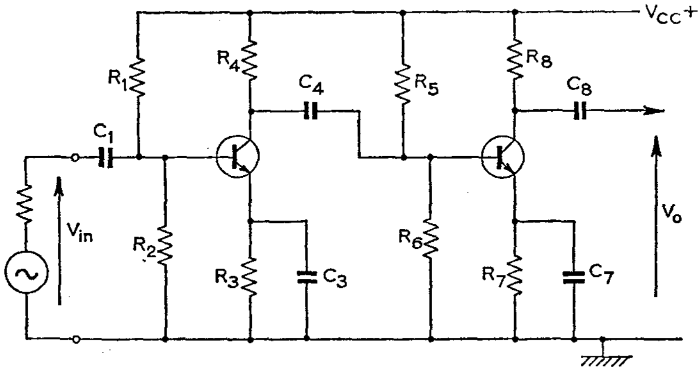


Fig. 21.13 TYPICAL TWO-STAGE *RC*-COUPLED TRANSISTOR AMPLIFIER

of the output current of the first transistor which flows into the second transistor base.

Similar considerations regarding overall bandwidth apply as in cascaded valve amplifiers.

Sometimes the complementary properties of transistors are used to give a cascaded circuit with direct coupling. This improves the

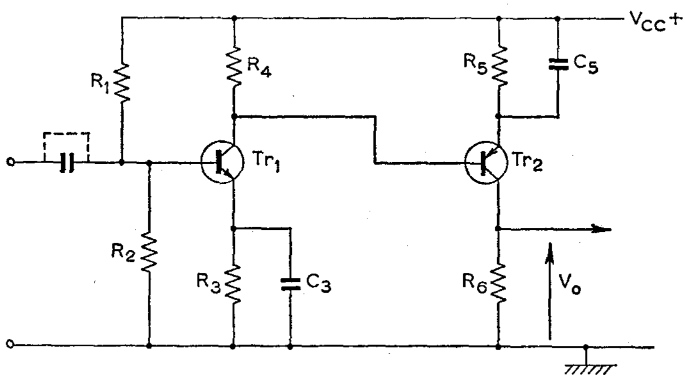


Fig. 21.14 USE OF COMPLEMENTARY SYMMETRY IN TWO-STAGE DIRECT-COUPLED AMPLIFIER

frequency characteristics, since it eliminates the high-pass *CR* coupling circuit. A typical circuit is shown in Fig. 21.14, where the first transistor is of the *n–p–n* type and the second is a *p–n–p*. The circuit is designed from the fact that the quiescent voltage at the emitter of Tr_2 combined with the emitter resistance, R_5, gives the operating current of Tr_2. The collector potential of Tr_1 must then be the emitter voltage of Tr_2 plus the base–emitter voltage of Tr_2.

For a chosen value of quiescent current in Tr_1 this then gives the required value of R_4, remembering that the quiescent base current of Tr_2 flows towards Tr_1.

PROBLEMS

21.1 A common-emitter amplifier uses a transistor which has the following common-base T-parameters: $r_b = 700\ \Omega$; $r_e = 17\ \Omega$; $r_c = 900\ k\Omega$; $\alpha = 0{\cdot}97$. The internal resistance of the signal source is 30 kΩ, the collector load resistance is 2·2 kΩ and the input resistance of the next stage is 1 kΩ. Calculate, from first principles and at mid-frequencies (1) the current gain, (2) the amplifier input resistance, (3) the voltage gain, (4) the power output for an input of 0·5 V r.m.s. Neglect the effect of bias circuit resistors on a.c. operation. (*Hint.* Draw the C.E. circuit, then replace the transistor by its equivalent-T circuit (common base) and solve). (H.N.D.)

Ans. −31·5; 1250 Ω; −17·4; 0·18 mW.

21.2 From common-base tests on a transistor the following results were obtained: (i) with I_E constant, an increase of 1·0 V in collector–base voltage caused an increase in emitter–base voltage of 1 mV and an increase in collector current of 2 μA; (ii) with collector–base voltage constant an increase of 75 mV in emitter–base voltage produced increases in base and collector currents of 50 μA and 5 mA respectively.

Determine the parameters of an equivalent T circuit for this transistor. Give two limitations in the use of this circuit.

Ans. $r_c = 0{\cdot}5\ M\Omega$; $r_b = 500\ \Omega$; $r_e = 9{\cdot}9\ \Omega$; $\alpha = 0{\cdot}99$.

21.3 A small-signal single-stage C.E. amplifier employs a transistor for which $h_{ie} = 1{\cdot}5\ k\Omega$, $h_{fe} = 80$, $h_{oe} = 20\ \mu S$. The collector load resistance is 4·7 kΩ. Determine the mid-frequency current gain of the stage. If the input source has an e.m.f. of 50 mV r.m.s. determine the output signal voltage if the internal source resistance is (*a*) 600 Ω, (*b*) 70 Ω. Assume that the base-bias circuit is equivalent to a resistance of 5 kΩ at the input terminals, and neglect h_{re}.

Ans. −56; 7·57 V; 10·8 V.

21.4 Repeat the calculations of Problem 21.3 for the 600 Ω source resistance, but include the feedback factor $h_{re} = 4 \times 10^{-4}$.

Ans. −49; 8·1 V.

21.5 A single-stage transistor C.E. amplifier employs a transistor for which $h_{feo} = 100$ and $f_1 = 0{\cdot}5$ GHz. Determine the frequency at which the current gain falls 3 dB below the mid-frequency value.

What will be the fall in current gain relative to the mid-frequency value at a frequency of (*a*) 10 MHz, (*b*) 20 MHz, (*c*) 40 MHz?

Ans. 5 MHz; 7·1 dB; 12·3 dB; 18·1 dB.

21.6 A single-stage transistor C.E. amplifier has a current gain which is 3 dB down on its mid-frequency value of 80 when the frequency is 12 MHz. Determine the approximate value of the base transit time and the gain-bandwidth product of the transistor.

Ans. 0·17 ns; 960 MHz.

21.7 The *Q*-factor of the tuned circuit of a tuned transistor C.E. amplifier is 80, the resonant frequency is 470 kHz, and the resonant impedance (dynamic resistance) is 100 kΩ. Determine (*a*) the bandwidth of the simple amplifier, (*b*) the coil tapping required for a bandwidth of 10 kHz, and (*c*) the tuning capacitance required. The transistor has a value of h_{oe} of 40 μS.

Ans. 29·5 kHz; 0·42; 271 pF.

21.8 Derive an expression for the lower cut-off frequency of a transistor *RC*-coupled C.E. amplifier, stating any approximations made.

A transistor for which $h_{fe} = 60$ and $h_{oe} = 50\ \mu$S is used in the first stage of an *RC* coupled C.E. amplifier. The input resistance of the second stage is 1·5 kΩ and the coupling capacitance is 2 μF. The gain falls by 3 dB at 16·8 Hz. Determine (*a*) the collector load resistance, (*b*) the mid-frequency gain, (*c*) the gain reduction and phase shift at 10 Hz. Neglect the effect of input coupling to the first stage. (H.N.D.)

Ans. 3·9 kΩ; -41; $-5{\cdot}9$ dB, 239°.

21.9 Derive the common-base *h*-parameters of a transistor from its common-emitter *h*-parameters assuming h_{re} to be negligible and that $h_{oe}h_{ie} \ll (1 + h_{fe})$.

Two identical transistors are used in the cascode arrangement of Fig. 21.15. The transistor parameters are $h_{ie} = 1$ kΩ; $h_{fe} = 80$; $h_{oe} = 10^{-4}$ S and h_{re} negligible. Calculate the current and power gains for the load resistor R_L at the

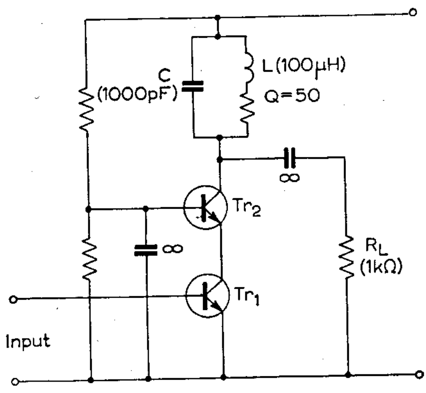


Fig. 21.15

resonant frequency of the tuned collector circuit. (*Hint*. Find the common-base *h*-parameters; then treat the circuit as having a C.E. input transistor followed by a C.B. output transistor).

Ans. $-65{\cdot}7$; 4,310.

21.10 Sketch and explain biasing and coupling arrangements for a two-stage common-emitter *RC*-coupled transistor amplifier.

Such an amplifier has a second stage which employs a transistor with $h_{fe} = 60$, $h_{ie} = 2$kΩ and $h_{oe} = 15\,\mu$S. If the collector load resistance is 12kΩ, determine the voltage gain of the second stage, neglecting bias and coupling components.

Ans. -305.

21.11 If the reverse voltage-feedback ratio in Problem 21.10 is $h_{re} = 2 \times 10^{-4}$, draw the exact equivalent circuit for the second stage and solve for the voltage gain.

Ans. -288.

21.12 Design a single-stage transistor amplifier using an *n–p–n* transistor for which $h_{ie} = 2\text{k}\Omega$; $h_{fe} = 60$; $h_{oe} = 25\mu\text{S}$. The collector circuit is to take 1 mA from a 10 V supply. Find the voltage gain. If the stage feeds a second, similar stage, to what value does the voltage gain fall? Assume a base–emitter voltage of 0·7 V, and a collector–emitter voltage equal to the voltage drop across the collector load resistance. (Start by choosing a suitable quiescent emitter voltage.)

Ans. Typical gains, −120; −40, with emitter resistance of 1 kΩ.

21.13 A common-emitter amplifier stage is fed through a *CR* coupling from a similar previous stage. The transistors used have $h_{ie} = 1{\cdot}5\text{k}\Omega\mu\text{S}$, $h_{fe} = 50$ and $h_{oe} = 40\mu\text{S}$. The collector-circuit load resistance is 5 kΩ. Determine the coupling capactance required for a low-frequency cut-off at 20 Hz. Neglect bias circuits and emitter decoupling.

Ans. Minimum value, 1·84 μF.

21.14 A common-emitter amplifier has a collector load resistance of 10 kΩ. Assuming typical values of $h_{feo} = 50$, $h_{oe} = 25\mu\text{S}$, determine the mid-frequency current gain. If the amplifier is to have an upper cut-off frequency of 3 MHz, determine a minimum value for f_1 and hence find the current gain at 4·5 MHz.

Ans. −40; 153 MHz; $22\underline{/124°}$.

21.15 Two common-emitter stages with *CR* coupling use transistors for which $h_{ie} = 2\text{k}\Omega$, $h_{re} = 5 \times 10^{-4}$, $h_{fe} = 60$ and $h_{oe} = 25\mu\text{S}$. The collector resistances are each 5 kΩ, the source impedance is 2·5 kΩ and the load is 2 kΩ. Determine the voltage gain (V_oE) and the current gain (i) neglecting h_{re}, (ii) including h_{re}, where E is the source e.m.f.

Ans. (i) 760; 1,710, (i) 756; 1,700.

21.16 A common-base amplifier is used as a current buffer stage to feed a load of 20 kΩ from a current source of 0·15 mA which has an internal resistance 10 kΩ. The transistor used has $h_{ib} = 35\Omega$, $h_{ob} = 1\mu\text{S}$ and $h_{fb} = -0{\cdot}98$. Determine the current in the 20 kΩ load if it is (*a*) connected directly to the current source, (*b*) used as the collector resistor of the common-base amplifier.

Ans. (*a*) 0·05 mA, (*b*) 0·147 mA.

21.17 Show that the two-transistor circuit of Fig. 21.16 is equivalent to a single transistor of equivalent overall gain $h_{fe} = h_{fe1} + h_{fe2} + h_{fe1} h_{fe2}$ and input

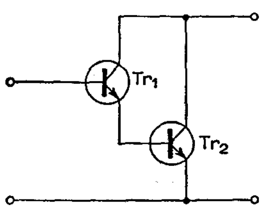


Fig. 21.16

impedance $h_{ie'} = h_{ie1} + (1 + h_{fe1}) h_{ie2}$, where the subscript 1 refers to the first and the subscript 2 refers to the second transistor. (This is known as the *Darlington connexion*).

Chapter 22

FEEDBACK AMPLIFIERS

In the simple amplifiers discussed so far, the input signal current or voltage controls the output, without itself being affected by the output. It has been seen, however, that in the exact *h*-parameter equivalent circuit of a transistor, one of the terms (which has been neglected in the present treatment) represents an effect which the output voltage has on the input circuit. Also the emitter circuit impedance was shown to couple output and input circuits. This effect is called *feedback*, and by the intentional use of feedback the characteristics of amplifiers can be considerably modified.

If the feedback effect reduces the value of the overall gain, it is called *negative feedback*. The advantage of negative feedback is that it enables some of the output characteristics to be stabilized, i.e. to become less dependent on supply voltage and active device parameter changes. Thus negative voltage feedback (where the feedback signal is proportional to the output voltage) will reduce the voltage gain, but make it less liable to vary, and will also increase the amplifier voltage-gain bandwidth. Negative current feedback (where the feedback signal is proportional to the output current) will reduce the current gain, but will stabilize it and will increase the bandwidth. In both cases the distortion and noise will be reduced in the same ratio as the gain, so that the signal-to-noise ratio is unaltered, but it is, in fact, possible to increase this ratio because the active device may not be driven into a non-linear part of its characteristic when negative feedback is applied. In general, the selected gain characteristics become more dependent on circuit components than on active device parameters. Negative feedback will also affect amplifier input and output impedances, as will be seen.

If the feedback effect causes an increase in the overall gain it is called *positive feedback*. In this case the amplifier stability suffers (i.e. the gain becomes more dependent on active device variations and supply voltage changes). If there is enough feedback the amplifier may become completely unstable and act as an oscillator. Often stray coupling (especially at high frequencies) can give unwanted positive feedback and result in undesired oscillation or instability of gain.

22.1 Series Voltage Feedback

Fig. 22.1 represents an amplifier having a constant voltage amplification A_v between the input and the output terminals, i.e.

$$A_v = \frac{V_o}{V'} \tag{22.1}$$

A fraction, β say, of the output voltage V_o is fed back and connected in series with the input voltage (e.g. by using a simple resistance

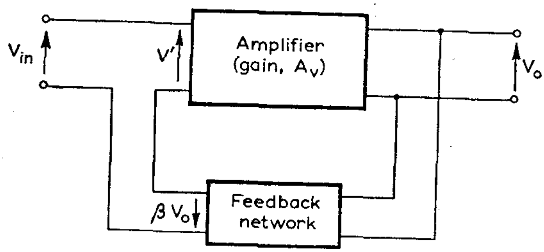


Fig. 22.1 SERIES VOLTAGE FEEDBACK IN GENERAL

voltage-divider). The effective input voltage to the amplifier, V', is then the complexor sum of the input voltage to the circuit, V_{in}, and the voltage which is fed back, βV_o:

$$V' = V_{in} + \beta V_o \tag{22.2}$$

Note that β can be a complex number if the feedback path contains reactive elements.

Substituting for V' in eqn. (22.1) gives

$$V_o = A_v(V_{in} + \beta V_o)$$

so that

$$V_o(1 - \beta A_v) = A_v V_{in}$$

and the overall gain, A_{vf}, with feedback is

$$A_{vf} = \frac{V_o}{V_{in}} = \frac{A_v}{1 - \beta A_v} \tag{22.3}$$

If the magnitude of the denominator on the right-hand side of this equation is greater than unity the amplification is reduced and the feedback effect is negative, or degenerative. If the magnitude of the denominator is less than unity the amplification is increased and the feedback is positive, or regenerative.

Positive feedback apparently gives an attractive method of increasing the gain of an amplifier. It is seldom used in such a role because it leads to instability of amplification.

Note that if the feedback is taken over any odd number of common-emitter stages and is fed back in series aiding with the input then the overall gain at mid-frequencies, A_v, will be negative and β will be positive. The denominator of eqn. (22.3) will then be greater than unity, and the feedback will be negative. A similar result is obtained by taking the feedback over an even number of stages but feeding it in series opposition to the input. In this case A_v is positive but β is negative, again giving overall negative feedback.

Also if A_v is large so that $\beta A_v \gg 1$ then eqn. (22.3) reduces to

$$A_{vf} \approx -\frac{1}{\beta} \tag{22.3a}$$

EXAMPLE 22.1 A 3-stage C.E. amplifier has a mid-frequency voltage gain, A_v of -10^5. If 2 per cent negative series voltage feedback is employed, determine the gain with feedback, A_{vf}. Also find the maximum percentage change in A_v so that A_{vf} will not decrease by more than 2 per cent.

From eqn. (22.3),

$$A_{vf} = \frac{-10^5}{1 + 0{\cdot}02 \times 10^5} \approx \frac{-1}{0{\cdot}02} = \underline{\underline{-50}}$$

Let A_v' be the value of A_v that results in a 2 per cent change in A_{vf}. Then

$$-0{\cdot}98 \times 50 = \frac{A_v'}{1 - \beta A_v'} \qquad \text{or} \qquad -49 + 0{\cdot}98 A_v' = A_v'$$

whence

$$A_v' = -\frac{49}{0{\cdot}02} = -\,2{,}450$$

so that the change in A_v is

$$\frac{10^5 - 2{,}450}{10^5} \times 100 = \underline{\underline{97{\cdot}5 \text{ per cent}}}$$

22.2 Shunt Voltage Feedback

In shunt voltage feedback the feedback signal is applied in parallel with the input signal, and is proportional to the output voltage. The general circuit fed from a constant-voltage source is shown in

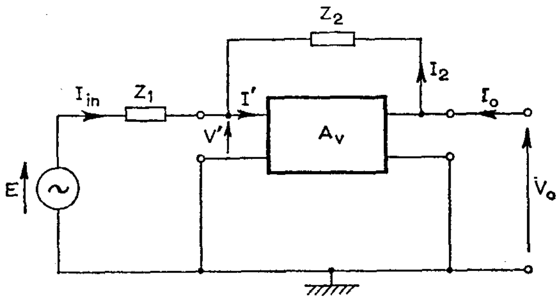


Fig. 22.2 SHUNT VOLTAGE FEEDBACK IN GENERAL

Fig. 22.2. In this circuit, feedback is achieved by adding the impedances Z_1 and Z_2. Then by Millman's theorem,

$$V' = \frac{E/Z_1 + V_o/Z_2}{1/Z_1 + 1/Z_2 + 1/R_{in}} = \frac{V_o}{A_v}$$

where R_{in} is the input resistance of the amplifier without feedback. Hence

$$\frac{E}{Z_1} + \frac{V_o}{Z_2} = \frac{V_o}{A_v}\frac{1}{Z_p}$$

where

$$Z_p = \frac{1}{\dfrac{1}{R_{in}} + \dfrac{1}{Z_1} + \dfrac{1}{Z_2}}$$

It follows that

$$\frac{V_o}{E} = \frac{A_v Z_p}{Z_1\left(1 - \dfrac{A_v Z_p}{Z_2}\right)}$$

Hence the overall voltage gain is

$$A_{vf} = \frac{V_o}{E} = \frac{\alpha A_v}{1 - \beta A_v} \tag{22.4}$$

where $\alpha = Z_p/Z_1$ and $\beta = Z_p/Z_2$.

This equation has the same form as eqn. (22.3). In the above Z_1 includes the source impedance. The derivation for a source of high impedance will be delayed until Section 22.7.

The current gain without feedback is I_o/I', and with feedback is I_o/I_{in}. It is shown in Section 22.7 that $I_{in} \gg I'$, so that shunt voltage feedback reduces the current gain.

Negative shunt voltage feedback is achieved automatically in a single-stage C.E. transistor amplifier when Z_1 and Z_2 are added, owing to the inherent 180° phase shift through the amplifier. It is also obtained if feedback is taken over any *odd* number of stages.

This type of amplifier is often referred to as a *virtual earth amplifier*, since if the gain of the amplifier is high, then for any normal output the input will be very small. The base of the first transistor must therefore be only a fraction of a volt above or below earth potential, and constitutes a *virtual earth* point (see Section 22.10).

22.3 Effect of Negative Feedback on Stability

If A_v (the gain of an amplifier without feedback) changes as the result of supply voltage changes or variation in the active device parameters then A_{vf} will also change, but by a much smaller fraction, i.e. the amplifier will have become more stable (see Example 22.1).

For example, consider an amplifier with shunt voltage feedback in which Z_2 and R_{in} are both very much greater than Z_1; then if $A_v = -100$, $\beta = 0{\cdot}1$, $\alpha = 1$, gives eqn. (22.4).

$$A_{vf} = \frac{-100}{1+10} = -9{\cdot}09$$

If A_v now falls to -90, the new value of overall gain is

$$A_{vf}' = \frac{-90}{1+9} = -9{\cdot}00$$

This shows that in this case a 10 per cent change in A_v gives less than 1 per cent change in A_{vf}. The increase in stability is achieved at the expense of the corresponding reduction in gain. If the term βA_v is large, eqn. (22.4) becomes

$$A_{vf} \approx \frac{\alpha A_v}{-\beta A_v} = -\frac{\alpha}{\beta} \tag{22.5}$$

i.e. the voltage gain is independent of the active device parameters and the supply voltage provided only that the gain is large enough.

EXAMPLE 22.2 For the circuit shown in Fig. 22.3, determine the voltage gain V_o/V_1 and the quiescent d.c. potentials of the base, emitter and collector terminals.

The transistor is a silicon *n–p–n* type for which $h_{ie} = 5\,\text{k}\Omega$, $h_{oe} = 40\,\mu\text{S}$

and $h_{fe} = 90$. The base-emitter voltage is 0·7V and $R_g = 50\Omega$; $R_1 = 5\text{k}\Omega$; $R_f = 250\text{k}\Omega$; $R_L = 7\text{k}\Omega$; and $R_E = 1{\cdot}5\text{k}\Omega$. If the production spread of h_{fe} is 90 — 100 determine the expected variation in gain.

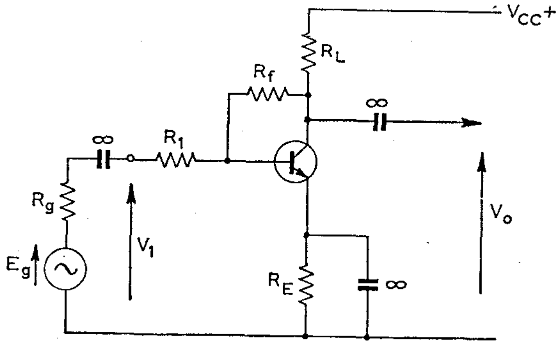


Fig. 22.3

In order to determine the d.c. operating potentials it may be assumed that the d.c. current gain, h_{FE}, is the same as the small signal value, h_{fe}, i.e. 90. Note that the feedback resistor, R_f, provides stabilization of the d.c. operating voltages, since its shunt feedback effect is operative on direct as well as alternating current. The direct currents through R_L and R_E are practically equal, so that, if this current is I,

$$V_{CE} = V_{CC} - IR_L - IR_E$$
$$V_{CB} = V_{CE} - 0{\cdot}7$$

and

$$I_B = \frac{I}{1 + h_{fe}} = \frac{V_{CB}}{R_f}$$

Hence

$$\frac{IR_f}{1 + h_{fe}} = V_{CC} - IR_L - I_{RE} - 0{\cdot}7$$

so that

$$I\left(\frac{R_f}{1 + h_{fe}} + R_L + R_E\right) = 9{\cdot}3$$

and

$$I = 0{\cdot}825\text{mA}$$

The various potentials can now be found.

Emitter potential $= 0{\cdot}825 \times 1{\cdot}5 \approx \underline{\underline{1{\cdot}2\text{V}}}$

Base potential $= 1{\cdot}2 + 0{\cdot}7 \approx \underline{\underline{1{\cdot}9\text{V}}}$

Collector potential $= 10 - (0{\cdot}825 \times 7{\cdot}0) \approx \underline{\underline{4{\cdot}2\text{V}}}$

The small-signal voltage gain may be found by finding the gain without feedback from eqn. (21.13) and substituting this in eqn. (22.4). This is relatively straightforward when the signal source is of low impedance as in this case. Thus, neglecting R_g, the gain without feedback is

$$A_v = \frac{-h_{fe}R_L}{(1 + h_{oe}R_L)h_{ie}} = -1{\cdot}1h_{fe}$$

Also $Z_p = h_{ie}\|R_f\|R_1 = 2{\cdot}5\Omega$; $\alpha = Z_p/R_1 = 0{\cdot}5$; and $\beta = Z_p/Z_f = 0{\cdot}01$. The gain with feedback is therefore

$$A_{vf} = \frac{-0{\cdot}5 \times 1{\cdot}1h_{fe}}{1 + (0{\cdot}01 \times 1{\cdot}1h_{fe})}$$

For $h_{fe} = 90$, $A_{vf} = \underline{\underline{-25}}$, and for $h_{fe} = 100$, $A_{vf} = \underline{\underline{-26}}$.

In this case the gain of the single stage without feedback is fairly low, and the variation of the transistor parameters still has a significant (though smaller) effect on gain with feedback. Frequently the required characteristic is obtained by using a three-stage amplifier with overall feedback from the last to the first stage.

22.4 Effect of Negative Feedback on Bandwidth

The expression for the gain of a single-stage C.E. transistor amplifier is obtained from eqn. (21.39) as

$$A_v = \frac{A_iR_L}{R_{in}} = \frac{-h_{feo}R_L}{(1 + h_{oe}R_L)(1 + jf/f_\beta)R_{in}} = \frac{A_{vo}}{1 + jf/f_\beta}$$

If negative feedback is applied, with feedback factor β, then the general expression for the gain with feedback is

$$A_{vf} = \frac{A_v}{1 - \beta A_v}$$

which at mid-frequencies becomes

$$A_{vf} = \frac{A_{vo}}{1 - \beta A_{vo}}$$

At a high frequency, f, the overall gain is

$$A_{vf}' = \frac{A_{vo}/(1 + jf/f_\beta)}{1 - \beta A_{vo}/(1 + jf/f_\beta)}$$

where f_β is the upper 3 dB frequency of the amplifier without feedback. Simplifying,

$$A_{vf}' = \frac{A_{vo}}{1 - \beta A_{vo} + jf/f_\beta} = \frac{A_{vo}}{(1 - \beta A_{vo})\left(1 + j\dfrac{f}{f_\beta(1 - \beta A_{vo})}\right)} \qquad (22.6)$$

This is 3 dB down on the mid-frequency value when

$$\frac{f}{f_\beta(1 - \beta A_{vo})} = 1$$

i.e. when

$$f = f_\beta(1 - \beta A_{vo}) \tag{22.7}$$

where the product βA_{vo} is negative with negative feedback, so that $f > f_\beta$.

This shows that the "voltage-gain" bandwidth of the amplifier (i.e. the bandwidth between 3 dB points on the voltage-gain/frequency characteristic) has been increased by a factor $(1 + \beta A_{vo})$, again at the expense of a corresponding reduction in voltage gain. If the feedback reduces the current gain, similar considerations show that "current-gain" bandwidth is increased in the same proportion. Note that the gain-bandwidth product remains constant for a single-stage amplifier since bandwidth is increased by the same factor as the gain is reduced by the feedback. This simple relationship does not apply to multi-stage amplifiers.

EXAMPLE 22.3 A single-stage amplifier has a voltage gain of $-100 + j0$ at 1 kHz and $-10 + j30$ at 50 kHz. Find the upper 3 dB cut-off frequency. If 20 per cent negative series voltage feedback is now introduced determine the gain at (i) 1 kHz, and (ii) 60 kHz.

Without feedback the gain at frequency f is

$$A_v = \frac{-100}{1 + jf/f_\beta}$$

Hence at 50 kHz,

$$-10 + j30 = \frac{-100}{(1 + j50 \times 10^3/f_\beta)}$$

so that

$$(-10 + j30)(1 + j50 \times 10^3/f_\beta) = -100$$

i.e.

$$-10 - 15 \times 10^5/f_\beta + j30 - j50 \times 10^4/f_\beta = -100$$

Equating reference terms,

$$\frac{15 \times 10^5}{f_\beta} = 90$$

so that $f_\beta = \underline{\underline{16{\cdot}7\,\text{kHz.}}}$

As a check, the quadrate terms give $f_\beta = 50 \times 10^4/30 = 16{\cdot}7$ kHz.
With 20 per cent negative feedback, $\beta = 0{\cdot}2$ and at 1 kHz the gain is

$$A_{vf} = \frac{-100}{1 + 20} = \underline{\underline{-4{\cdot}8}}$$

At 60 kHz the gain is, from eqn. (22.6),

$$A_{vf}' = \frac{-4\cdot8}{1 + j\dfrac{60}{16\cdot7} \times \dfrac{1}{21}} = \frac{-4\cdot8}{1 + j0\cdot17} = \underline{\underline{4\cdot7\angle 170^\circ}}$$

The upper 3 dB frequency limit with feedback is, of course, $f_B(1 + 20) =$ 350 kHz.

22.5 Effect of Negative Feedback on Distortion and Noise

Distortion occurs in amplifiers when the signal is so large that the active device operates outside the linear range of its characteristics. The output then contains components which are not present at the input. These components are harmonics of the signal frequencies. Generally the last stage of an amplifier is the one in which most harmonic distortion occurs.

Noise also will occur in amplifiers, where "noise" is taken to mean unwanted voltages appearing at the output. These noise voltages may be due to (*a*) poor smoothing in the rectifier supply, (*b*) 50 Hz or 100 Hz hum pick-up, (*c*) mechanical vibrations of valve electrodes (*microphony*), (*d*) drift in transistors due to changes in temperature, (*e*) thermal noise in resistors due to molecular vibrations, (*f*) *shot noise* in the active device (due to the fact that the charge carriers are emitted randomly). Generally the first stage of an amplifier determines its noise performance, since it is in this stage that the signal is smallest, and hence the signal/noise ratio is also smallest. The size of the input signal which can be usefully amplified is determined largely by the noise in the first stage, since for any useful gain it must be possible to detect the signal in the presence of the noise. Hence generally the minimum input signal must be greater than the noise. When low signal levels are to be amplified special low-noise input circuits must be employed.

An important amplifier criterion is the ratio of signal/noise power at the input to that at the output. This is called the *noise factor*, F (in decibels it is $10 \log_{10} F$). Thus for a signal power S_i at the input and a power gain G in the amplifier the output signal power is GS_i. The input noise power N_i is also amplified and added to by the internal noise so that the output noise power, N_o, will be greater than GN_i. Then

$$F = \frac{S_i/N_i}{GS_i/N_o} = \frac{N_o}{GN_i} \tag{22.8}$$

Noise and distortion generated within an amplifier can be represented by a constant-voltage generator having an e.m.f. e_n, in series with the output as shown in the series-feedback circuit of Fig. 22.4. This e.m.f. can be assumed constant if the output signal is maintained at the level which it had without feedback (i.e. by increasing the input voltage).

Without feedback the harmonic or noise voltage output is e_n. Let this become me_n when feedback is present, where m is a fraction.

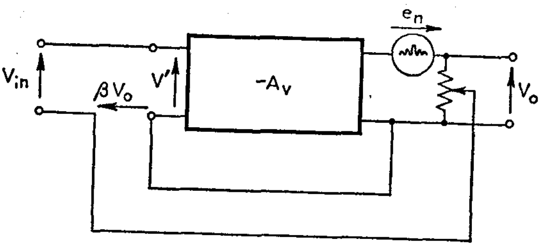


Fig. 22.4 FEEDBACK OF DISTORTION AND NOISE

Then the fraction of harmonic or noise voltage which is fed back is βme_n. This is amplified in the amplifier (of gain $-A_v$) to give an output harmonic or noise voltage of $-A_v\beta me_n$, so that

$$me_n = e_n - A_v\beta me_n$$

Hence

$$m = \frac{1}{1 + \beta A_v} \tag{22.9}$$

where negative feedback has been assumed. This shows that the noise or distortion is reduced in the same ratio as the gain. Note that adding feedback alone does not improve the signal/noise ratio at the amplifier output. Indeed, since noise and harmonic distortion voltages may well cover a wider bandwidth than the signal, they may be present at the lower and upper frequency ranges where the phase shift of a multi-stage amplifier is such that the feedback becomes positive. In this case the noise may be increased by the feedback, and careful design of the frequency characteristics of the feedback path is essential to prevent this. Note that the same arguments apply to drift in d.c. amplifiers—i.e. the percentage drift is not reduced by negative feedback.

Generally feedback in the output stage of an amplifier can be used to reduce distortion, because extra stages are inserted before the feedback stage in order to restore the overall gain. Since these are low-signal-level stages they will hardly contribute any distortion themselves, and hence the overall distortion will be reduced, for the same output power.

Consideration of the shunt voltage feedback circuit gives the same results as those derived above.

22.6 Current Feedback

SERIES CURRENT FEEDBACK

The simplest way to obtain series current feedback (in which the feedback signal is proportional to the load current) is to have an unbypassed resistor in the emitter or cathode circuit. Then the

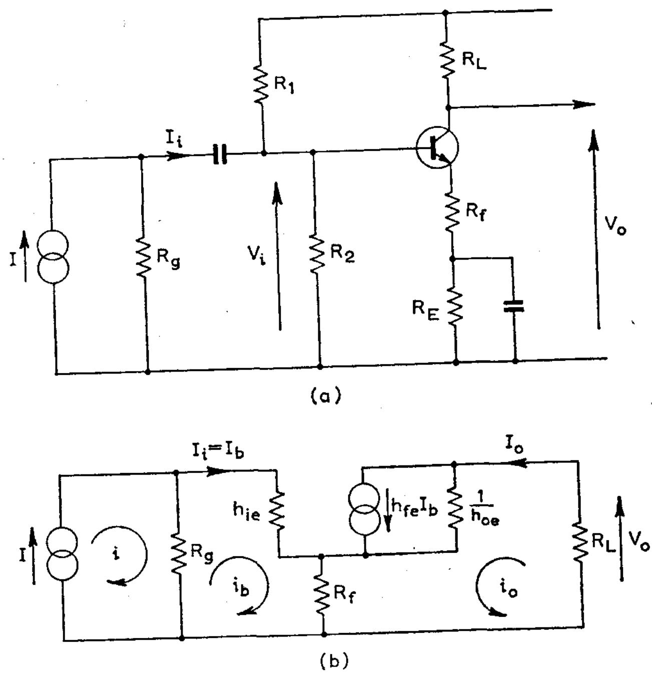


Fig. 22.5 SERIES CURRENT FEEDBACK

output current flowing through this resistor gives rise to a voltage (proportional to current) in series with the input. A typical transistor circuit with series current feedback is shown in Fig. 22.5(*a*).

It will be seen that the voltage drop across the unbypassed portion of the emitter-circuit resistor is common to both the input and output circuits, and hence causes feedback which an inspection of relative current directions shows is negative. Neglecting the effect of $R_p\,(=R_1\|R_2)$ for simplicity, the small-signal equivalent circuit becomes that shown at (*b*). Using the mesh equations (with small lettters to represent the complexor mesh currents),

$$\text{Input mesh} \qquad 0 = (R_g + h_{ie} + R_f)i_b - R_g i + R_f i_o \tag{22.10}$$

$$\text{Output mesh} \quad 0 = \left(R_L + \frac{1}{h_{oe}} + R_f\right) i_o + R_f i_b - \frac{h_{fe} i_b}{h_{oe}} \tag{22.11}$$

Since $i_b \ll i_o$, eqn. (22.11) can be simplified by neglecting $R_f i_b$ which is small compared to $R_f i_o$, so that

$$\frac{h_{fe} i_b}{h_{oe}} \approx \left(R_f + R_L + \frac{1}{h_{oe}}\right) i_o$$

For the mesh currents shown, the conductor currents I_b and I_o are equal to the corresponding mesh currents i_b and i_o, so that the internal current gain with feedback is

$$A_{if} = \frac{V_o/R_L}{I_b} = \frac{-I_o}{I_b} = \frac{-h_{fe}}{1 + h_{oe}(R_L + R_f)} \tag{22.12}$$

If $R_f \ll R_L$, the current gain is thus hardly affected by the feedback (without feedback $A_i = -h_{fe}/(1 + h_{oe}R_L)$).

The voltage gain with feedback is given by

$$A_{vf} = \frac{V_o}{V_i} = \frac{-I_o R_L}{I_b(h_{ie} + R_f) + I_o R_f} \approx \frac{-I_o R_L}{I_b h_{ie} + I_o R_f} \quad (\text{since } I_b \ll I_o)$$

$$= \frac{-I_o R_L}{I_b h_{ie} + \dfrac{I_b h_{fe} R_f}{1 + h_{oe}(R_L + R_f)}}$$

$$= \frac{-I_o R_L}{I_b h_{ie}\left\{1 + \dfrac{h_{fe} R_f}{h_{ie}(1 + h_{oe}(R_L + R_f))}\right\}} = \frac{A_v}{1 - \beta A_v} \tag{22.13}$$

where

$$\beta = \frac{R_f}{R_L} \frac{1 + h_{oe} R_L}{1 + h_{oe}(R_L + R_f)}$$

$$\approx \frac{R_f}{R_L} \qquad \text{and} \qquad A_v = \frac{-h_{fe} R_L}{h_{ie}(1 + h_{oe} R_L)}$$

is the voltage gain without feedback. Since eqn. (22.13) has the same form as eqn. (22.5), the same considerations regarding stability and bandwidth will apply.

SHUNT CURRENT FEEDBACK

With negative shunt current feedback the signal that is fed back to the input is proportional to the output current, and is fed back in parallel with the input. One possible circuit is shown in Fig. 22.6(*a*), where the effective a.c. circuit elements only are shown.

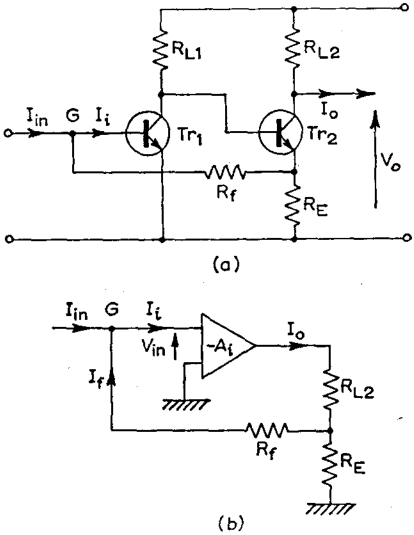


Fig. 22.6 SHUNT CURRENT FEEDBACK

In this case it may be shown that the current gain I_o/I_{in} is reduced by a factor of $(1 - \beta A_1)$, where $\beta = R_E/R_f$ and A_1 is the magnitude of the current gain over the two stages without feedback. The input impedance is reduced by the same factor, and the output impedance is increased.

For the general circuit shown at (*b*) the current gain without feedback is the ratio of output current I_o to input current I_i, i.e.

$$A_i = \frac{I_o}{I_i} = \frac{I_o}{I_{in}} \qquad \text{where } I_i = I_{in} = \text{source current}$$

When shunt current feedback is applied the current gain (assuming the same source current, I_{in}) becomes

$$A_{if} = \frac{I_{of}}{I_{in}} \qquad \text{where } I_{of} \text{ is the output current with feedback}$$

In this case the input current to the first transistor is

$$I_{if} = I_{in} + I_f = \frac{I_{of}}{A_i}$$

Since G is a virtual earth,

$$I_f \approx \frac{I_{of}R_E}{R_E + R_f} = \frac{I_{of}R_E}{R_f} \approx \beta I_{of}$$

provided that $R_f \gg R_E$ and where $\beta = R_E/R_f$. Hence

$$I_{in} + \beta I_{of} = \frac{I_{of}}{A_i}$$

so that

$$A_{if} = \frac{I_{of}}{I_{in}} = \frac{A_i}{1 - \beta A_i} \tag{22.14}$$

In the absence of feedback the input impedance of the circuit is

$$Z_{in} = \frac{V_{in}}{I_{in}} = \frac{V_{in}}{I_i} = \frac{A_i V_{in}}{I_o}$$

Assume now that feedback is applied and that the input current is increased until the output current has the same value as without feedback. The input current to the first transistor will be $I_i = I_o/A_i$ and the input voltage will be V_{in}—the same values as without feedback. The source current I_{in} must now be changed to

$$I_{in\,f} = I_i - I_f = I_i - \beta I_o = I_i(1 - \beta A_i)$$

so that the input impedance with feedback becomes

$$Z_{in\,f} = \frac{V_{in}}{I_{in\,f}} = \frac{V_{in}}{I_i(1 - \beta A_i)} = \frac{Z_{in}}{(1 - \beta A_i)} \tag{22.15}$$

The input impedance is thus changed in the same ratio as the gain. For negative feedback the input impedance is *reduced* in the same ratio as the gain. Note that for positive feedback the input impedance will be *increased* and may become negative after passing through the value infinity.

22.7 Effect of Feedback on Impedance Levels

It has been seen in the last section that negative shunt current feedback reduces the input impedance to an amplifier circuit. In general, feedback can affect both the input and output impedances of amplifiers. The way in which the feedback is introduced into the input circuit will affect the input impedance (i.e. the ratio of input voltage to input current). Thus, for series negative feedback, the input voltage required for a given input current must be increased—i.e. the input impedance is increased. For shunt negative feedback the input current can divide between the amplifier and the feedback path, and hence input impedance is reduced. Both of these effects are reversed if the feedback is positive, and indeed in this case the input resistance can become negative.

The effect of feedback on output impedance depends upon whether the feedback signal is proportional to output voltage or output current. In general voltage feedback reduces the output impedance while current feedback increases it.

Output impedance can be obtained by inspection of the voltage gain expression for an amplifier. Thus if the voltage gain is written in the form

$$A_v = \frac{PZ_L}{Q + Z_L} \tag{22.16}$$

where Z_L is the load, then P is the intrinsic voltage gain as $Z_L \to \infty$, and Q is the output impedance. The Thévenin equivalent output circuit is as shown in Fig. 22.7(*a*). The Norton equivalent circuit

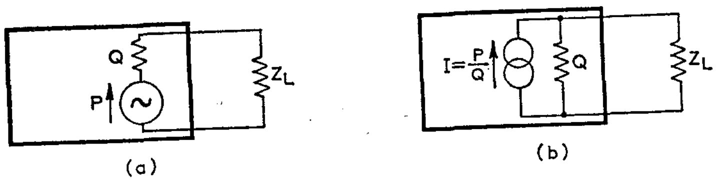


Fig. 22.7 EQUIVALENT OUTPUT CIRCUITS OF AN AMPLIFIER

is shown at (*b*). The output impedance seen by any other load which may be connected across Z_L is then simply Q in parallel with Z_L.

An alternative method of finding the output impedance is to apply a constant-voltage or a constant-current generator to the output terminals with the input signal source represented by its internal impedance. The output impedance is then the ratio of voltage at the output to the current flowing into the output terminals.

SERIES VOLTAGE FEEDBACK

In the series voltage feedback circuit of Fig. 22.8(*a*), let the amplifier (which will usually be multistage) have an intrinsic gain A_v and an output impedance of Z_o without feedback. A constant-voltage source V is applied across the output terminals, and the input

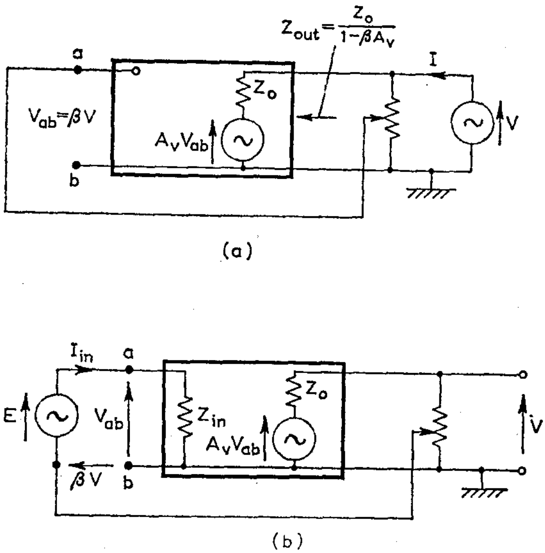


Fig. 22.8 IMPEDANCE LEVELS FOR SERIES VOLTAGE FEEDBACK

terminals are short-circuited. A fraction, β, of the voltage V at the output is fed back to the input. Then the voltage V_{ab} at the input is βV, and neglecting the impedance of the feedback network (which must be purely resistive for pure voltage feedback), the current, I, from the source is

$$I = \frac{V - A_v V_{ab}}{Z_o} = \frac{V - \beta V A_v}{Z_o}$$

Hence the output impedance is

$$Z_{out} = \frac{V}{I} = \frac{Z_o}{1 - \beta A_v} \tag{22.17}$$

If the feedback is negative the output impedance is smaller when feedback is applied than without feedback. Note that if A_v is positive the feedback is still negative if β is negative—this normally requires a transformer coupling.

The input impedance with feedback is obtained by considering the circuit at (*b*). A signal is applied to the input, and the output is open-circuited. Then neglecting the impedance of the feedback network, the feedback voltage is $\beta V = \beta V_{ab} A_v$ and the new input impedance is

$$Z_{in\,f} = \frac{E}{I_{in}} = \frac{V_{ab} - \beta V}{I_{in}} = \frac{V_{ab}(1 - \beta A_v)}{I_{in}}$$
$$= Z_{in}(1 - \beta A_v) \tag{22.18}$$

where Z_{in} is the input impedance without feedback. For negative feedback βA_v is negative, and the input impedance is therefore increased.

EXAMPLE 22.4 A transistor amplifier uses a transistor for which $h_{ie} = 1\,\text{k}\Omega$ in the first stage, which is in common-emitter connexion. The overall current gain without feedback is $-15{,}000$. If there is 0·05 p.u. negative series voltage feedback, determine the input impedance of the feedback amplifier if the load resistance is 75 Ω. If the output impedance without feedback is 20 kΩ, determine the ouput impedance when feedback is applied.

The input resistance, R_{in}, without feedback will be almost equal to h_{ie} (neglecting the d.c. bias resistors).

The overall voltage gain without feedback is

$$A_v = \frac{A_i \times R_L}{R_{in}} = \frac{-15{,}000 \times 75}{1{,}000} = -1{,}125 \quad (\text{since } R_{in} \approx h_{ie})$$

Hence

$$Z_{in\,f} = R_{in}(1 - \beta A_v) \quad (\text{eqn. } (22.18))$$
$$= 1{,}000(1 + 56{\cdot}3) = \underline{\underline{57{\cdot}3\,\text{k}\Omega}}$$

From eqn. (22.17) the output impedance with feedback applied is

$$Z_{out\,f} = \frac{Z_o}{1 - \beta A_v} = \frac{20{,}000}{57{\cdot}3} = \underline{\underline{350\,\Omega}}$$

SHUNT VOLTAGE FEEDBACK

The input impedance of the shunt voltage feedback circuit shown in Fig. 22.9 is

$$Z_{in\,f} = \frac{E}{I_{in}} = Z_1 + Z_{GE}$$

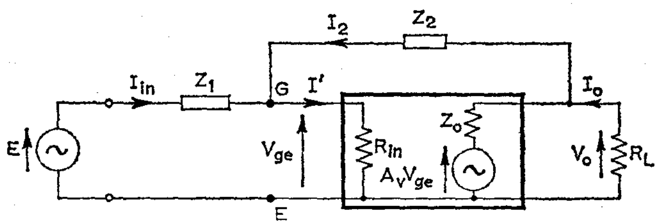


Fig. 22.9 MULTISTAGE AMPLIFIER WITH SHUNT VOLTAGE FEEDBACK

where $Z_{GE} = V_{GE}/(I' - I_2)$. The admittance looking into the amplifier at G is

$$Y_{GE} = \frac{I' - I_2}{V_{GE}} = \frac{1}{R_{in}} - \frac{I_2}{V_{GE}}$$
$$= \frac{1}{R_{in}} - \frac{(A_v V_{GE} - V_{GE})}{Z_2 V_{GE}}$$
$$= \frac{1}{R_{in}} + \frac{1 - A_v}{Z_2} \tag{22.19}$$

where $A_v = V_o/V_{GE}$.

This represents the admittance of a circuit consisting of a resistance R_{in} in parallel with an impedance $Z_2/(1 - A_v)$. Hence the impedance at G is decreased by the negative feedback. The input impedance is thus $Z_1 + 1/Y_{GE}$. Note that if A_v is large, the impedance of the feedback path through Z_2 as seen looking in at point G is approximately $Z_2/(-A_v)$, which will generally be small.

The gain of the amplifier fed from a high-impedance source can now be found. The circuit is that of Fig. 22.9, but Z_1 is now the source impedance. The voltage gain A_v ($= V_o/V_{GE}$) is almost unaltered by the feedback, provided that $I_2 \ll I_o$, since the amplifier input terminals are now G and E.

With feedback applied the input current, I_{in}, divides between R_{in} and the equivalent impedance of Z_2 as seen looking into point G (i.e. $Z_2/(1 - A_v)$, from eqn. 22.19). The current gain is now

$$A_{if} = \frac{I_o}{I_{in}} = \frac{I_o}{I'}\frac{I'}{I_{in}}$$
$$= A_i \frac{Z_2}{(1 - A_v)} \frac{1}{R_{in} + Z_2/(1 - A_v)}$$
$$= \frac{A_i}{\dfrac{R_{in}}{Z_2}(1 - A_v) + 1}$$
$$\approx \frac{A_i}{\beta A_v} \quad \text{(assuming } A_v \gg 1\text{)} \tag{22.20}$$

and where $\beta = R_{in}/Z_2$.

The current gain thus is reduced by the negative feedback.

MILLER EFFECT

It should be noted that shunt voltage feedback may occur unintentionally in single-stage amplifiers owing to the stray capacitance between input and output circuits providing the feedback impedance, Z_2. This stray capacitance is called the *Miller capacitance*. In this case the impedance Z_2 is $1/j\omega C$, where C is the Miller capacitance. The admittance of the feedback circuit seen at the input to the amplifier is, from eqn. (22.19),

$$Y_f = \frac{1 - A_v}{Z_2} = j\omega C(1 - A_v) \tag{22.21}$$

Three special cases are of interest. First, if A_v is a negative reference quantity, then Y_f is a capacitive susceptance of magnitude $\omega C(1 + A_v)$. Second, if A_v has a phase angle between 90° and 180° so that it can be represented as $A_v = -P + jQ$, then

$$Y_f = j\omega C(1 + P - jQ) = \omega CQ + j\omega C(1 + P) \tag{22.22}$$

i.e. it consists of a conductance (frequency dependent), ωCQ, in parallel with a capacitive susceptance, $\omega C(1 + P)$. Third, if A_v has a phase angle between 180° and 270° (e.g. in a single stage with an inductive load such as a tuned circuit operating below resonance), so that it can be represented as $A_v = -P - jQ$, then

$$Y_f = j\omega C(1 + P + jQ) = -\omega CQ + j\omega C(1 + P) \tag{22.23}$$

i.e. there is a negative component of input conductance. If this is larger than the conductance of the source the amplifier will be unstable.

EXAMPLE 22.5 In the transistor amplifier shown in Fig. 22.10(*a*), find expressions for the input admittance and the output impedance. There are stray capacitances C_{cb} between collector and base and C_{be} between base and earth, and the collector load is an inductance $j\omega L$.

The small-signal equivalent circuit is shown at (*b*) neglecting R_4 and R_2 and assuming that the emitter resistor, R_3, is adequately decoupled. From this, the current gain without feedback is

$$A_i = \frac{-h_{fe}}{1 + h_{oe}Z_L} \quad \text{where } Z_L = j\omega L$$

Hence the voltage gain without feedback is

$$A_v = \frac{-h_{fe}Z_L}{h_{ie}(1 + h_{oe}Z_L)} \approx \frac{-h_{fe}j\omega L}{h_{ie}}$$

so that the input admittance with feedback is (from eqn. (22.21))

$$Y_{inf} = Y_{in} + Y_f = j\omega C_{be} + \frac{1}{h_{ie}} + j\omega C_{cb}\left(1 + \frac{jh_{fe}\omega L}{h_{ie}}\right)$$

$$= \frac{-\omega^2 LC_{cb}h_{fe}}{h_{ie}} + \frac{1}{h_{ie}} + j\omega(C_{be} + C_{cb})$$

where $Y_{in} = j\omega C_{be} + 1/h_{ie}$ = input admittance without feedback.

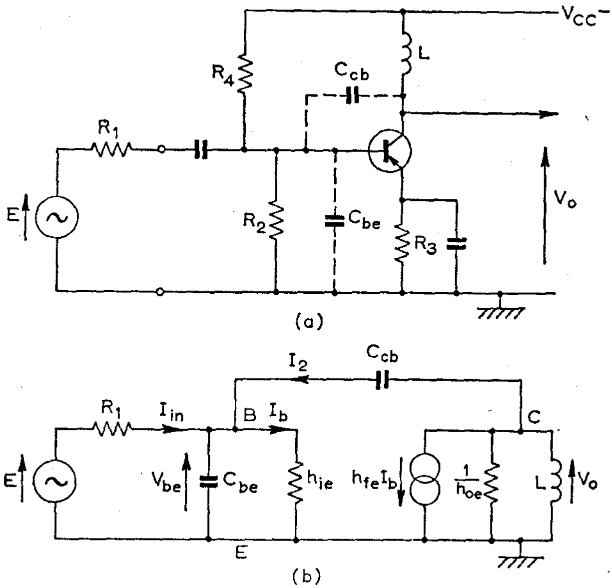


Fig. 22.10 MILLER CAPACITANCE

For a low-impedance source ($R_1 \ll h_{ie}$), the output impedance, obtained by replacing the load L by a constant-voltage source, V, and representing the input generator by its internal impedance, R_1, is very nearly $1/h_{oe}$ in parallel with $1/j\omega C_{cb}$.

For a high-impedance source ($R_1 \gg h_{ie}$) and neglecting C_{be}, the same procedure yields an alternating input current to the transistor of approximately $Vj\omega C_{cb}$ (assuming that $j\omega C_{cb} \gg 1/h_{ie}$). Hence at point C in Fig. 22.10(*b*) the currents are as follows:

$I_2 \approx Vj\omega C_{cb}$

Current in $h_{oe} = Vh_{oe}$

Current in constant-current generator, $h_{fe}I_b \approx h_{fe}I_2$

The sum of these gives the current, I, from the constant-voltage source which replaces L. Hence

$$I = Vj\omega C_{cb} + Vh_{oe} + h_{fe}Vj\omega C_{cb},$$

so that the output impedance with feedback is

$$Z_{out} = \frac{V}{I} \approx \frac{1}{h_{oe} + j\omega C_{cb}(1 + h_{fe})}$$

which is even smaller than in the case of a low value of R_1.

SERIES CURRENT FEEDBACK

With this form of feedback the calculation of output impedance follows closely that used for series voltage feedback, but here a constant-current generator delivering a current I is connected across the output as shown in Fig. 22.11.

The voltage across R_f is IR_f and this is fed back to the input to give an input voltage of $-IR_f$. If the intrinsic voltage gain is A_v, then from the diagram,

$$V = I(Z_o + R_f) + A_v V_1 = I(Z_o + R_f) - IR_f A_v$$

Hence the output impedance with feedback is

$$Z_{out} = \frac{V}{I} = Z_o + (1 - A_v)R_f \tag{22.24}$$

For negative feedback A_v is negative, and hence the output impedance is greater with feedback than it is without it.

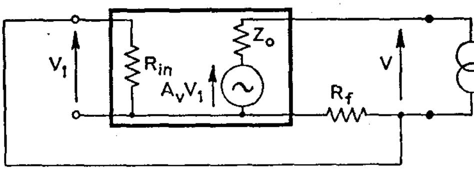


Fig. 22.11 TO FIND OUTPUT IMPEDANCE WITH SERIES CURRENT FEEDBACK

The input impedance of the amplifier can be found as before and will be higher with feedback than without it.

EXAMPLE 22.6 In the transistor series current-feedback amplifier of Fig. 22.5, $R_f = 100\Omega$; $R_L = 2{,}000\Omega$; $h_{ie} = 1{\cdot}5\text{k}\Omega$; $h_{fe} = 80$; and $h_{oe} = 10^{-4}$S. Determine approximately the input impedance, the voltage gain, and the output impedance.

Referring to Fig. 22.5(b), the input impedance with feedback is

$$Z_{inf} = \frac{V_i}{I_b} \approx \frac{I_b h_{ie} + I_0 R_f}{I_b} \quad \text{(assuming } I_b \ll I_o\text{)}$$

$$= \frac{I_b h_{ie} - A_i I_b R_f}{I_b}$$

$$= h_{ie} - A_i R_f$$

where A_i = current gain without feedback $= \dfrac{-h_{fe}}{1 + h_{oe}R_L}$.

Hence

$$Z_{inf} = h_{ie}(1 - \beta A_v)$$

where $\beta = R_f/R_L$ and $A_v = A_i R_L/R_{in} = A_i R_L/h_{ie}$. Note that this expression for impedance is in the same form as that obtained in eqn. (22.18) for series voltage feedback.

Inserting numerical values,

$$Z_{inf} = 1{,}500 + \frac{80 \times 100}{1 + 10^{-4} \times 2 \times 10^3} = \underline{\underline{8{,}200\Omega}}$$

From eqn. (22.13) the voltage gain with feedback is

$$A_{vf} = \frac{A_v}{1 - \beta A_v} = \frac{-h_{fe}R_L/h_{ie}(1 + h_{oe}R_L)}{1 + R_f h_{fe}/h_{ie}(1 + h_{oe}R_L)}$$

$$= \frac{-h_{fe}R_L}{h_{ie}\left(1 + h_{oe}R_L + \dfrac{R_f h_{fe}}{h_{ie}}\right)}$$

$$= \frac{-(h_{fe}/h_{ie}h_{oe})R_L}{\dfrac{1}{h_{oe}} + \dfrac{R_f h_{fe}}{h_{oe}h_{ie}} + R_L} \qquad \text{(i)}$$

Comparing this with eqn. (22.16), it will be seen that the output impedance is

$$Z_{out} = \frac{1}{h_{oe}} + \frac{R_f h_{fe}}{h_{oe}h_{ie}}$$

$$= 10^4 + \frac{10^4 \times 100 \times 80}{1{,}500} = \underline{\underline{63\text{k}\Omega}}$$

This is the same result as would be obtained by the direct application of eqn. (22.24).

Inserting numerical values in (i), the voltage gain with feedback is

$$A_{vf} = \underline{\underline{-16{\cdot}5}}$$

22.8 Compound Feedback

In some instances it may be desired to achieve the increased stability that negative feedback gives together with definite levels of input

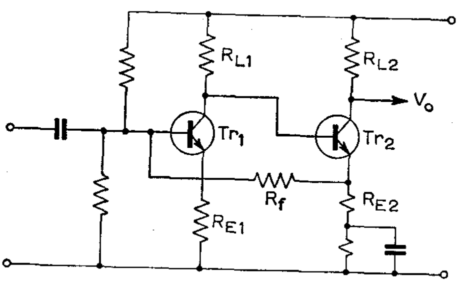


Fig. 22.12 TYPICAL COMPOUND FEEDBACK CIRCUIT

and output impedance. In such cases it may be necessary to use more than one feedback path. This is called *compound feedback.*

A typical circuit is shown in Fig. 22.12. The resistor R_{E1} gives negative series current feedback for the first stage (giving an increase in input impedance of the stage), and R_{E2} and R_f provide negative shunt current feedback over both stages (giving a decrease in input impedance). By suitable choice of components the input and output impedance levels may be made to have designed values.

An alternative arrangement is to remove R_{E2} and connect R_f between the collector of Tr_1 and the emitter of Tr_2. Note, however, that such a connexion no longer provides shunt-applied feedback.

EXAMPLE 22.7 In the circuit of Fig. 22.12 the value of R_{E1} is chosen to give 0·02 p.u. negative feedback in the first stage, and the values of R_f and R_{E2} give 0.01 p.u. negative feedback over the two stages. If h_{ie} for the first transistor is 1·2 kΩ determine the new value of input impedance. The loaded voltage gain of the first stage is $A_{v1} = -100$, and that of the second stage is $A_{v2} = -120$, both without feedback.

The resistor R_{E1} in the first stage gives negative series current feedback. This increases the input impedance of the first stage from $R_{in} \approx h_{ie}$ to

$$Z_{in} \approx h_{ie}(1 - \beta A_{v1})$$

The gain of the first stage with the series current feedback falls to

$$A_{v1f} = \frac{A_{v1}}{1 - \beta_1 A_{v1}} = \frac{-100}{1 + 2} = -33{\cdot}3$$

Hence the overall gain of both stages with feedback in the first stage only is

$$A_v' = A_{v2} \times A_{v1f} = (-120)(-33{\cdot}3) = 4{,}000$$

The new overall input impedance is reduced by the second form of feedback from $h_{ie}(1 - \beta A_{v1})$ to

$$Z_{inf} = \frac{h_{ie}(1 - \beta_1 A_{v1})}{1 - \beta_2 A_v'} = \frac{1{,}200\,(1 + 2)}{1 + 40} = \underline{\underline{88\ \Omega}}$$

22.9 The Emitter Follower (Common Collector)

The *emitter follower* circuit employs 100 per cent negative series voltage feedback, and therefore from Section 22.7(*a*) it will have a high input and a low output impedance. The voltage gain will be shown to be slightly less than unity, with no phase reversal, and the circuit is frequently used as a buffer stage between a high-impedance source and a low-impedance load. The circuit is shown in Fig. 22.13(*a*) and the small-signal a.c. equivalent circuit at (*b*). The circuit is often called the *common-collector circuit*, since the collector terminal is common to input and output circuits (through the low impedance of the d.c. supply).

The d.c. conditions are set by the resistors R_1 and R. Thus for a quiescent collector current I_C ($\approx I_E$) the emitter voltage is $I_C R$ and

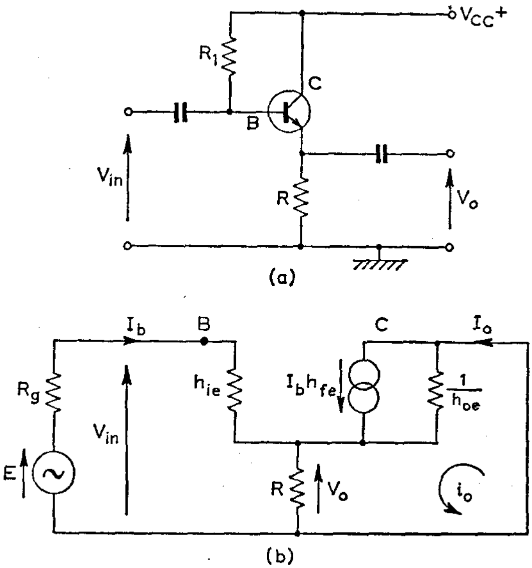


Fig. 22.13 EMITTER FOLLOWER

the base voltage is $I_C R + V_{BE}$, so that for a forward d.c. short-circuit current gain h_{FE} (to give $I_B = I_C/h_{FE}$), the value of R_1 is

$$R_1 = \frac{(V_{CC} - I_C R - V_{BE})h_{FE}}{I_C}$$

For the signal, the output mesh equation can be written

$$0 = \left(\frac{1}{h_{oe}} + R\right) I_o - \frac{h_{fe}}{h_{oe}} I_b \quad \text{(assuming } I_o R \gg I_b R\text{)}$$

so that

$$\frac{I_o}{I_b} = \frac{h_{fe}}{1 + h_{oe}R} = A_i \tag{22.25}$$

The current gain is thus the same as that derived for the common-emitter amplifier using the simplified equivalent circuit.

In the input mesh, the input signal voltage is

$$\begin{aligned} V_{in} &= I_b h_{ie} + I_o R \\ &= I_o \left\{\frac{(1 + h_{oe}R)h_{ie}}{h_{fe}} + R\right\} \\ &= \frac{V_o}{R}\left\{\frac{(1 + h_{oe}R)h_{ie}}{h_{fe}} + R\right\} \end{aligned}$$

Hence the voltage gain is

$$A_v = \frac{V_o}{V_{in}} = \frac{h_{fe}R}{h_{ie} + (h_{fe} + h_{oe}h_{ie})R} \tag{22.26}$$

Note that

$$\frac{V_o}{E} = \frac{h_{fe}R}{h_{fe}R + (1 + h_{oe}R)(h_{ie} + R_g)} \tag{22.27}$$

There is no phase reversal, and A_v is obviously smaller than unity.

The input impedance is readily obtained from the relation

$$\begin{aligned} V_{in} &= I_b h_{ie} + I_o R \\ &= I_b \left(h_{ie} + \frac{h_{fe}R}{1 + h_{oe}R} \right) \end{aligned}$$

so that the input impedance is

$$Z_{in} = h_{ie} + \frac{h_{fe}R}{1 + h_{oe}R} \tag{22.28}$$

The simplest way of determining the output impedance is by inspection of eqn. (22.27) and comparison with eqn. (22.16). Thus the ratio of V_o to E can be written

$$\frac{V_o}{E} = \frac{h_{fe}R}{\{h_{fe} + h_{oe}(h_{ie} + R_g)\}\left\{\dfrac{h_{ie} + R_g}{h_{fe} + h_{oe}(h_{ie} + R_g)} + R\right\}}$$

It follows that the output impedance with feedback is

$$Z_{out} = \frac{h_{ie} + R_g}{h_{fe} + h_{oe}(h_{ie} + R_g)} \tag{22.29}$$

and the output admittance is

$$Y_{out} = h_{oe} + \frac{h_{fe}}{h_{ie} + R_g} \tag{22.29a}$$

For a constant-voltage source ($R_g \to 0$) Y_{out} becomes h_{fe}/h_{ie}, while for a constant-current source ($R_g \to \infty$), $Y_{out} \to h_{oe}$.

The frequency characteristic of the emitter-follower amplifier is obtained by substituting $h_{fe} = h_{feo}/(1 + jf/f_\beta)$ (i.e. from eqn. 21.38) in a simplified form of eqn. (22.26). Thus

$$A_v \approx \frac{h_{fe}R}{h_{ie} + h_{fe}R} \quad \text{(assuming } h_{oe}h_{ie} \ll h_{fe}\text{)}$$

i.e.

$$A_v = \frac{h_{feo}R}{h_{ie}(1 + jf/f_\beta) + h_{feo}R}$$

This is 3 dB down on the mid-frequency gain when

$$\frac{h_{ie}f}{f_\beta} = h_{ie} + h_{feo}R$$

i.e. when

$$f = f_\beta\left(1 + \frac{h_{feo}R}{h_{ie}}\right) \approx h_{feo}f_\beta \approx f_1 \qquad (22.30)$$

assuming R is of the same order as h_{ie}.

22.10 The Operational Amplifier

The shunt voltage feedback amplifier can be used to perform certain linear and non-linear "mathematical" operations (e.g. adding,

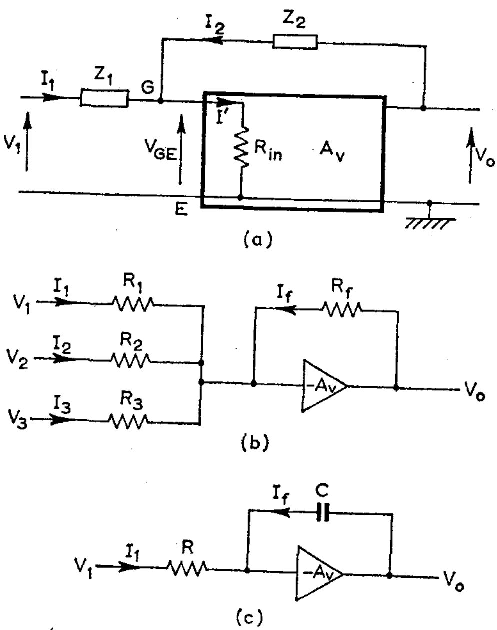


Fig. 22.14 OPERATIONAL AMPLIFIER

integrating and inverting) provided that the gain, A_v, of the am without feedback is negative and very high (generally at leas and usually higher). The amplifier is then called a *virtual operational* amplifier. The basic circuit is shown in Fig. 2

Several important features follow the assumption of high gain.

(a) If A_v $(= V_o/V_{GE})$ is very large it follows that, since V_o is limited by the supply voltage to the order of 100 V, V_{GE} must be very small and G is virtually at earth potential.

(b) The impedance of the feedback path seen looking into G is $Z_2/(1 - A_v)$ (from eqn. (22.19)), and as $A_v \to -\infty$ this tends to zero.

(c) Since $Z_2/(1 - A_v)$ will in general be very small compared to the input impedance, R_{in}, of the amplifier without feedback, it follows that $I' \ll -I_2$, so that all the input current can be assumed to flow in Z_2, or in other words $I_2 = -I_1$.

(d) Since $V_{GE} \to 0$, the input voltage, V_{in}, must all appear across Z_1 and the output voltage V_o across Z_2. Hence

$$V_{in} = I_1 Z_1 \quad \text{and} \quad V_o = I_2 Z_2$$

or

$$\frac{V_o}{V_{in}} = \frac{I_2 Z_2}{I_1 Z_1} = \frac{-Z_2}{Z_1} \tag{22.31}$$

At (b) the amplifier is shown arranged to add several input signals. For this circuit

$$I_f = -I_1 - I_2 - I_3$$

Hence

$$\frac{V_o}{R_f} = -\frac{V_1}{R_1} - \frac{V_2}{R_2} - \frac{V_3}{R_3}$$

or

$$V_o = -\frac{R_f}{R_1} V_1 - \frac{R_f}{R_2} V_2 - \frac{R_f}{R_3} V_3 \tag{22.32}$$

The output voltage is thus the inverse of the sum of the input voltages, taken in ratios determined by the input and feedback resistances. Obviously, if these resistances are all of equal magnitude,

$$V_o = -(V_1 + V_2 + V_3)$$

If only one input is used the amplifier becomes a simple invertor $(V_o = -V_1)$ when the input and feedback resistances are equal.

At (c) the feedback element is a capacitor. In this case, using instantaneous values,

$$v_1 = i_1 R \quad \text{and} \quad v_o = \frac{1}{C}\int_0^t i_f \, dt$$

Since $i_1 = -i_f$, it follows that

$$v_o = -\frac{1}{CR}\int_0^t v_1 \, dt \tag{22.33}$$

The output voltage in this case is the time integral of the input voltage multiplied by the constant $1/CR$ (which is usually chosen as unity). If the input consists of a d.c. step, the output will be a ramp, which will rise linearly until the amplifier becomes saturated.

When a capacitor is used to provide the input impedance and a resistor to provide the feedback impedance, an analysis similar to the above shows that the amplifier will differentiate the input signal. Differentiators are not commonly used since they tend to introduce instability and noise because the gain rises with frequency.

It should be noted that it has been assumed that the gain of the amplifier without feedback is very much greater than unity, and is negative. For low-gain amplifiers, an analysis must use eqn. (22.4), and the addition and integrating properties are adversely affected.

22.11 D.C. Amplifiers—the Long-tailed Pair

The amplifiers described in Section 22.10 must have a high gain down to zero frequency—i.e. they are *d.c. amplifiers*. Such amplifiers find applications in other areas also, such as instrumentation and power-supply stabilizers. In general d.c. amplifiers are inherently more difficult to design and stabilize than the corresponding a.c. amplifiers.

In capacitively-coupled or transformer-coupled a.c. amplifiers the coupling capacitors or transformers act as d.c. blocks, so that small, slow changes in the quiescent point of one stage (due to slow change of temperature or supply voltage, for example) are not passed on to the next stage; sudden changes are not passed on permanently. Such couplings cannot be used in d.c. amplifiers which must be direct coupled, and hence any change in the quiescent point will affect the output and will not be distinguishable from a change in the d.c. input signal. Any change in the output of a d.c. amplifier that is not due to a change in the input is called *drift*. Amplifier drift is usually expressed in terms of the equivalent input signal that will produce the same change of output. Drift is one of the main problems in d.c. amplifier design.

The chief sources of drift are:

(*a*) Changes in temperature, which (*i*) can change transistor leakage currents and base–emitter voltages, (ii) can give rise to changes in contact potentials, and (iii) can alter the ohmic value of resistors (hence silicon transistors and low-temperature-coefficient resistors should be used, together with temperature-compensating techniques).

(*b*) Changes in the supply voltage which alter the quiescent point (hence the supplies to d.c. amplifiers should be very well stabilized.

(*c*) Random changes and ageing of components (hence stable, long-life components should be used, together with circuits balanced to give zero output with zero input.

In multi-stage d.c. amplifiers any drift in the first stage is amplified in succeeding stages, so that the overall drift is largely determined by that of the first stage.

Drift cannot be reduced by straightforward feedback since this reduces the gain in proportion; the technique of compensation may however be employed. In this method drift in one part of a circuit is balanced against corresponding drift in another part. The long-tailed or emitter-coupled pair circuit shown in Fig. 22.15 is an example.

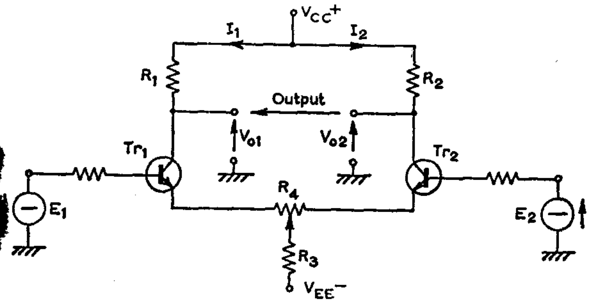


Fig. 22.15 BASIC LONG-TAILED PAIR CIRCUIT

In this circuit, ideally, R_1 and R_2 are equal and the two transistors are matched and mounted together so that they are both subject to the same temperature variations. The "tail" resistor, R_3, should be large (ideally the tail should be a constant-current source). The circuit operates either as a *difference amplifier* (giving a differential output, $(V_{o1} - V_{o2})$), which is proportional to the difference $(E_1 - E_2)$, between the isolated inputs), or (when one of the inputs is set at zero) as a *single-ended stage*.

Drift is minimized by this configuration. Thus if the supply voltage changes, both halves of the circuit are equally affected, and the differential output is (ideally) unchanged. Similarly, temperature changes will (ideally) give the same change in leakage current and base–emitter voltage in both transistors, so that again the current changes in both halves of the circuit will be the same and the differential output will not alter. Drift due to changes in the current gain, h_{fe}, will also tend to cancel, and may be further reduced by operating at low bias currents, using low source impedances and having transistors with high values of matched current gain. In a practical

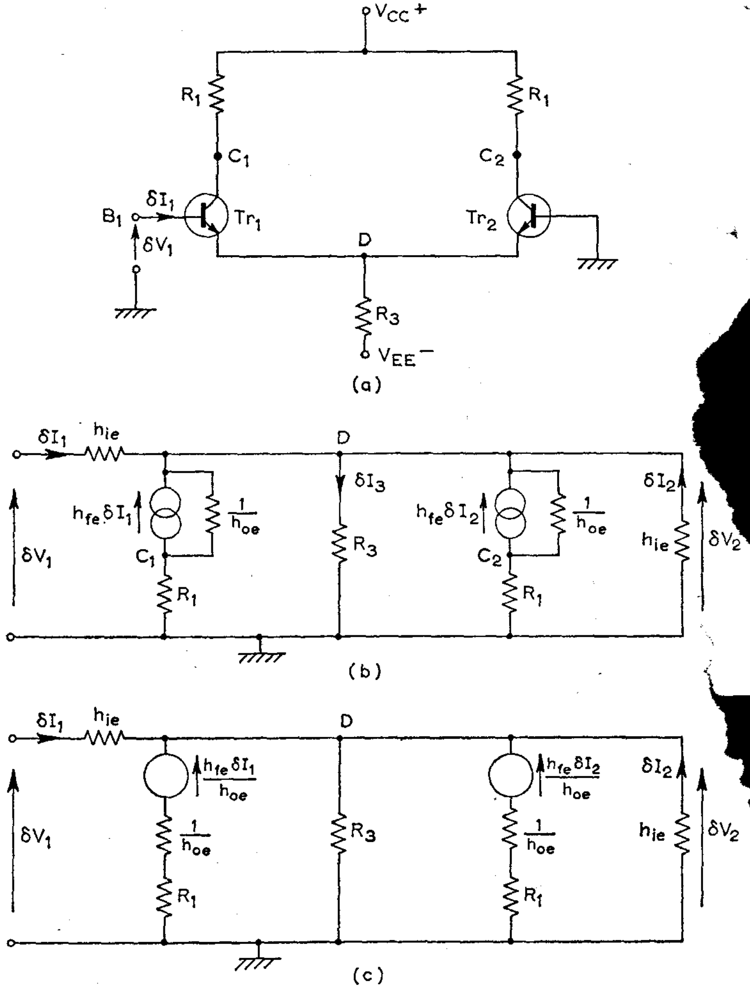


Fig. 22.16 EQUIVALENT CIRCUITS FOR LONG-TAILED PAIR

circuit there will, of course, be some residual drift, but the above compensating actions will operate to keep this small.

The initial balance of the circuit is achieved by setting both inputs to zero (or e_1 to zero in the single-ended case) and adjusting R_4 until the differential output is zero.

Consider a single-ended input applied to Tr_1 with the base of Tr_2 held at a fixed potential. Tr_1 acts as an emitter-follower stage with the emitter load consisting of the tail resistor R_3 in parallel with the input impedance of the Tr_2 circuit, which acts as a common-base circuit. If R_3 is large compared to this input impedance, then any change in current in Tr_1 gives rise to an almost equal and opposite change of current in Tr_2. Hence if V_{o1} rises, V_{o2} falls by an almost equal amount, and the circuit will act as a phase splitter with antiphase outputs δV_{o1} and δV_{o2}, where $\delta V_{o1} \approx -\delta V_{o2}$.

Assuming linear operation, suppose that the input to Tr_2 now varies by an equal but opposite amount to the input to Tr_1. The total change in output voltage will be twice that for the single-ended input. In general, therefore, the differential output will be proportional to the difference between the input signals, i.e. to the differential input.

If the two inputs change together by the same amount and in the same sense there will ideally be no change in the differential output voltage. The quality of a practical differential amplifier may be measured in terms of its ability to amplify only the difference between the input signals and to reject the signal common to both inputs. The property is called *common-mode rejection*, and numerically it is expressed as a ratio:

$$\text{Common-mode rejection ratio} = \frac{\text{Gain for antiphase inputs}}{\text{Gain for in-phase inputs}} \tag{22.34}$$

The higher this ratio (which is normally expressed in decibels) the better is the amplifier.

An approximate analysis of the long-tailed pair circuit of Fig. 22.16(*a*) follows from the approximate equivalent circuit at (*b*), where the transistors are represented by their common-emitter equivalents, neglecting h_{re}. The two collector resistors are assumed to be equal. Then for a single-ended input change of δV_1, the current changes will be

$$\delta I_1 = \frac{\delta V_1 - \delta V_2}{h_{ie}} \quad \text{and} \quad \delta I_2 = -\frac{\delta V_2}{h_{ie}}$$

The constant-current sources, $h_{fe}\delta I_1$ and $h_{fe}\delta I_2$, and their shunt resistors, $1/h_{oe}$, may be replaced (Norton/Thévenin conversion) by the constant-voltage sources of e.m.f. (h_{fe}/h_{oe}), $(h_{fe}/h_{oe})\delta I_2 I_1$ and series resistors $1/h_{oe}$ as shown at (*c*). Then, using Millman's theorem,

$$\delta V_2 = \frac{\delta V_1/h_{ie} + h_{fe}\delta I_1/h_{oe}(R_1 + 1/h_{oe}) + h_{fe}\delta I_2/h_{oe}(R_1 + 1/h_{oe})}{2/h_{ie} + 1/R_3 + 2/(R_1 + 1/h_{oe})}$$

Substituting for δI_1 and δI_2 and collecting like terms,

$$\delta V_2 \left\{\frac{2}{h_{ie}} + \frac{1}{R_3} + \frac{2}{(R_1 + 1/h_{oe})} + \frac{2h_{fe}}{h_{oe}h_{ie}(R_1 + 1/h_{oe})}\right\}$$

$$= \delta V_1 \left\{\frac{1}{h_{ie}} + \frac{h_{fe}}{h_{oe}h_{ie}(R_1 + 1/h_{oe})}\right\} \qquad (22.35$$

Since $2/h_{ie}$, $1/R_3$ and $2/(R_1 + 1/h_{oe})$ are all very much smaller $h_{fe}/h_{oe}h_{ie}(R_1 + 1/h_{oe})$, it follows that

$$\delta V_2 \approx \frac{\delta V_1}{2} \qquad (22.$$

To find the differential gain, the changes in the collector voltages at C_1 and at C_2 must be determined, and this can be done by referring back to Fig. 22.16(*b*). The change in output voltage at collector C_1 is then, since $1/h_{oe} \gg R_1$,

$$\delta V_{C1} = -h_{fe}R_1\delta I_1 = -h_{fe}R_1(\delta V_1 - \delta V_2)/h_{ie} \approx \frac{-h_{fe}R_1}{2h_{ie}}\delta V$$

and at collector C_2 is

$$\delta V_{C2} = -h_{fe}R_1\delta I_2 = +\frac{h_{fe}R_1\delta V_2}{h_{ie}} \approx \frac{h_{fe}R_1}{2h_{ie}}\delta V_1$$

The differential output voltage is therefore

$$\delta V_{C1} - \delta V_{C2} = -\frac{h_{fe}R_1\delta V_1}{h_{ie}} \qquad (22.37)$$

By superposition, the gain with equal antiphase inputs, δV_1, will be

$$\delta V_{C1} - \delta V_{C2} = -\frac{2h_{fe}R_1\delta V_1}{h_{ie}} \qquad (22.38)$$

The input impedance of the amplifier for single-ended operation will be

$$R_{in} = \frac{\delta V_1}{\delta I_1} = \frac{\delta V_1 h_{ie}}{(\delta V_1 - \delta V_2)} \approx 2h_{ie} \tag{22.39}$$

since for differential input $(\delta V_1 - \delta V_2) = \delta V_1 - (-\delta V_1) = 2\delta V_1$.

Differential gains of several hundred can be readily achieved, with input impedances of well over a kilohm. The long-tailed pair forms the basis of all integrated-circuit operational amplifiers.

An alternative approach to the problem of d.c. amplifiers is to convert the d.c. input signal to a.c. by a mechanical or electronic chopper. The a.c. signal is then amplified in a conventional amplifier, and is reconstructed to d.c. by a phase-sensitive rectifier at the output. e main disadvantage of this approach is the reduction in high-uency response which results. D.C. amplifiers can also be de-ed in which the chopper and a.c. amplifier are placed in the back loop. These are known as *chopper-stabilized amplifiers*, have a considerably greater bandwidth than the simple chopper ifiers, though still less than that of the long-tailed pair.

PROBLEMS

22.1 In an amplifier having an overall amplification of 50,000 without feedback, 0·02 per cent of the output voltage is fed back in antiphase to the input. Calculate the percentage reduction in amplification if the overall amplification, without feedback, falls to 40,000.

Ans. 2·2 per cent.

22.2 A resistance-capacitance-coupled amplifier has a voltage gain of −80. The coupling capacitor has negligible reactance at the frequency of operation. A very-high-resistance voltage divider across the output taps off one-tenth of the output voltage, and this is fed back and connected in series with the input voltage to give negative feedback. If there is an output ripple voltage of 1·8 V r.m.s. due to the supply to the amplifier without feedback calculate the ripple voltage appearing in the output of the amplifier when feedback is applied.

Ans. 200 mV.

22.3 A single-stage *RC* amplifier has a mid-frequency voltage gain of 46 dB and an upper 3 dB cut-off frequency of 1·5 MHz. Determine the mid-frequency gain and the bandwidth if 0·01 p.u. negative shunt voltage feedback is applied. Also evaluate the overall gain and bandwidth for two such stages in cascade.

Ans. −66·7; 4·5 MHz; 4,600; 2·9 MHz.

22.4 An amplifier without feedback has a voltage gain A_v of −1,500. Determine the percentage negative feedback that must be applied so that the sensitivity of the gain to changes in A_v is reduced to 0·01 of the value without feedback. What is the new overall gain?

Ans. 6·6 per cent; −15.

22.5 In the circuit of Fig. 22.17 the component values are $R_1 = 33$ kΩ, $R_2 = 10$ kΩ; $R_L = 4{\cdot}7$ kΩ, $R_E = 330$ Ω. The h-parameters of the transistor are $h_{ie} = 2{\cdot}5$ kΩ; $h_{fe} = 60$; $h_{oe} = 50$ μS. Determine the input impedance, current gain, voltage gain and output impedance. Do not neglect the effect of R_1 and R_2 on the equivalent a.c. circuit. Assume a constant-current source.

Ans. 5·4 kΩ; −14·3; −12·3; 50·4 kΩ.

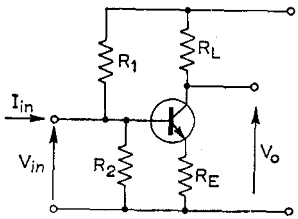


Fig. 22.17

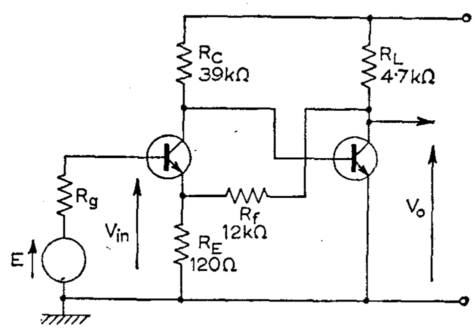


Fig. 22.18

22.6 Calculate the mid-frequency voltage gain and the input impedance of the simplified amplifier circuit of Fig. 22.18. For the transistors it may be assumed that $h_{fe} = 70$, $h_{ie} = 1{\cdot}2$ kΩ, $h_{oe} = h_{re} = 0$. The d.c. bias circuit may be assumed to give an equivalent resistance of 30 kΩ at the input terminals.

Ans. 98·8; 25·2 kΩ.

22.7 The circuit shown in Fig. 22.19 has a two-stage amplifier with an overall gain of +100. The input and output impedances are very large. If $R_1 = 0{\cdot}1$ MΩ and $C = 2{\cdot}5\mu$F, determine the value of R_2 which is required if 100e is to be equal to $K\int V_{in}\,dt$ and find the value of K.

The amplifier saturates when $e = 1{\cdot}0$ V. Find the maximum time over which an input step of 5 V can be integrated. (Note that the feedback is positive.)

Ans. 9·9 MΩ; $K = 400$; 50 ms.

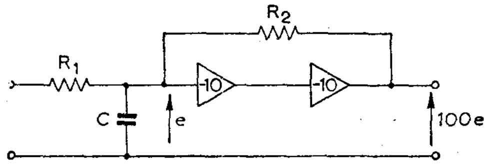


Fig. 22.19

22.8 A two-stage d.c. amplifier has a gain $A_{vo} = 1{,}550/(1 - f^2 10^{-8} + j2f10^{-4})$ where f is in hertz. Negative voltage feedback using a network of pure resistances s applied to reduce the d.c. gain to 50. Determine the fraction of the output ltage that is fed back, and calculate the gain with and without feedback at 0 and 50kHz. Sketch the modulus of the gain/frequency response with and ut feedback, and comment on the shapes. (*L.U.*)
ote that the feedback becomes positive as the frequency increases.]

s. 0·019; no feedback, $775\underline{/-90^\circ}$; $92\underline{/-152^\circ}$; $60\underline{/-157^\circ}$; with feedback, $50\underline{/1^\circ}$; $83\underline{/-25^\circ}$; $121\underline{/-51^\circ}$.

22.9 The load on a C.E. transistor amplifier is a tuned circuit of dynamic resistance 100kΩ. The amplifier has series current feedback through an unbypassed 100Ω resistor in the emitter lead. The resonant circuit is tapped in the ratio 1:6 for matching to the transistor. If $1/h_{oe} = 20\,\text{k}\Omega$, $h_{ie} = 1\,\text{k}\Omega$, $h_{fe} = 60$, and the source generator has an internal resistance of 5kΩ, determine the effective additional damping due to the transistor circuit, and the dynamic efficiency. (Dynamic efficiency = (power in tuned circuit)/(transistor output power).)

Ans. 710kΩ; 0·87p.u.

22.10 A transistor amplifier uses shunt voltage feedback. It is fed from a source of internal resistance 2kΩ. The feedback resistance is 10kΩ and the collector load resistance is 1kΩ. If $h_{ie} = 1{\cdot}2\,\text{k}\Omega$, $1/h_{oe} = 40\,\text{k}\Omega$ and $h_{fe} = 70$, determine the output voltage if the signal generator e.m.f. is 200mV.

Ans. 0·81V.

22.11 An emitter-follower circuit is fed from a C.E. amplifier which has a collector circuit resistance of 5kΩ. The emitter-circuit load of the emitter follower is 500Ω. For a given current input to the first transistor, determine the ratio of the output voltage of the emitter follower to the voltage across a 500Ω resistor used as the load on the first stage with the emitter follower disconnected. Also find the output impedance of the emitter follower. Both transistors have $h_{ie} = 1\,\text{k}\Omega$, $h_{oe} = 50\,\mu\text{S}$ and $h_{fe} = 90$.

Ans. 8·1; 60Ω.

22.12 A transistor amplifier has a $2{\cdot}5\,\mathrm{k}\Omega$ load resistor. Bias is obtained through a $60\,\mathrm{k}\Omega$ resistor connected between collector and base. This resistor is divided into two $30\,\mathrm{k}\Omega$ sections, with the mid-point connected to the supply rail by a very large capacitor C. Calculate the current gain and input impedance when (*a*) C is connected (hence no negative signal feedback), (*b*) C is removed. The transistor has $h_{ie} = 800\,\Omega$; $h_{oe} = 60\,\mu\mathrm{S}$; $h_{fe} = 100$.

Ans. 87, $780\,\Omega$; 19, $17\,\Omega$.

Chapter 23

OSCILLATIONS AND PULSES

In the previous chapter it was seen that negative feedback in general improves the stability but decreases the gain of an amplifier. If the feedback is positive the stability will decrease but the gain will increase. This increase in gain may be such that an a.c. output is obtained without an input, and the amplifier has become an oscillator, giving an a.c. output whose energy is obtained from the d.c. supply.

Oscillators may be divided into two types according to the output waveform:

(*a*) *Sinusoidal* oscillators, giving pure sine-wave outputs.
(*b*) *Relaxation* oscillators, giving pulse or square-wave outputs.

All oscillators work on a negative-resistance principle, and each of the above types may be further divided into (i) *retroaction* oscillators, where the negative resistance is obtained by use of the feedback principle, and (ii) *dynatron* oscillators, where the device itself introduces the negative resistance. Basically, if a tuned circuit is connected in parallel with a negative conductance (i.e. a system in which an *increase* in current gives rise to a *reduction* in voltage), then when the negative conductance is equal to the equivalent parallel conductance of the tuned circuit, the total parallel conductance is zero. The circuit damping will then be zero and oscillations will be sustained in the tuned circuit at its self-resonant frequency.

Only retroactive oscillators will be considered in this chapter.

23.1 General Condition for Oscillation

The block diagram of a voltage amplifier of gain A_v without feedback is shown in Fig. 23.1(*a*). As in the previous chapter, the expression for gain with feedback is

$$A_{vf} = \frac{V_o}{V_i} = \frac{A_v}{1 - \beta A_v}$$

If the feedback is positive the term βA_v is positive, and A_{vf} is greater than A_v, and indeed rises towards infinity when $1 - \beta A_v = 0$, or

$$\beta A_v = 1 \tag{23.1}$$

To investigate this condition further, consider the feedback loop broken as shown in Fig. 23.1(*b*). With a given a.c. signal, V',

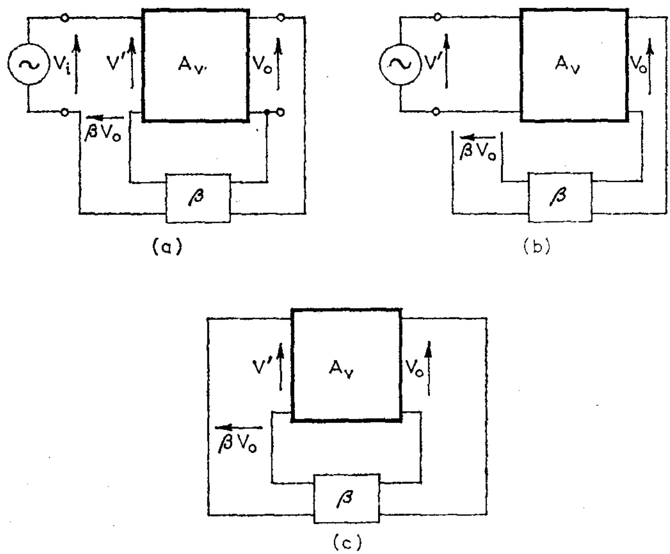


Fig. 23.1 DEVELOPMENT OF A FEEDBACK OSCILLATOR

applied, the loop gain is defined as $\beta V_o/V'$. If this is unity, then the input source can be removed and replaced by the feedback voltage as shown at (*c*). The circuit will continue to give a signal output, and has become an oscillator. In this condition the source of the signal is in fact the noise in the circuit. This noise is selectively amplified to give the output. For this to occur,

$$\frac{\beta V_o}{V'} = 1 \qquad \text{or} \qquad \beta A_v = 1 \quad \text{(as above)}$$

Since in general both β and A_v may be complex, eqn. (23.1) implies two conditions:

(*a*) In magnitude, $\beta A_v = 1$.
(*b*) The phase of βA_v is zero or any multiple of 2π.

For sinusoidal oscillators the open loop is designed to be frequency dependent, so that only at one frequency is the condition $\beta A_v = 1$ fulfilled, and this is the frequency of oscillation.

If the magnitude of βA_v is less than unity when the phase angle is zero or an integral multiple of 2π, then insufficient voltage will be fed back to maintain the output voltage in the absence of an input. If the external input voltage, V_i, shown at (*a*) is removed, the output will fall to zero.

If $\beta A_v > 1$ when the phase angle of βA_v is zero or $2\pi n$ then the feedback voltage is more than enough to maintain the output voltage even when the external input is removed, and any oscillation will increase in magnitude. With the large voltage swings which will become established, the amplifier will operate in a non-linear mode, so that A_v will fall. Eventually this fall in A_v will bring βA_v to unity, and the amplitude of the output voltage will become stabilized at the level at which $\beta A_v = 1$. Any further increase in the output voltage will cause A_v to fall further, and the output voltage will also tend to fall to its stable level. Similarly, any random decrease in the output voltage will cause A_v to rise, so that the output voltage will also tend to rise to its stable level. Thus, if the condition $\beta A_v > 1$ when arg $\beta A_v = 0$ or $2\pi n$ is satisfied for small voltage variations (i.e. for operation in the linear region of the amplifier), oscillations will build up to a stable value.

In addition, if $\beta A_v > 1$, not only will oscillations be maintained but they will be self-starting. This is due to the noise voltages which must exist in any circuit, and which can be shown to cover a wide range of frequencies. Selective amplification causes the build-up of oscillations at the frequency for which the phase condition arg $\beta A_v = 0$ or $2\pi n$ is satisfied.

EXAMPLE 23.1 A three-stage transistor amplifier uses transistors for which, at frequency f, $h_{fe} = 30/(1 + jf/10^6)$; $h_{ie} = 1\,\text{k}\Omega$; and $h_{oe} = 50\,\mu\text{S}$. The final stage feeds a load of $1\,\text{k}\Omega$. Find the maximum series voltage feedback factor, β, that can be used if instability is to be avoided at high frequencies.

The collector load resistances are 5 kΩ in each case, and base-bias resistors can be neglected.

The effective load, R_L', of each stage is 5 kΩ in parallel with 1 kΩ, i.e. 840 Ω. The voltage gain per stage without feedback is thus, from eqn. (21.13),

$$A_v = \frac{-h_{fe}R_L'}{(1 + h_{oe}R_L')h_{ie}} = \frac{-30 \times 840}{(1 + jf/10^6)(1 + 0{\cdot}04) \times 1{,}000}$$

$$= \frac{24\underline{/180^\circ}}{(1 + jf/10^6)}$$

At frequencies for which $f/10^6 \ll 1$, each stage introduces 180° phase shift, so that the three stages introduce a total phase shift of 360° + 180°, i.e. a resultant phase shift of 180°. Any feedback from the last stage to the first will then be negative. If, however, each stage introduces a further phase shift of ±60°, the resultant phase shift will be zero, and the feedback will be positive, so that the amplifier may be unstable. In this case the phase angle of $(1 + jf/10^6)$ must be 60°. This means that

$$\arg(1 + jf/10^6) = 60^\circ$$

or

$$f/10^6 = \tan^{-1} 60^\circ = \sqrt{3}$$

or

$$f = \sqrt{3} \times 10^6\,\text{Hz}$$

At this frequency the value of A_v per stage is $24/\sqrt{(1 + 3)} = 12$, so that for the three stages it is $12^3 = \underline{\underline{1{,}730}}$.

Hence, if the feedback factor, β, is greater than 1/1,730, eqn. (23.1) will be satisfied and oscillations will build up at a frequency of $\sqrt{3}$ MHz.

It should be noted that, in an *RC*-coupled multistage amplifier, the coupling capacitors introduce phase shift at low frequencies. Oscillation may occur if this amounts to 60° per stage, and if the gain at such frequencies is large enough.

23.2 General Theory of Transistor *LC* Oscillators

In all oscillators the basic criterion is that the critical loop gain is unity (*Barkhausen criterion*); i.e. power, voltage and current gains are all unity. With bipolar junction transistors the input impedance is not infinite and it is most convenient to take the condition for oscillation that the loop current gain is unity, where this gain is calculated by opening the feedback loop and terminating it in an impedance, h_{ie}, equal to the input impedance of the transistor. The current in this terminating impedance with the open feedback loop will then be the input current to the first transistor when the loop is closed.

Consider the general circuit shown in Fig. 23.2(*a*), where only the a.c. components are shown. The a.c. equivalent circuit is shown at (*b*), where the resistance $1/h_{oe}$ has been assumed large compared

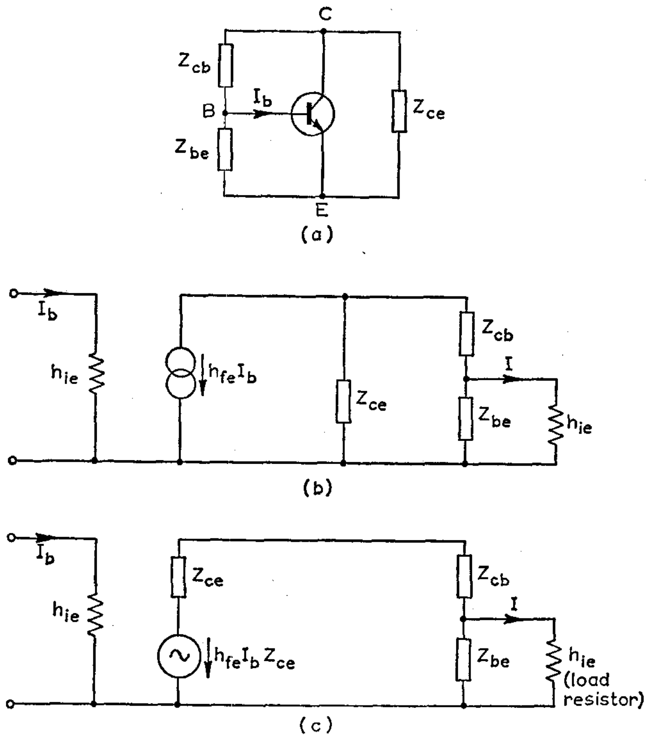


Fig. 23.2 GENERALIZED TRANSISTOR *LC* OSCILLATOR

with the load impedance, Z_{ce}. Thévenin's theorem is now used to transform the constant-current source $h_{fe}I_b$ and impedance Z_{ce} to the constant-voltage equivalent as shown at (*c*). The circuit is loaded by a resistor h_{ie} and the feedback connexion is not made. In this case the current I in the load resistor h_{ie} is

$$I = \frac{-h_{fe}I_bZ_{ce}}{Z_{ce} + Z_{cb} + Z_{be}h_{ie}/(Z_{be} + h_{ie})} \frac{Z_{be}}{(Z_{be} + h_{ie})}$$

$$= \frac{-h_{fe}Z_{ce}Z_{be}I_b}{(Z_{be} + h_{ie})(Z_{ce} + Z_{cb}) + Z_{be}h_{ie}}$$

Now, for oscillation to take place, $I = I_b$. Hence

$$I_b = \frac{-h_{fe}Z_{ce}Z_{be}I_b}{Z_{be}Z_{ce} + Z_{be}Z_{cb} + h_{ie}(Z_{ce} + Z_{cb} + Z_{be})}$$

so that

$$Z_{be}Z_{ce} + Z_{be}Z_{cb} + h_{ie}(Z_{ce} + Z_{cb} + Z_{be}) = -Z_{ce}Z_{be}h_{fe}$$

If it is now assumed that Z_{be}, Z_{ce} and Z_{cb} are pure reactances, it follows that the terms $Z_{bc}Z_{ce}$, $Z_{be}Z_{cb}$ and $Z_{ce}Z_{be}h_{fe}$ are positive or negative reference terms, while $h_{ie}(Z_{be} + Z_{cc} + Z_{eb})$ is quadrate. Equating reference terms and substituting reactances for impedances where appropriate*,

$$X_{be}X_{ce} + X_{be}X_{cb} = -X_{ce}X_{be}h_{fe}$$

or

$$X_{cb} + (1 + h_{fe})X_{ce} = 0 \tag{23.2}$$

so that X_{cb} and X_{ce} must be opposite types of reactance, i.e. if X_{cb} is inductive then X_{ce} must be capacitive and vice versa.

Equating the quadrate term to zero,

$$h_{ie}(X_{ce} + X_{cb} + X_{be}) = 0$$

and since $h_{ie} \neq 0$ it follows that

$$X_{ce} + X_{cb} + X_{be} = 0 \tag{23.3}$$

Substituting for X_{cb} from eqn. (23.2) in eqn. (23.3),

$$X_{ce} - (1 + h_{fe})X_{ce} + X_{be} = 0$$

or

$$X_{be} = h_{fe}X_{ce}$$

so that X_{be} and X_{ce} must be reactances of the same type.

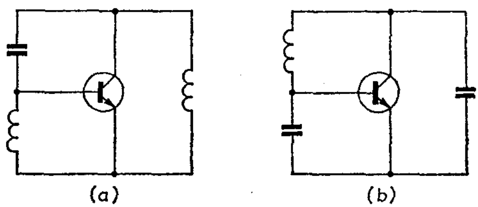


Fig. 23.3 THE TWO BASIC *LC* OSCILLATOR CIRCUITS

Note that eqn. (23.3) gives the frequency condition for oscillation, while eqn. (23.2) gives the necessary gain condition for the maintenance of oscillation. Also, the relation between the reactances is

$$X_{ce}:X_{be}:X_{cb} = 1:h_{fe}:-(1 + h_{fe}) \tag{23.4}$$

There are therefore basically two types of *LC* oscillator as shown in Fig. 23.3. These lead to several variants, depending on the position of the d.c. supply.

* X_{be} will represent $\pm jX_{be}$, and so on.

23.3 Colpitts and Hartley Oscillators

These oscillators illustrate the application of the general theory very simply. The Colpitts oscillator is basically the two-capacitor type shown in Fig. 23.4. The bias circuit is required to ensure that the transistor will be biased "on" when the supply is connected, so that

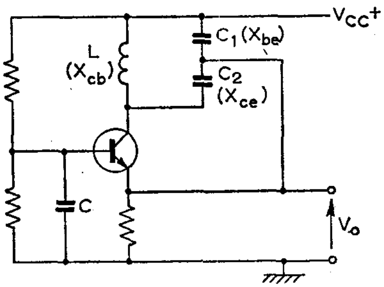


Fig. 23.4 PRACTICAL COLPITTS OSCILLATOR

the oscillator will be self-starting. Capacitor C acts as a base-decoupling capacitor for the signal. For this circuit, eqn. (23.3) gives the frequency condition as

$$\frac{1}{j\omega C_2} + j\omega L + \frac{1}{j\omega C_1} = 0$$

or

$$\omega^2 = \frac{1}{L}\left(\frac{1}{C_1} + \frac{1}{C_2}\right) \tag{23.5}$$

while eqn. (23.2) gives the gain condition as

$$j\omega L + \frac{(1 + h_{fe})}{j\omega C_2} = 0$$

i.e.

$$1 + h_{fe} = \omega^2 LC_2 = \frac{C_2}{C_1} + 1$$

Hence

$$h_{fe} \geqslant \frac{C_2}{C_1} \tag{23.6}$$

Frequencies over 5 GHz are possible with this oscillator. Similar expressions can readily be obtained for the Hartley oscillator shown in Fig. 23.5. It is left as an exercise for the reader to show that, if

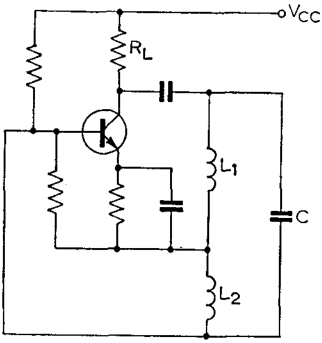


Fig. 23.5 PRACTICAL HARTLEY OSCILLATOR

there is no mutual coupling, M, between the coils L_1 and L_2, then the frequency of oscillation is given by

$$\omega^2 = \frac{1}{C(L_1 + L_2)} \tag{23.7}$$

and the gain condition by

$$h_{fe} \geqslant \frac{L_2}{L_1} \tag{23.8}$$

If $M \neq 0$, these expressions are modified to $\omega^2 = 1/C(L_1 + L_2 + 2M)$ and $h_{fe} = (L_1 + M)/(L_2 + M)$.

23.4 Frequency Stability

In the foregoing derivations, the frequency at which an oscillator operates has been found as a function of circuit components only. In practical oscillators, external loading, non-linearity of active device and component losses all introduce effects which make the frequency dependent on supply voltage, active device parameter changes and temperature. The frequency stability depends upon how rapidly the phase criterion expressed by eqn. (23.13) changes with frequency near the critical frequency. If the change is not rapid, small changes in parameters may cause a large frequency change as the closed loop adjusts itself to maintain the phase and gain response required for oscillation. Note that, if $\beta A_{vo} > 1$ when $\arg \beta A_{vo} = 0$ or $2\pi n$, then oscillations build up in amplitude until the amplifier operates in a non-linear region so that A_{vo} falls until $\beta A_{vo} = 1$. To achieve frequency stability one of the frequency-determining elements is often chosen as a tuned circuit, operated very slightly off resonance, so that it behaves either as an inductance (at frequencies below resonance) or a capacitance (when operated slightly above resonance). The higher the Q-factor of the tuned circuit, the greater will be the change in phase with frequency near the critical frequency, and hence the greater will be the frequency stability. Expressed mathematically the condition for frequency stability

$$\frac{d}{d\omega}(X_{ce} + X_{cb} + X_{be})$$

is as high as possible when

$$X_{ce} + X_{cb} + X_{be} \approx 0$$

Since a piezo-electric crystal behaves as a tuned circuit of very high Q-factor, oscillators in which one of the elements is a crystal have very high frequency stability.

The piezo-electrical effect is the mechanical contraction or elongation along one axis of a crystal when a voltage is applied across another axis (and vice versa). If a voltage pulse is applied then on its removal the mechanical strain relaxes and the strain energy is converted to stored electrical energy—a voltage appears on the voltage axis. This in turn leads to further mechanical stress. The crystal will continue to oscillate between mechanical and electrical stored energy states until internal losses use up all the original stored energy. It is thus equivalent to a tuned circuit as shown in Fig 23.6(a). Q-factors of the order of 50,000 are readily obtained from quartz crystals. Because of their larger electrical output crystals of Rochelle salt are commonly used in piezo-electric pick-ups, microphones and transducers. The self-resonant frequency of crystals is determined largely by their thickness—the thicker the crystal the lower the self-resonant frequency.

A simple piezo-electric oscillator is shown at (b). This is effectively a Colpitts oscillator—with the coil replaced by the crystal, which must therefore operate slightly below its self-resonant frequency. This is often called the Pierce oscillator. A similar circuit in which the capacitors C_1 and C_2 are replaced by coils gives the crystal equivalent of the Hartley oscillator.

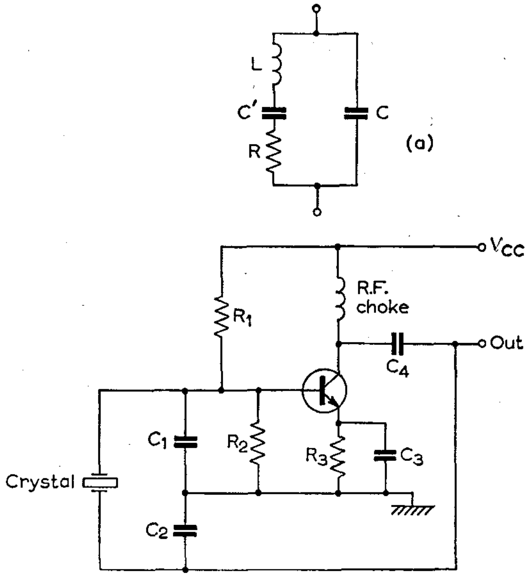


Fig. 23.6 PIERCE OSCILLATOR

23.5 Oscillators with Mutual Coupling

Mutual coupling is often used in oscillator circuits. When generalized theory is applied in these cases the mutual inductance should be replaced, as far as the calculations are concerned, by the equivalent π-circuit shown in Fig 23.7.

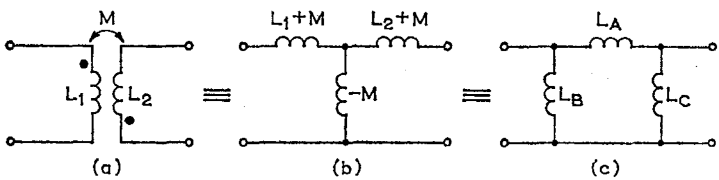


Fig. 23.7 EQUIVALENT CIRCUITS OF A MUTUAL COUPLING

By considering mesh equations it is easy to verify that the circuit at (*b*) with no mutual coupling is equivalent to that at (*a*), and the star–delta transformation then gives the equivalent π-circuit, where

$$L_A = \frac{L_1L_2 - M^2}{-M} \qquad L_B = \frac{L_1L_2 - M^2}{L_2 + M} \qquad L_C = \frac{L_1L_2 - M^2}{L_1 + M} \tag{23.9}$$

Note that L_A is a negative inductance. This means that it has a negative reactance which increases with frequency. A negative inductance in parallel with a positive inductance does not, however, form a tuned circuit with a resonant frequency, but simply gives a *larger* positive inductance (if the negative reactance is larger than the positive reactance).

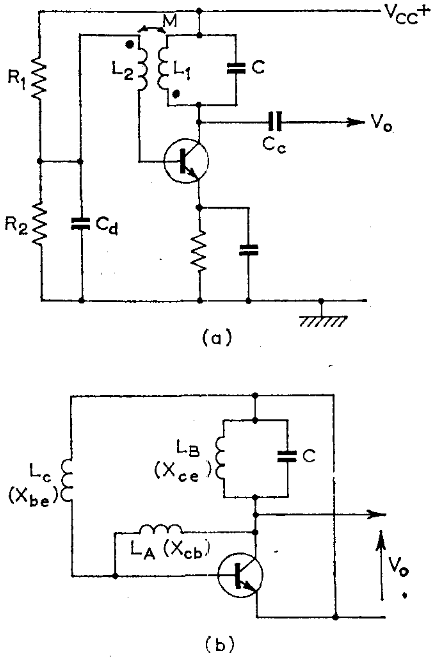


Fig. 23.8 TUNED-COLLECTOR OSCILLATOR

(*a*) Actual circuit (*b*) Equivalent a.c. circuit

Using the above relations it is seen that the tuned collector oscillator of Fig. 23.8(*a*) reverts to the general form as shown at (*b*). The inductance L_C is positive and L_A is negative (i.e. X_{be} is positive and X_{cb} is negative). Hence X_{ce} must be positive, i.e. L_B in parallel with C must operate *below* the resonant frequency, $f = 1/2\pi\sqrt{(L_BC)}$. The conditions for oscillation are then fulfilled provided that the gain is high enough. In a practical transistor oscillator of this type the tuned circuit coil will normally be tapped in order to reduce the loading of the circuit due to the transistor. A third winding is often used for the output.

The frequency of the oscillator will be slightly below $1/2\pi\sqrt{(L_B C)}$, i.e. slightly above $1/2\pi\sqrt{(L_1 C)}$. The frequency stability of the circuit can be high, but because the rate of change of reactance with frequency of the negative inductance, L_A, is negative, the stability is not so high as it would be if a capacitor were to replace L_A (the rate of change of reactance of the capacitor with frequency would be positive). Mathematically,

$$\frac{d}{d\omega}(X_{ce} + X_{cb} + X_{be})$$

is smaller if X_{cb} is the reactance of a negative inductance than if it is the reactance of a capacitor, and so the frequency stability is lower.

23.6 Transistor Wien Bridge Oscillator

For the circuit shown in Fig. 23.9(a) the voltage transfer function is

$$\frac{V_f}{V'} = \frac{1}{3 + j\omega CR + 1/j\omega CR}$$

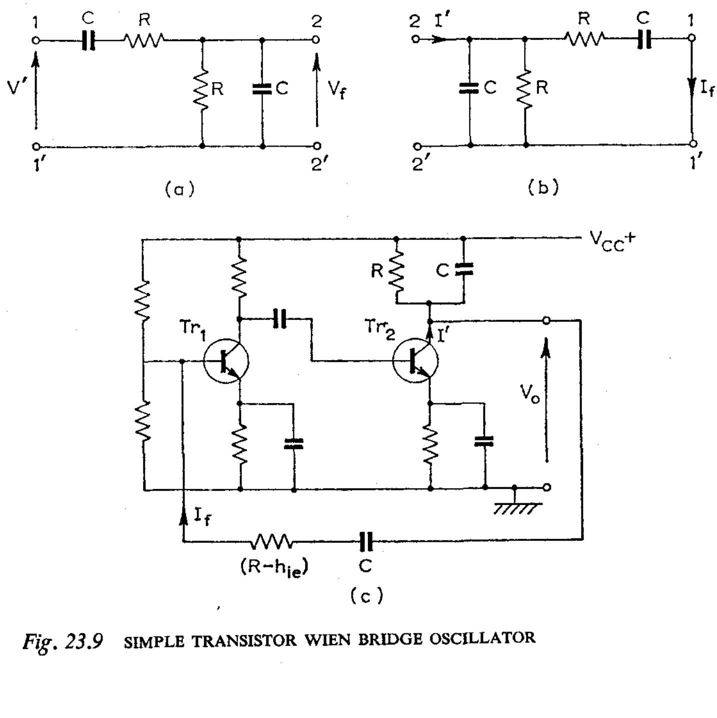


Fig. 23.9 SIMPLE TRANSISTOR WIEN BRIDGE OSCILLATOR

The current dual of this circuit is obtained from a simple application of the *principle of duality*. It follows from this principle that the ratio of output to input voltage for an *open-circuited* two-port passive network is the same as the ratio of the current through a *short-circuit* across the input terminals to the current fed into the output terminals. Thus, if the output terminals 2, 2′ in Fig. 23.9(*a*) are made the current input terminals at (*b*) of the current dual, it is easy to verify that

$$\frac{I_f}{I'} = \frac{1}{3 + j\omega CR + 1/j\omega CR}$$

In the oscillator circuit shown at (*c*) the feedback resistance is chosen to be $(R - h_{ie})$, so that when the transistor input resistance, h_{ie}, is taken into account, the circuit is equivalent to a capacitor C and resistor R in series feeding into a short-circuit. The input base current of transistor Tr_1 is

$$I_f = \frac{I'}{3 + j\omega CR + 1/j\omega CR} \tag{23.10}$$

neglecting the effect of the base bias resistors. Hence, if

$$\frac{I'}{I_f} > 3 \tag{23.11}$$

the circuit will oscillate at the frequency which makes the quadrate terms in eqn. (23.10) vanish; i.e. at a frequency given by

$$\omega CR = \frac{1}{\omega CR}$$

or

$$\omega = \frac{1}{CR} \tag{23.12}$$

It is, of course, relatively simple to achieve a current gain of 3 in a two-stage amplifier, but the frequency stability of this circuit is poor. For this reason the full Wien bridge oscillator shown in Fig. 23.10 is to be preferred.

This requires a differential amplifier whose output is a constant, A_v times the *difference* between the inputs; i.e.

$$V_o = A_v(V_1 - V_2) \tag{23.13}$$

The gain A_v must be positive reference number—i.e. the differential amplifier must have an even number of stages, as in the half Wien bridge. The bridge is *almost* balanced, so that the voltage across AB is small.

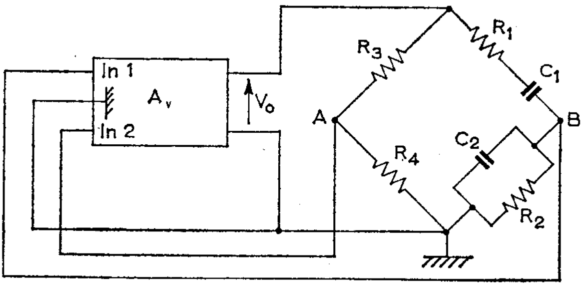


Fig. 23.10 THE FULL WIEN BRIDGE OSCILLATOR

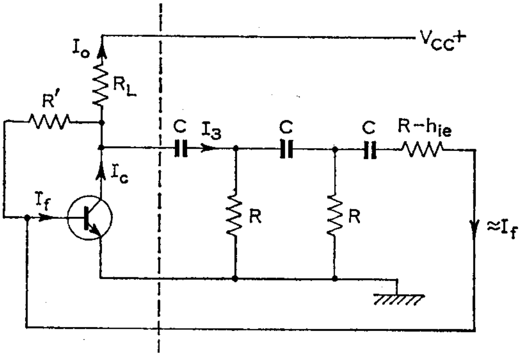


Fig. 23.11 SINGLE-STAGE *RC* OSCILLATOR

Analysis of the circuit shows that the frequency of oscillation is given by

$$f = \frac{1}{2\pi\sqrt{(C_1R_1C_2R_2)}}$$

and, since the voltage across AB is small, the condition of oscillation will be that A_v is large enough to satisfy eqn. (23.10) (high gains are however, easily obtained). The frequency stability of the circuit is high; i.e. the effective Q-factor of the selective network is much greater than that of the half Wien bridge.

23.7 Single-stage Transistor *RC* Oscillator

The circuit of a single-stage *RC* phase-shift oscillator is shown in Fig. 23.11. The circuit to the left of the broken line is a simple amplifier, in which the bias is obtained by means of the collector–base resistor R'. Since the amplifier is in the common emitter connexion, the collector current, I_c, will be approximately in antiphase with the base input current, I_f. Hence, for oscillation, the phase-shift network of C's and R's must introduce a further 180° phase shift, and the transistor current gain must compensate for the network attenuation. Assuming a transistor input resistance of h_{ie}, the following equation can be deduced (neglecting any current through the bias resistor R') by considering the current division between R_L and the feedback chain, and the current division in the feedback chain itself.

$$\begin{aligned} I_c &= I_o + I_3 \\ &= I_f\left\{3 + \frac{R}{R_L} - \frac{1}{\omega^2C^2R}\left(\frac{1}{R} + \frac{5}{R_L}\right) + \frac{1}{j\omega C}\left(\frac{4}{R} + \frac{6}{R_L}\right)\right. \\ &\qquad \left. - \frac{1}{j\omega^3C^3R^2R_L}\right\} \end{aligned} \tag{23.14}$$

Oscillation will occur at the frequency which makes the quadrate term zero, provided that the current gain makes I_c equal to I_f times the reference term. For zero quadrate term,

$$\frac{1}{\omega C}\left(\frac{4}{R} + \frac{6}{R_L}\right) = \frac{1}{\omega^3C^3R^2R_L}$$

or

$$\omega^2 = \frac{RR_L}{C^2R^2R_L(4R_L + 6R)} = \frac{1}{C^2R(4R_L + 6R)} \tag{23.15}$$

so that

$$\omega = \frac{1}{C\sqrt{[R(4R_L + 6R)]}} \tag{23.16}$$

When this condition is satisfied,

$$\frac{I_c}{I_f} = 3 + \frac{R}{R_L} - \frac{(4R_L + 6R)(R_L + 5R)}{RR_L} \tag{23.17}$$

(substituting for ω^2C^2R from eqn. (23.15) in the reference part of eqn. (23.14)).

The current gain, A_i, must be greater than this in order to sustain oscillations; i.e.

$$A_i > 3 + \frac{R}{R_L} - \frac{4R_L}{R} - 26 - \frac{30R}{R_L}$$
$$> -\left(23 + \frac{29R}{R_L} + \frac{4R_L}{R}\right)$$

Neglecting $1/h_{oe}$, since $1/h_{oe} \ll R_L$, and assuming $I_c \approx -h_{fe}I_b \approx -h_{fe}I_f$,

$$h_{fe} > 23 + \frac{29R}{R_L} + \frac{4R_L}{R} \tag{23.18}$$

It has been assumed that $1/h_{oe} > 10R_L$, and that $I_c \approx -h_{fe}I_b$.

For example, if $C = 0{\cdot}05\mu\text{F}$, $R_L = 3{\cdot}3\,\text{k}\Omega$ and $R = 5{\cdot}6\,\text{k}\Omega$, then the oscillation frequency is approximately

$$f = \frac{10^6}{2\pi \times 0{\cdot}05\sqrt{[3{\cdot}3(13{\cdot}2 + 33{\cdot}6) \times 10^6]}} = 250\,\text{Hz}$$

and $h_{fe} > 23 + 49 + 2 > 74$.

It is left as an exercise for the reader to show that if the resistor $(R - h_{ie})$ is omitted, the frequency of oscillation is given by

$$\omega^2 = \frac{1}{C^2(h_{ie}R_L + 3RR_L + 3R^2 + 3h_{ie}R)}$$

and the gain condition is that

$$h_{fe} > 1 + \frac{2h_{ie}}{R} + \frac{h_{ie}}{R_L} - \frac{(R_L + h_{ie} + 4R)(h_{ie}R_L + 3RR_L + 3R^2 + 3h_{ie}R)}{R^2R_L}$$

If $h_{ie} \to 0$ and $R_L = R$, these expressions reduce to

$$\omega^2 = \frac{1}{6C^2R^2} \quad \text{and} \quad h_{fe} > 29$$

In general, *RC* oscillators are convenient at low frequencies, since no large tuning inductance is required. Further, since the frequency of oscillation is inversely proportional to the capacitance (while in *LC* oscillators it is inversely proportional to the square root of the capacitance), the same capacitance change gives a larger frequency variation in *RC* than in *LC* oscillators.

23.8 Astable Multivibrator (Relaxation Oscillator)

In the *astable multivibrator*, two cascaded *RC* coupled stages are given 100 per cent positive feedback, and this gives rise to a conduction process whereby one stage is fully conducting while the other is switched off, conduction being transferred from one stage to the other by the discharge of a capacitor through a resistor. The output consists of square waves or pulses whose repetition frequency depends upon the *CR* time constants involved.

A simple circuit is shown in Fig. 23.12. The basic design criterion is that when either transistor is conducting it will operate in the

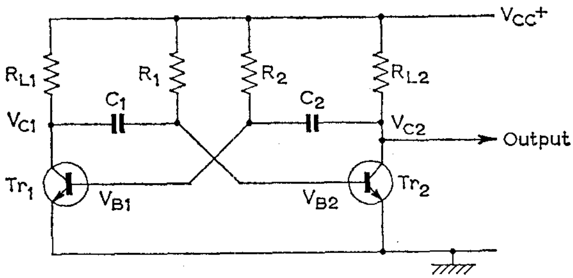


Fig. 23.12 CROSS-COUPLED ASTABLE MULTIVIBRATOR

saturation region. Up to the saturation value of collector current the relation $I_C \approx h_{FE} I_B$ applies (where h_{FE} is the large-signal d.c. current gain before saturation). If I_B is further increased there will be no corresponding increase in I_C since it has reached its saturation value (all that will happen is that I_E will increase slightly). Hence the criterion for saturation is $h_{FE} I_B > I_C$.

For transistor Tr_1, the base current in the saturated "on" state is approximately V_{CC}/R_2 (neglecting the transistor base–emitter voltage), and the collector current is V_{CC}/R_{L1}. Hence, for this state,

$$\frac{h_{FE}}{R_2} \gg \frac{1}{R_{L1}} \qquad \text{or} \qquad R_{L1} \gg \frac{R_2}{h_{FE}}$$

Similarly $R_{L2} \gg R_1/h_{FE}$ (normally R_{L1} and R_{L2} are chosen as equal).

The initial setting of the circuit can be seen by assuming that, when the d.c. supply is switched on, both transistors carry equal rising currents: some random variation will be bound to slightly increase or reduce one current with respect to the other—suppose that the current in Tr_2 is increased above that in Tr_1 by a small amount. The following reactions will occur:

(*a*) The voltage drop across R_{L2} will increase, so that V_{C2} will fall, and since the voltage across C_2 cannot change instantaneously, this fall in voltage is also effective at the base of Tr_1.

(*b*) Since V_{B1} has fallen, the base current in Tr_1 falls and the collector current is reduced, so that the collector voltage V_{C1} rises, this rise being transferred across C_1 to the base of Tr_2.

(*c*) Since V_{B2} rises it causes the current in Tr_2 to increase further, and the process repeats until Tr_2 is saturated and Tr_1 is cut off (provided that $R_{L2} \gg R_1/h_{FE}$).

The system is thus inherently unstable. It may be taken that the changes take place instantaneously since there are no reactive

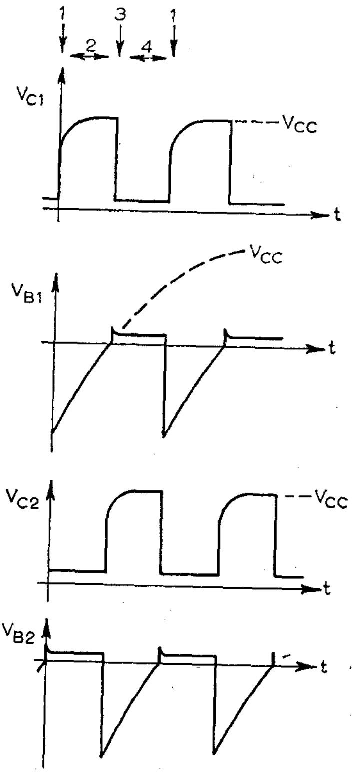


Fig. 23.13 WAVEFORMS OF ASTABLE MULTIVIBRATOR

elements except C_1 and C_2, whose instantaneous charges are constant. It is, of course, purely random which transistor becomes saturated and which is cut off initially.

The wave diagrams shown in Fig. 23.13 commence at the instant when Tr_1 is cut off, so that, to a fair approximation, $V_{B2} = 0$;

$V_{C2} = 0$; V_{B1} is rapidly going negative with respect to earth; and V_{C1} is rapidly rising to the supply voltage, V_{CC}. The reactions from this instant are as follows:

1. The positive-going voltage at the collector of Tr_1 is applied across C_1 and gives a base current flow in Tr_2, so that C_1 can rapidly charge through the forward-biased base–emitter diode of Tr_2. V_{B2} rises sharply above earth but rapidly falls to just above earth potential, so giving the small positive spike shown. It is convenient to assume that, just prior to the present voltage changes, V_{C2} has been at the supply voltage, V_{CC}, and that V_{B1} has been at zero voltage; this may not be so initially but becomes so after only a very few oscillations. Then the change in the collector voltage of Tr_2 is from V_{CC} (when Tr_2 is non-conducting) to zero (when Tr_2 is saturated). Since the charge on C_2 cannot change instantaneously, the base potential of Tr_1 must fall by the same amount, i.e. from about zero initially to $-V_{CC}$, so that the voltage across R_2 is $2V_{CC}$.

2. In the interval marked "2" in Fig. 23.13, relatively static conditions prevail. V_{C2} remains at about zero voltage, V_{C1} at $+V_{CC}$, and V_{B2} at about zero. However the voltage across R_2 is now $2V_{CC}$, and so a current flows in it which charges C_2 from $-V_{CC}$ towards $+V_{CC}$, the time constant of the exponential charging current being C_2R_2. The equation of this "relaxation" is

$$\begin{aligned} V_{B1} &= -V_{CC} + 2V_{CC}(1 - \mathrm{e}^{-t/C_2R_2}) \\ &= V_{CC} - 2V_{CC}\,\mathrm{e}^{-t/C_2R_2} \qquad (23.19) \end{aligned}$$

However, as soon as its base comes to zero voltage, Tr_2 starts conducting and a second stage of precipitate amplification (labelled "3") occurs. The variation of base voltage during phase 2 is shown, the broken curve indicating the "prospective" rise, which is cut short at instant 3.

3. As soon as V_{B1} reaches zero voltage, Tr_1 starts conducting. Hence V_{C1} falls, and the fall is transferred across C_1 (whose voltage cannot change instantaneously) to the base of Tr_2, which therefore becomes cut off. V_{C2} rises (short time constant, C_2R_{L2}) to V_{CC} so that (since the voltage across C_2 cannot change instantaneously) the base voltage of Tr_1 also rises, to saturate Tr_1 and cause its collector voltage to fall to about zero. The base voltage of Tr_2 falls to $-V_{CC}$ volts.

4. As soon as this precipitate amplification period is over, a further quiescent period follows, with $V_{C1} \approx 0$, $V_{C2} \approx V_{CC}$ and $V_{B1} \approx 0$. V_{B2} now rises (time constant, C_1R_1) from $-V_{CC}$ towards $+V_{CC}$. The equation for the rise is

$$V_{B2} = V_{CC} - 2V_{CC}\,\mathrm{e}^{-t/C_1R_1} \qquad (23.20)$$

The whole process repeats as soon as V_{B2} reaches zero, when a further stage of precipitate amplification takes place.

The output from either collector is thus a series of rectangular pulses, whose repetition frequency can now be found. Thus for the first relaxation period, the time, τ_1, required for the base voltage to rise from $-V_{CC}$ to zero is given, from eqn. (23.19), by

$$0 = V_{CC} - 2V_{CC}\, e^{-\tau_1/C_2R_2}$$

Hence

$$\exp(\tau_1/C_2R_2) = 2 \qquad \text{or} \qquad \tau_1 = C_2R_2 \log_e 2 \tag{23.21}$$

Similarly the time for the second relaxation period is

$$\tau_2 = C_1R_1 \log_e 2$$

and the *pulse repetition frequency* (p.r.f.) is

$$\text{P.R.F.} = \frac{1}{\tau_1 + \tau_2} = \frac{1}{(C_1R_1 + C_2R_2)\log_e 2} \tag{23.22}$$

For the symmetrical multivibrator, $C_1R_1 = C_2R_2 = CR$, and hence

$$\text{P.R.F.} = \frac{1}{2CR \log_e 2} = \frac{1}{1{\cdot}4CR} \quad \text{pulses/second} \tag{23.23}$$

The p.r.f. can be altered by (*a*) changing the CR constant or (*b*) returning the base resistors R_1 and R_2 to a different potential than that of the supply rail (increasing the voltage to which they are returned increases the frequency).

The frequency stability of transistor multivibrators is not high, and in particular the temperature sensitivity of the reverse leakage current I_{CEO} (especially in germanium transistors) causes frequency drift. This is because the leakage current at the base of the "off" transistor affects the charging rate of the "relaxing" capacitor. The basic circuit is not suitable for use with silicon planar transistors if $V_{CC} > 6$V since the base–emitter junction may break down when it becomes reverse biased.

23.9 Synchronizing

It is often desirable to lock the frequency of a multivibrator to that of a more frequency-stable oscillator that can produce short pulses (*clock pulses*). This can be done in the circuit of Fig. 23.12 by feeding positive clock pulses through diodes to the base connexions. The transition between states can then be made to correspond to clock impulses as shown in Fig. 23.14. If the clock pulse frequency is slightly higher than that of the multivibrator, the multivibrator will be brought into exact synchronism. If the clock pulse frequency is several times the multivibrator basic frequency, the multivibrator will act as a dividing circuit (as in Fig. 23.14) giving one output pulse

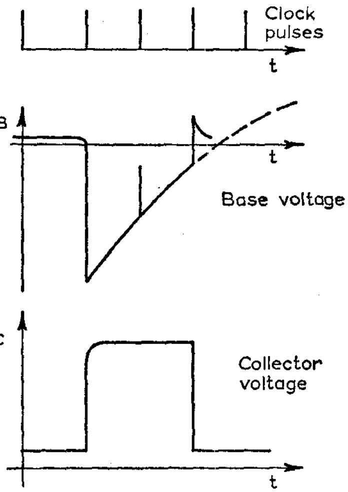


Fig. 23.14 SYNCHRONIZING WAVEFORMS FOR A MULTIVIBRATOR

for every 2, 3, 4 . . . etc. clock pulses. This division is not usually carried beyond 10, since it is then difficult to make synchronizing determinate and stable.

23.10 Monostable Multivibrator

The astable multivibrator has been seen to have no stable state, conduction being transferred continuously from one transistor to the other and back. The *monostable multivibrator* (Fig. 23.15(*a*)) has one stable state to which it will always revert. When an external trigger pulse is applied the circuit switches over, and then, after some delay, returns to the stable state until a further trigger pulse is applied. The circuit is used to give a time delay action, or to produce a known pulse width.

In the circuit shown, the voltage of the base of Tr_1 is set by the resistor chain R_L, R_2 and R_3 to be below the emitter voltage. Thus Tr_1 is normally cut off. The base of Tr_2 is connected through R_1 to the positive supply, so that Tr_2 is normally conducting (i.e. in the "on" state). This is the stable state, in which $V_{C1} = V_{CC}$ and $V_{C2} \approx 0$.

When a negative trigger pulse is applied to the base of Tr_2, the current in this transistor falls, V_{C2} rises, and since the voltage across C_2 cannot change instantaneously, the voltage V_{B1} also rises. This starts Tr_1 conducting; its collector voltage falls, and since the voltage across C_1 cannot change instantaneously, the base of Tr_2 also falls

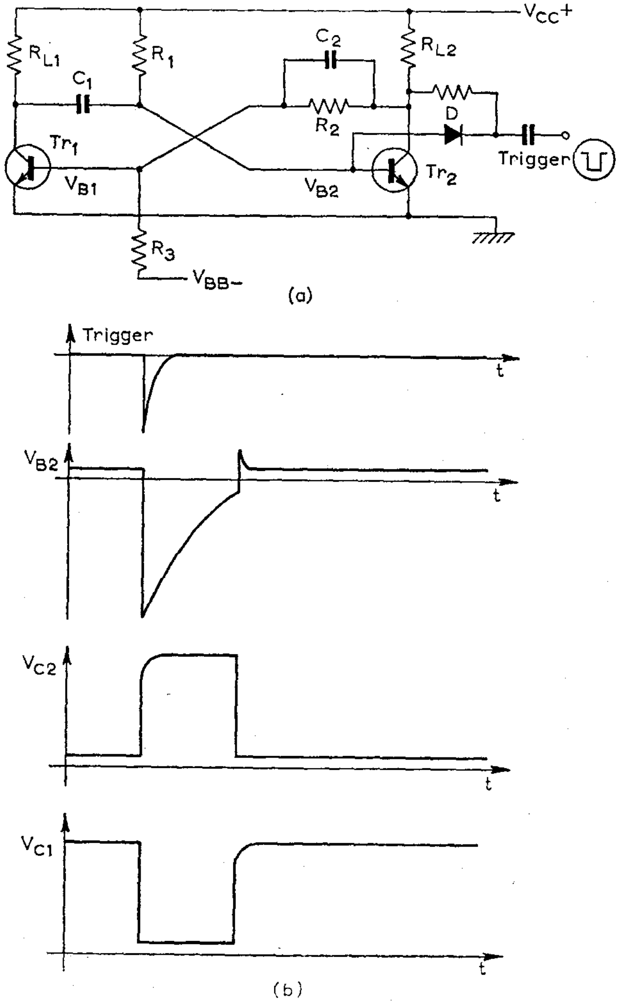


Fig. 23.15 MONOSTABLE MULTIVIBRATOR

to maintain Tr_2 in the cut-off condition. This precipitate amplification ends with Tr_1 on and Tr_2 off, so that $V_{C1} \approx 0$; $V_{C2} \approx V_{CC}$ (neglecting the voltage drop across R_{L2} due to the resistor chain current); $V_{B1} \approx 0$; and $V_{B2} = -V_{CC}$.

The quasi-stable state follows, as C_1 charges towards V_{CC} (time constant, C_1R_1), with $V_{C1} \approx 0$ and $V_{C2} \approx V_{CC}$. The duration of this state is the time required for V_{B2} to rise from $-V_{CC}$ to zero, when aiming at $+V_{CC}$; i.e. it is given by eqn. (23.21) as

$$\tau = C_1R_1 \log_e 2$$

As soon as V_{B2} reaches zero, Tr_2 starts conducting, and a further stage of precipitate amplification occurs to change Tr_1 to "off" and Tr_2 to "on." The circuit is then back in its stable state and awaits a further trigger pulse before again switching over. The waveforms are shown in Fig. 23.15(*b*).

Base leakage current in Tr_2 causes the same effect as in the astable circuit, and results in some temperature dependence of the output pulse-width.

EXAMPLE 23.2 In the monostable circuit of Fig. 23.15, $R_{L1} = R_{L2} = 1{\cdot}8\,\text{k}\Omega$; $C_1 = C_2 = 10{,}000\,\text{pF}$; $R_1 = 22\,\text{k}\Omega$; $R_2 = 15\,\text{k}\Omega$; $R_3 = 120\,\text{k}\Omega$. The supply voltage, V_{CC}, is 6V and the bias voltage is -6V. Verify the stable-state currents and voltages, and show that Tr_1 is bottomed in the quasi-stable state. The transistors used have $h_{FE} = 70$.

In the quiescent state Tr_1 is off and Tr_2 is on. Assuming that Tr_2 is saturated (bottomed) and that the voltage drop across it is then negligible, the collector current in Tr_2 will be

$$I_{C2} = \frac{V_{CC}}{R_{L2}} = \frac{6}{1{,}800}\text{A} = 3{\cdot}3\,\text{mA}$$

The base current is then

$$I_{B2} \approx \frac{V_{CC}}{R_1} = \frac{6}{22{,}000}\text{A} = 0{\cdot}27\,\text{mA}$$

(neglecting the base–emitter voltage).

Since $I_{B2} > I_{C2}/h_{FE}$, the transistor is saturated.

The base voltage of Tr_1 is determined by the voltage-dividing action of R_2 and R_3 across V_{C2} and $-V_{BB}$, i.e.

$$V_{B1} = \frac{-V_{BB}R_2}{R_2 + R_3} = -\frac{6 \times 15}{135} = -0{\cdot}7\,\text{V} \quad (\text{since } V_{C2} \approx 0)$$

Hence the base of Tr_1 is at a lower potential than the emitter, and so Tr_1 must be cut off. Note also that the drain of R_2 and R_3 is approximately

$$I = \frac{V_{BB}}{R_2 + R_3} \approx 45\,\mu\text{A}$$

which gives a negligible voltage drop across R_{L2}.

The width of the output pulse when triggering occurs is

$$\tau = C_1R_1 \log_e 2 = 10^{-8} \times 22 \times 10^3 \times 0{\cdot}69\,\text{s} = 152\,\mu\text{s}$$

In the quasi-stable state Tr_2 is off. *If* there were no base current in Tr_1 the voltage at the mid-point of R_2 and R_3 would be

$$-V_{BB} + \frac{V_{CC}R_3}{R_{L2} + R_2 + R_3} = -6 + \frac{12 \times 120 \times 10^3}{137 \times 10^3} = +4{\cdot}5\,\text{V}$$

Hence the base must be forward biased, and neglecting the base–emitter voltage drop, the actual base voltage is zero. The actual current through R_2 is therefore

$$I_{R2} = \frac{6}{1{,}800 + 15{,}000}\,\text{A} = 0{\cdot}357\,\text{mA}$$

The current through R_3 is, similarly, $6/120{,}000\,\text{A} = 0{\cdot}05\,\text{mA}$, so that the base current is $0{\cdot}375 - 0{\cdot}05 = 0{\cdot}325\,\text{mA}$. Since this is greater than I_{C1}/h_{FE} the transistor is bottomed.

23.11 The Bistable Multivibrator

The *bistable multivibrator* has two stable states, since either transistor may conduct and in doing so cuts the other one off. The circuit is shown in Fig. 23.16(*a*). Each transistor is d.c. coupled to the other. The capacitors C_1 and C_2 are called "speed-up capacitors" and act simply to ensure precipitate amplification during the transition from one stable state to the other.

Suppose that Tr_1 is on. The voltage-divider resistor chain is designed so that the low collector voltage of the "on" transistor is sufficient to keep the other transistor off under steady-state conditions. Hence if Tr_1 is on, Tr_2 is held off. If a negative trigger pulse is now routed through diode D_1 to the base of Tr_1, the base potential falls and Tr_1 takes less current. The collector potential of Tr_1 rises, and hence the base potential of Tr_2 rises. Tr_2 starts conducting and precipitate amplification then follows to turn Tr_2 fully on and to cut Tr_1 off. The circuit remains in this state until a negative reset pulse is applied through D_2 to the base of Tr_2. This will turn Tr_2 off and Tr_1 on. The relevant waveforms are shown in Fig. 23.16(*b*).

By interconnecting the trigger and reset pulse circuits, the input negative pulses may be routed by the diodes to whichever of the two bases is the more positive, so that successive input pulses will cause the bistable multivibrator to change state. The output at either collector will then be at half the pulse repetition frequency of the input pulse train. The circuit is then a *scale-of-two* counter or a *divide-by-two* circuit, which has many important applications in computers and static switching.

Leakage currents do not affect the operation of the bistable multivibrator, which is inherently stable and largely independent of temperature variations.

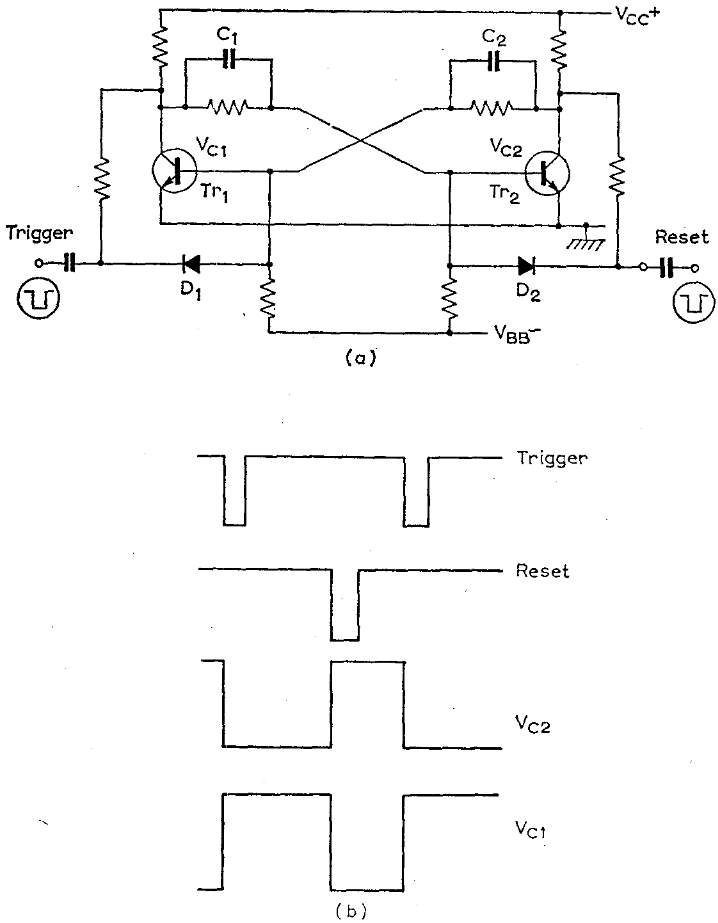


Fig. 23.16 BISTABLE MULTIVIBRATOR

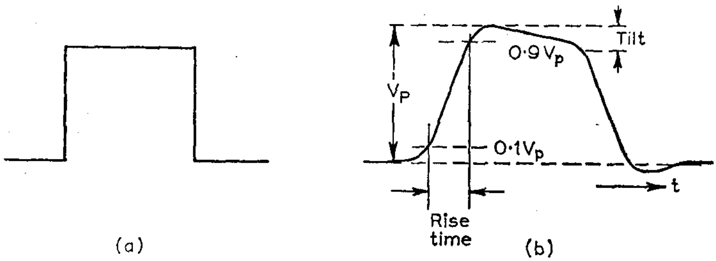


Fig. 23.17 IDEAL AND PRACTICAL PULSES
(*a*) Ideal (*b*) Practical

23.12 Rise Time and Tilt of Pulse Waveforms

An ideal pulse is rectangular in shape, as shown in Fig. 23.17(*a*). In practice it is not possible for a pulse to rise or fall instantaneously, and often the top of the pulse is not level. These effects are exaggerated at (*b*) for clarity.

When the leading edge of any practical pulse is observed on a cathode-ray oscilloscope with a fast enough time base, it is impossible to determine the exact instant at which the rise starts, since (as shown at (*b*)) the rise starts on an asymptotic curve. For the same reason it is also impossible to determine the time at which the peak is first reached. It has therefore become usual to specify the rise time of a pulse as the time taken for the pulse to grow from 0·1 to 0·9 of its peak value.

The tilt of a pulse waveform is a measure of the slope of the top of the pulse. The per-unit tilt is the difference between the peak value and the value at the start of the trailing edge expressed as a fraction of the peak value.

23.13 Pulse Response of a Low-pass *CR* Network

The voltage transfer function of the low-pass circuit shown in Fig. 23.18 is

$$\frac{V_o}{V_i} = \frac{1/j\omega C}{(R + 1/j\omega C)} = \frac{1}{1 + j\omega CR}$$

If ω is small so that $\omega \ll 1/CR$, the voltage transfer function is unity. V_o/V_i will fall by 3 dB from this value when $\omega CR = 1$. The frequency, f_H, at which this occurs may be taken as a measure of the upper limit of frequency of the circuit, where

$$f_H = \frac{1}{2\pi CR} \tag{23.24}$$

Hence, for any frequency f,

$$\frac{V_o}{V_i} = \frac{1}{1 + j\dfrac{f}{f_H}} \tag{23.25}$$

When a pulse waveform is applied to the circuit, transient analysis shows that the output will be in the form of an exponential growth and decay. If the time constant is less than 0·2 of the pulse-width, the output pulse will eventually reach about 99 per cent of the peak value of the input as at (*b*) (a *CR* transient may be considered to reach

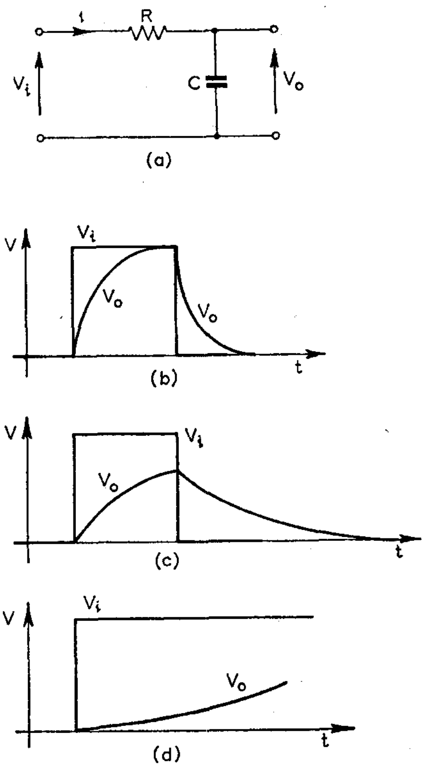


Fig. 23.18 PULSE RESPONSE OF LOW-PASS *RC* NETWORK

a steady value in five time constants). The expression for the leading edge of the output voltage waveform for a pulse of height V_i is

$$v_o = V_i(1 - e^{-t/CR})$$

The rise time is $(t_2 - t_1)$, where t_1 is the time for v_o to reach $0{\cdot}1V_i$, and t_2 is the time to reach $0{\cdot}9V_i$. Thus

$$0{\cdot}1V_i = V_i(1 - e^{-t_1/CR}) \qquad \text{or} \qquad t_1 = 0{\cdot}1CR$$

and

$$0{\cdot}9V_i = V_i(1 - e^{-t_2/CR}) \qquad \text{or} \qquad t_2 = 2{\cdot}3CR$$

so that

$$\text{Rise time} = t_1 - t_2 = 2{\cdot}2CR = \frac{2{\cdot}2}{2\pi f_H} \quad \text{seconds} \tag{23.26}$$

On the trailing edge the output decays according to

$$v_o = V_i\,\mathrm{e}^{-t/CR)}$$

where t is reckoned from the instant when the trailing edge of the input pulse commences.

If the pulse width is less than five time constants, the output pulse will never reach the peak of the input. This condition is shown in Fig. 23.18(c). The decay will, however, last at least five time constants, unless the next pulse arrives before this.

If the time constant, CR, is very large, the output of the circuit will be approximately equal to the time integral of the input for integrating times which are small compared to CR. Thus, for the circuit of Fig. 23.18(a), the instantaneous output voltage, v_o, is given by

$$v_o = \frac{q}{C}$$

where q is the instantaneous charge on C. Therefore

$$\frac{dv_o}{dt} = \frac{i}{C}$$

and

$$V_i = iR + v_o = CR\frac{dv_o}{dt} + v_o$$

If CR is large enough, so that $v_o \ll V_i$, then

$$v_o = \frac{1}{CR}\int_0^t V_i\,dt \tag{23.27}$$

The disadvantage of this form of integrating circuit is that the output voltage is very small. The output of such a circuit with a step input is shown at (d), and is simply the initial part of an exponential growth curve of long time constant.

23.14 Pulse Response of a High-pass *CR* Network

In the high-pass CR circuit shown in Fig. 23.19(a), the a.c. voltage transfer function is

$$\frac{V_o}{V_i} = \frac{R}{R + 1/j\omega C} = \frac{1}{1 + 1/j\omega CR}$$

For very high frequencies this ratio is unity. It falls 3 dB below unity at a frequency, f_L, for which

$$2\pi f_L = \frac{1}{CR} \tag{23.28}$$

The voltage transfer function can hence be written for any frequency, f, as

$$\frac{V_o}{V_i} = \frac{1}{1 - j\dfrac{f_L}{f}} \tag{23.29}$$

When a pulse waveform is applied to this circuit, on the rising edge the change of input causes the output voltage to rise instantaneously by the same amount, since the voltage across C cannot change

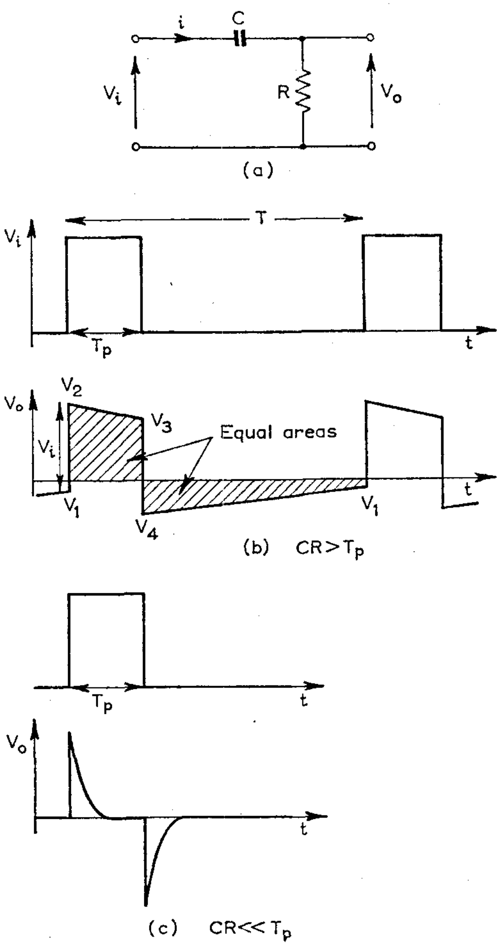


Fig. 23.19 PULSE RESPONSE OF HIGH-PASS *RC* NETWORK

instantaneously. C charges through R, and hence the output voltage falls exponentially towards zero. The fall in voltage stops when the negative edge of the input pulse arrives. A similar process then occurs at this falling edge as shown in Fig. 23.19(*b*). Since C acts as a d.c. block, the mean output voltage must be zero, so that the shaded areas above and below the time axis in the diagram must be equal.

The per-unit tilt of the output pulse can be found readily if CR is very much greater than the pulse-width, T_p. In this case $T_p/CR \ll 1$ and the exponent in $e^{-T_p/CR}$ is approximately equal to $1 - T_p/CR$. Also, if the output voltage rises from V_1 to V_2, then

$$V_2 = V_1 + V_i$$

where V_1 will be negative, and

$$V_3 = V_2\, e^{-T_p/CR} \approx V_2\left(1 - \frac{T_p}{CR}\right)$$

so that the per-unit tilt is

$$\frac{V_2 - V_3}{V_i} \approx \frac{V_2 - V_2(1 - T_p/CR)}{V_i} = \frac{V_2 T_p}{V_i CR} \qquad (23.30)$$

If V_1 is small, $V_i \approx V_2$ so that

$$\text{Per-unit tilt} = \frac{T_p}{CR} = 2\pi f_L T_p \qquad (23.30a)$$

For a CR time constant which is less than 0·2 of the pulse-width, the capacitor C can completely charge before the negative-going edge of the pulse arrives, and the output voltage will consist of positive- and negative-going spikes as shown at (*c*).

If the output voltage, v_o, is a very small fraction of the input voltage, v_i, then to a good approximation all the input voltage appears across the capacitor C so that the alternating current in C is

$$i \approx C\frac{dv_i}{dt}$$

Thus, if there is no external loading,

$$v_o = iR \approx CR\frac{dv_i}{dt} \qquad (23.31)$$

i.e. the circuit acts as approximate differentiating circuit. Note, however, that if the input is an ideal pulse with zero rise time, the derivative is infinite, and this cannot be obtained from the given circuit. In other words, the circuit acts as an approximate differentiator if (*a*) the CR time constant is small compared to the input rise time, (*b*) the output voltage is very small compared to the input, and (*c*) the rise time of the input is finite.

EXAMPLE 23.3 A pulse waveform of peak value 10 V is applied to the high-pass *CR* network of Fig. 23.19(*a*). The pulse-width is 100 μs and the *pulse repetition frequency* (p.r.f.) is 2,000 pulses/second. If $C = 0{\cdot}1\ \mu$F and $R = 100\,\text{k}\Omega$, determine the maximum positive and negative output voltage and the per-unit tilt.

Since the p.r.f. is 2,000 the pulse period is

$$T = \frac{1}{\text{p.r.f.}} = \frac{1}{2{,}000}\,\text{s} = 500\,\mu\text{s}$$

The output waveform is that shown in Fig. 23.19(*b*), where

$$V_2 = V_1 + V_i \qquad \text{(i)}$$

$$V_3 = V_2\,\mathrm{e}^{-T_p/CR} \qquad \text{(ii)}$$

$$V_4 = V_3 - V_i$$

and

$$V_1 = V_4\,\mathrm{e}^{-(T-T_p)/CR} \qquad \text{(iii)}$$

For equal areas above and below the axis,

$$\int_0^{T_p} V_2\,\mathrm{e}^{-t/CR}\,dt + \int_0^{T-T_p} V_4\,\mathrm{e}^{-t/CR}\,dt = 0$$

Hence

$$-CRV_2(\mathrm{e}^{-T_p/CR} - 1) = CRV_4(\mathrm{e}^{-(T-T_p)/CR} - 1)$$

Substituting the given numerical values,

$$-V_2(\mathrm{e}^{-0\cdot01} - 1) = V_4(\mathrm{e}^{-0\cdot04} - 1)$$

and, to a close approximation,

$$-V_2(1 - 0{\cdot}01 - 1) = V_4(1 - 0{\cdot}04 - 1) \quad \text{or} \quad V_2 = -4V_4$$

From eqn. (iii),

$$V_1 \approx V_4(1 - 0{\cdot}04) = 0{\cdot}96V_4$$

Solving these equations for V_1 and V_2 gives $V_1 = -1{\cdot}9\,\text{V}$ and $V_2 = 8{\cdot}1\,\text{V}$.

The per-unit tilt is found by substituting in eqn. (23.30):

$$\text{Per-unit tilt} \approx \frac{V_2T_p}{V_iCR} = \frac{8{\cdot}1 \times 10^{-4}}{10 \times 10^{-7} \times 10^5} = \underline{\underline{0{\cdot}008}}$$

23.15 Preservation of Pulse Waveform

When a pulse waveform is applied to an *RC*-coupled amplifier the output will in general have a longer rise time than the input and will also have tilt. The Fourier analysis of a pulse waveform shows that it contains a large number of harmonics, and hence it is essential that the amplifier should have a wide bandwidth if the output is not to be distorted.

The high-frequency response of a single-stage amplifier determines the rise time of the pulse output. This is equivalent to the response of a low-pass *RC* network, and eqn. (23.25) gives the relation between rise time and the upper cut-off frequency, f_H, as

$$\text{Rise time} = \frac{2{\cdot}2}{2\pi f_H}$$

The low-frequency response of a single-stage amplifier determines the per-unit tilt of the output, given by eqn. (23.30*a*) for a pulse-width, T_p, as

$$\text{Per-unit tilt} \approx 2\pi f_L T_p$$

provided that the pulse period, T, is long enough to ensure that $(T - T_p) > 5CR$.

Hence for best response f_H must be as high as possible and f_L as low as possible.

If a pulse waveform is applied to a circuit which has appreciable input capacitance, an increase in the rise time will occur unless the source is of zero impedance.

This is of particular importance if pulses are to be observed on a cathode-ray oscilloscope, since the trace obtained may not then correspond to the actual input, owing to the input capacitance of the oscilloscope. This effect may be partly overcome by the use of a ***compensated attenuator***.

Consider the simple attenuator shown in Fig. 23.20(*a*). Using the normal transient analysis for a step input, it may be readily shown that the output for a step input is given by

$$V_o = \frac{V_i R_2}{(R_1 + R_2)}(1 - \mathrm{e}^{-t/CR_p})$$

where $R_p = R_1R_2/(R_1 + R_2)$. Hence the output rise time is $2{\cdot}2CR_p$ seconds.

If the attenuator is compensated by the addition of capacitor C_1, as shown at (*b*), then for r.m.s. voltages,

$$\frac{V_i}{V_o} = 1 + \frac{Y_p}{Y_s} = 1 + \frac{Z_s}{Z_p}$$

where $Y_s = 1/R_1 + j\omega C_1$ and $Y_p = 1/R_2 + j\omega C_2$. Hence

$$\frac{V_i}{V_o} = 1 + \frac{R_1(1 + j\omega C_2R_2)}{R_2(1 + j\omega C_1R_1)}$$

If $C_1R_1 = C_2R_2$ this expression becomes independent of frequency, and will therefore give no distortion of the pulse waveform. For this case,

$$\frac{V_o}{V_i} = \frac{1}{1 + \dfrac{R_1}{R_2}} = \frac{R_2}{R_1 + R_2} \tag{23.32}$$

If the attenuator is undercompensated ($C_1R_1 < C_2R_2$), there will be degradation of the rise time. If it is overcompensated ($C_1R_1 > C_2R_2$), there will be an overshoot on the output waveform. These effects are shown at (*b*).

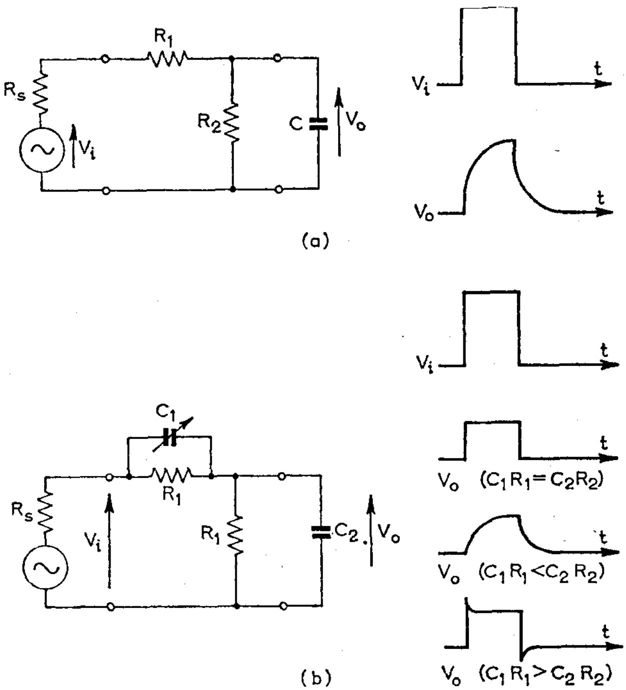


Fig. 23.20 PULSE RESPONSE OF ATTENUATOR
(*a*) Uncompensated (*b*) Compensated

23.16 Sawtooth Generator

In many applications a voltage that increases linearly with time is required (e.g. for the deflexion of the spot on a cathode-ray tube; in television scanning circuits; in radar displays; etc.). Normally a slow *sweep* followed by a very rapid *flyback* is wanted, such as the sawtooth waveform shown in Fig. 23.21(*a*). This waveform may be

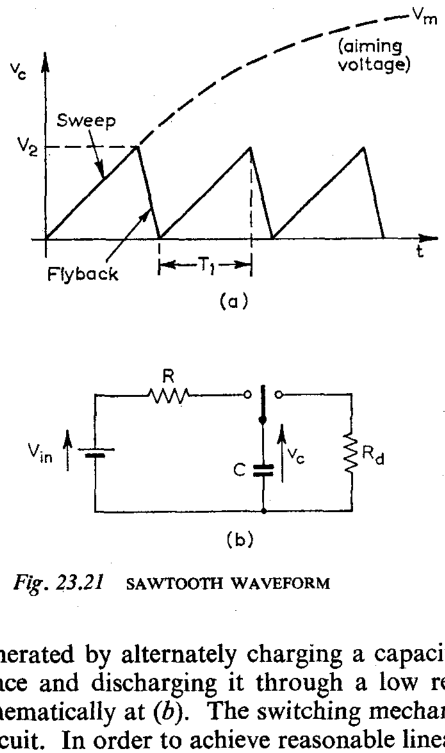


Fig. 23.21 SAWTOOTH WAVEFORM

generated by alternately charging a capacitor through a high resistance and discharging it through a low resistance. This is shown schematically at (*b*). The switching mechanism will be an electronic circuit. In order to achieve reasonable linearity, the flyback must be initiated when the capacitor voltage has reached a value V_2 which is a small fraction of the *aiming voltage*, V_m.

From transient theory, the instantaneous capacitor voltage, v_c, is

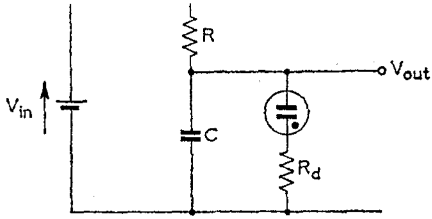
$$
\begin{aligned}
v_c &= V_m(1 - e^{-t/CR}) \\
&= V_m\left(1 - 1 + \frac{t}{CR} - \tfrac{1}{2}\left(\frac{t}{CR}\right)^2 + \ldots\right) \\
&= \frac{V_m t}{CR}\left(1 - \tfrac{1}{2}\frac{t}{CR} + \ldots\right)
\end{aligned}
$$

For true linearity the capacitor voltage would be $V_m t/CR$, so that the term $\frac{1}{2}t/CR$ represents a fractional error. Since the neglected terms in the exponential series alternate in sign, and (provided that $t/CR \ll 1$) become smaller, the actual fractional error will always be less than $\frac{1}{2}t/CR$. If flyback occurs at time T_1 after the commencement of the sweep, when the capacitor voltage is V_2, then

$$V_2 \approx \frac{V_m T_1}{CR}\left(1 - \frac{1}{2}\frac{T_1}{CR}\right) \tag{23.33}$$

Hence the error in linearity is small provided that T_1 is a small fraction of the CR time constant of the charging circuit. If a constant-current source is used to charge C, there will be no error in linearity, since effectively the aiming voltage will then be infinite. The rate of growth of voltage can be varied by altering V_m, C or R.

An extremely simple sawtooth generator using a neon discharge tube is shown in Fig. 23.22(*a*). The capacitor, C, charges through



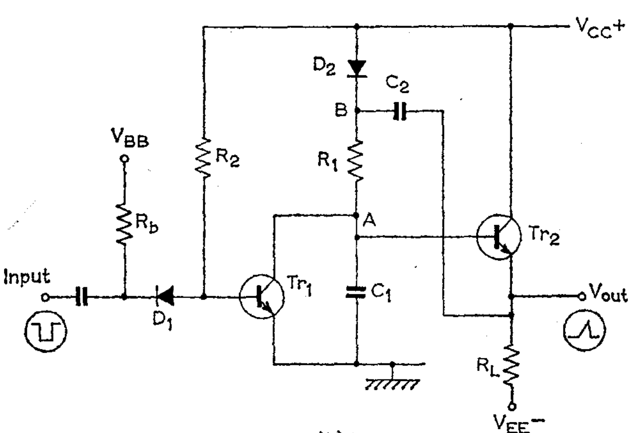


Fig. 23.22 SIMPLE SAWTOOTH GENERATORS
(*a*) Neon (*b*) Bootstrap integrator

R until the breakdown voltage of the neon is reached, when it discharges rapidly through the protective resistance in the neon circuit. The neon is extinguished and the process repeats. The voltage rise will, of course, start from the extinction voltage of the neon and not from zero. For good linearity, V_m requires to be very large, and the sweep time, T_1, must be small compared to CR.

A more sophisticated circuit is the bootstrap integrator shown in Fig. 23.22(*b*). This relies on the fact that the integral of a square wave is a linearly rising curve. In principle, capacitor C_1 charges slowly through R_1 when transistor Tr_1 is non-conducting, and discharges rapidly through Tr_1 when this transistor is switched into conduction by the rising edge of the input square wave. The function of transistor Tr_2 is to give a constant voltage across the charging resistor R_1 (and hence to give a constant charging current for C_1). Since Tr_2 is in the emitter-follower connexion, it has a low-impedance output, so that the output waveform is very little affected by any load at the output terminals.

With no input, Tr_1 conducts, so that the voltage at A is almost zero (and hence C_1 is discharged), while the voltage across R_1 is almost V_{CC}. The output voltage is also almost zero, owing to emitter-follower action.

When the input voltage goes negative, Tr_1 becomes cut off, and C_1 starts to charge through R_1. The voltage at A rises, and owing to emitter-follower action, so does the output voltage. The output voltage rise is fed back through the large capacitor C_2 to point B, so that, if the gain of the emitter follower is unity, the voltages at A and B will both increase by the same amount and the charging current will remain constant. The rising voltage at B will cut off diode D_2. This linear rise continues as long as Tr_1 is cut off—the output of Tr_2 may be thought of as lifting itself during this time by its own bootstraps. The linearity depends on how closely the gain of the emitter follower approaches unity, and on the fact that $C_2 \gg C_1$.

PROBLEMS

23.1 The open loop gain of a feedback amplifier is given by

$$\beta A_{vo} = -10/(1 + jf(10^7)^3$$

Determine whether the amplifier will be unstable. Determine the frequency at which the amplifier will become unstable if the numerator becomes -5.

Ans. 17 MHz.

23.2 For the Colpitts oscillator shown in Fig. 23.4 the value of L is 20 μH and the two tuned circuit capacitors each have a value of 500 pF. Determine from the generalized conditions for oscillation the frequency of oscillation and the minimum value of h_{fe} required for the transistor.

Ans. 2·25 MHz; 1.

23.3 Repeat Problem 23.2 for the Hartley oscillator of Fig. 23.5 given that $L_1 = 500\,\mu$H, $L_2 = 200\mu$H and $C = 1{,}000$ pF, and assuming no mutual coupling between the coils.

Ans. 190 kHz; 2·5.

23.4 Determine the figure of merit

$$\frac{d}{d\omega}(X_{ce} + X_{cb} + X_{be})$$

for frequency stability of the oscillator of Problem 23.3.

Ans. 40×10^{-6}.

23.5 Repeat Problem 23.4 for the Hartley oscillator of Problem 23.3.

Ans. $1{\cdot}38 \times 10^{-3}$.

23.6 In the tuned-collector oscillator of Fig. 23.6 the tuned circuit has a dynamic resistance of R_0. Show that the circuit will oscillate at a frequency given by $\omega_0 \approx 1/\sqrt{(L_1C)}$ when

$$n(h_{ie} + R)(1/R_0 + h_{oe}) + 1/n = h_{fe}$$

where R is the resistance of R_1 and R_2 in parallel. The transformer has a turns ratio of n:1 and a coupling coefficient of unity.

23.7 Sketch the circuit and describe the operation of a free-running (astable) multivibrator.

In a symmetrical multivibrator the collector resistors are 1·8 kΩ and the coupling capacitors are 0·01 μF. The transistors are *p–n–p* and the base resistors of 27 kΩ are connected to the negative supply rail. Calculate from first principles the pulse repetition frequency. Also determine the rise time of the output pulses.

Ans. 2,650 pulses/second; 39·5 μs.

23.8 Explain the operation of a monostable transistor multivibrator.

For the circuit shown in Fig. 23.30 determine the output pulse-width and the

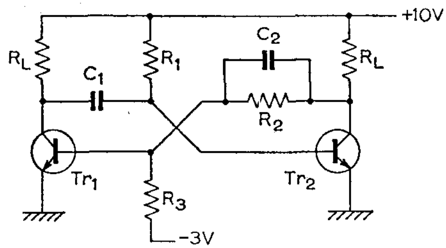


Fig. 23.23

base and collector voltages and currents in the quiescent state. $R_L = 4{\cdot}7\,\text{k}\Omega$, $C_1 = C_2 = 0{\cdot}001\,\mu$F; $R_1 = R_2 = 100\,\text{k}\Omega$; $R_3 = 47\,\text{k}\Omega$. Assume a base–emitter voltage of 0·7 V and negligible collector–emitter voltage at saturation.

Ans. 69 μs; $V_{C1} = 10$ V; $V_{B1} = -0{\cdot}96$ V; $V_{C2} = 0$; $V_{B2} = 0{\cdot}7$ V; $I_{C2} =$ 2·13 mA; $I_{B2} = 93\,\mu$A.

23.9 A train of positive pulses, of amplitude 20V, pulse-width 0·2ms and repetition frequency 1kHz, is applied to a high-pass *CR* network in which $C = 0{\cdot}01\,\mu$F and $R = 60\,\text{k}\Omega$. Sketch the output waveform and determine the maximum positive and negative values of the output voltage. (*H.N.C.*)

Ans. 18·2V; −7V.

23.10 A square-wave pulse train (1:1 mark-to-space ratio) is applied to a low-pass *RC* network. The input pulses have a peak-to-peak amplitude of 10V. If $R = 100\,\text{k}\Omega$, $C = 0{\cdot}1\,\mu$F and the pulse repetition frequency is 150 pulses per second, determine the peak-to-peak output voltage. Sketch the input and output waves.

Ans. 0·28V.

23.11 Explain the action of a compensated attenuator.

In a compensated attenuator as shown in Fig. 23.20(*b*), $C_2 = 250$pF, $R_1 = 10\,\text{k}\Omega$, $R_2 = 15\,\text{k}\Omega$. Determine the correct value of C_1. If C_1 is 10 per cent high, find the fractional overshoot when a rectangular step is applied at the input.

Ans. 375pF; 0·054.

23.12 The symmetrical astable multivibrator shown in Fig. 23.24 uses *n–p–n* transistors. Determine the pulse repetition frequency if $R_L = 4{\cdot}7\,\text{k}\Omega$, $R = 100\,\text{k}\Omega$ and $C = 500$pF.

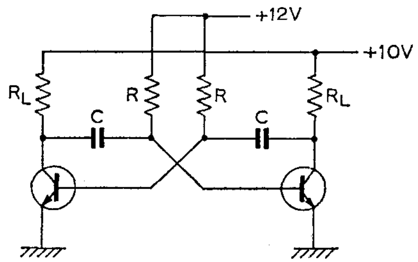


Fig. 23.24

Ans. 16,000 pulses/second.

Chapter 24

POWER AMPLIFIERS

Power amplifiers can be generally classed into two groups:

(*a*) Those intended to deliver power over a wide frequency range.
(*b*) Those intended to deliver power at one particular frequency only.

The first type is principally used for audio-frequency power (30-30,000 Hz) and may be either transformer coupled or direct coupled; the second type is principally used for radio-frequency power (100 kHz upwards) and is generally a tuned amplifier.

The power amplifier stage in an equipment is generally the final stage which must drive a given load, e.g. a mechanical indicator, a loudspeaker, cathode-ray-tube magnetic deflexion coils, a transmitting aerial of a communication network. Power amplifiers are also used in regulated power supplies, chopper amplifiers, motor controllers, etc. In general the load will demand a certain power from the amplifier and the output stage must be capable of dealing with this power. The design of an equipment often commences with the output stage and with its required output power. The preceding stage is then designed to give the input required to drive the output stage. Amplifiers are then added until there is the required amplification between the initial input signal and the output power amplifier.

In a power amplifier a large portion of the active-device characteristic is used (i.e. there is large alternating input voltage or current) to make efficient use of the device. This gives rise to non-linear distortion in the device due to the curvature of the characteristics.

It is important to keep the non-linear distortion within tolerable limits.

The power transistor is rapidly replacing the valve as the active device in most power amplifier applications. Triodes and tetrodes, however, are still used as output valves in broadcast transmitters, and special thermionic devices are still required at frequencies in the microwave range, where valves (Klystrons) having continuous output powers of over 50 kW at 10 GHz are used. It is not proposed to discuss these special devices in this book.

Power transistors are now available for operation at voltages in excess of 1 kV, or at currents in excess of 250 A (at low voltage). They can be obtained in complementary *n–p–n* and *p–n–p* configurations, which in many applications, gives them further advantages over valves, besides the great advantage of requiring no heater supply. The value of f_T for power transistors has been continuously increased over the past few years due to improvements in manufacturing technology, and for many nominally "low-frequency" power devices f_T exceeds 15 MHz.

The construction of the power transistor may differ mechanically from that of the small-signal transistor, in that the collector is usually directly connected to the metal outside case. This facilitates direct connexion (thermally and electrically) to a metal heat sink and enables the highest possible power rating to be achieved. At elevated temperatures care must be taken to observe the requirements of the derating curves supplied by the manufacturers.

24.1 Class A Power Amplifier

A Class A amplifier is one in which the current flows in the output valve or transistor at all instants throughout the cycle—other types will be described in following sections. A Class A valve amplifier is shown in Fig. 24.1(*a*). A triode, a pentode or a beam tetrode may be used—the triode has less distortion, but requires a larger alternating input grid voltage. For small powers (less than, say, 25 W) a pentode is more common, but for higher powers a triode or tetrode becomes essential as it is difficult to dissipate any appreciable power from the screen grid of a pentode.

A transistor power output stage is shown at (*c*). In transistor stages particular attention must be paid to heat dissipation (to prevent thermal runaway of the transistor), and to voltage breakdown.

Since power amplifiers may operate over wide ranges of the active-device characteristics, it is no longer permissible to use the small-signal linear equivalent circuit—actual device characteristic curves must be used. The load resistor, R, is usually connected through a step-down transformer in order to achieve maximum power output. This has the additional advantage that it prevents the flow of direct current through the load.

Typical output characteristics are shown at (*b*) and (*d*). In both cases the resistance of the transformer primary has negligible effect on the static load line, which is therefore the vertical ordinate from the supply voltage V_S. The quiescent (no signal) operating point is determined by the bias conditions, and is chosen to give equal

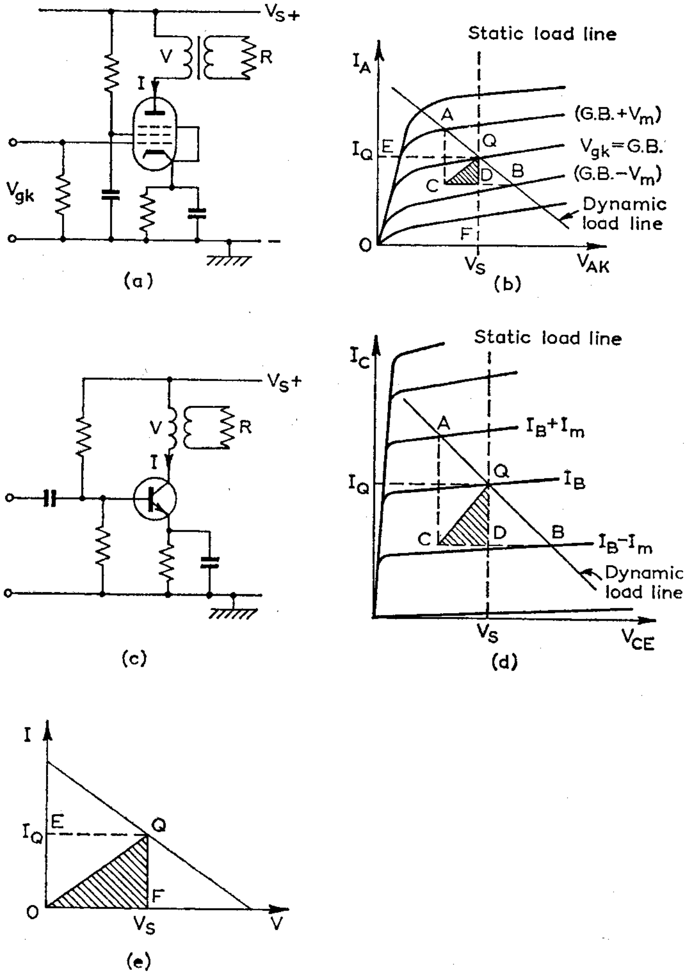


Fig. 24.1 CLASS A AMPLIFICATION

excursions along the dynamic load line above and below Q. For signal currents the load resistance referred to the primary of the output transformer is $k_t^2 R$, where k_t $(=N_1/N_2)$ is the turns ratio. The dynamic load line is drawn through Q with a slope of $-1/k_t^2 R$, assuming that the transformer winding resistance is negligible (this may not always be true, however, particularly for small transformers).

For an alternating input the operating point moves along the dynamic load line between A and B, and the variation of the primary current of the output transformer is then proportional to AC, and the corresponding voltage variation, to CB.

Assuming sinusoidal variations, the r.m.s. alternating component of the transformer primary current, I, is proportional to AC/$2\sqrt{2}$. Also the r.m.s. alternating component of the transformer primary voltage, V, is proportional to BC/$2\sqrt{2}$. Therefore

$$\text{Alternating power developed} = VI \propto \frac{\mathrm{AC}\,.\,\mathrm{BC}}{8} \tag{24.1}$$

By geometry, this alternating power is represented by one-quarter of the area ABC, or, if distortion is neglected, by the shaded area CQD.

The power drawn from the supply is the supply voltage times the mean direct current, i.e. VI_Q. This is represented by the rectangle OEQF shown at (*b*).

$$\text{Efficiency of amplifier} = \frac{\text{Signal power output}}{\text{Power from supply}} = \frac{VI}{VI_Q}$$

$$= \frac{\frac{1}{4} \times \text{area ABC}}{\text{area OEQF}} \approx \frac{\text{area CQD}}{\text{area OEQF}} \tag{24.2}$$

For the cases depicted the efficiency is less than $12\frac{1}{2}$ per cent.

$$\text{Power dissipated in active device} = V_S I_Q - VI \tag{24.3}$$

If there is no signal input, the entire supply power, $V_S I_Q$, will be dissipated in the valve or transistor, and the quiescent operating point must be chosen so that this power does not exceed the permissible limit for the device.

Due to the low efficiency and the possibility of thermal runaway in the absence of an input signal, transistors are not commonly used in Class A power operation.

24.2 Maximum Efficiency in Class A

As the signal input increases, the output power and efficiency will increase. Neglecting distortion, and assuming that the output characteristics extend to the current and voltage axes in Figs. 24.1(*b*) and (*d*), maximum power output is obtained if the entire dynamic load line is used as at (*e*). Then

$$\text{Efficiency} = \frac{\text{area OQF}}{\text{area OEQF}} = 0{\cdot}5 \tag{24.4}$$

This is the theoretical maximum efficiency of any Class A amplifier. In valve amplifiers it is unusual for efficiencies to exceed 0·25 owing to the large amount of distortion produced with large grid-voltage swings. Some transistor amplifiers achieve a value much nearer the theoretical maximum.

Note that, if the input signal is a square wave or a pulse waveform, the maximum theoretical efficiency is increased to 100 per cent. Pulse operation of a power amplifier has been called Class D operation.*

EXAMPLE 24.1 The anode characteristics of a triode are shown in Fig. 24.2. The triode is to be used as a power amplifier valve feeding a load resistance of 50Ω through a 10:1 step-down transformer. The supply voltage, V_B, is 530V. The maximum power dissipation at the anode of the triode is 20W.

(i) Draw a line across the triode characteristics representing the maximum anode dissipation.

(ii) Choose a suitable grid-bias voltage and determine the corresponding cathode bias resistor.

(iii) Draw the dynamic load line.

(iv) Determine the power output and efficiency for peak grid voltages of (*a*) 20V, and (*b*) 40V. The transformer may be assumed perfect.

(i) For a power of 20W at the anode, $V_A I_A = 20$, where I_A is the anode current in amperes. The line representing 20W on the anode characteristics is the rectangular hyperbola given by the above equation and detailed in the following table:

V_A (volts)	200	300	400	500	600	800
I_A (mA)	100	66·7	50	40	33·3	25

* An oscillator is said to operate in Class D when the active device is switched on and off at successive angles of current flow of 180°. The maximum theoretical efficiency, like that of Class C, is 100 per cent. The power is limited only by the permissible peak current and voltage, and by the switching speed obtainable, as distinct from Class C, where the power tends to zero as the efficiency approaches 100 per cent.

Class D operation may be applied to amplifiers. The angle of current flow in the on (or off) condition is modulated by the signal to give a pulse-width-modulated output from which the signal is recovered by means of a low-pass filter.

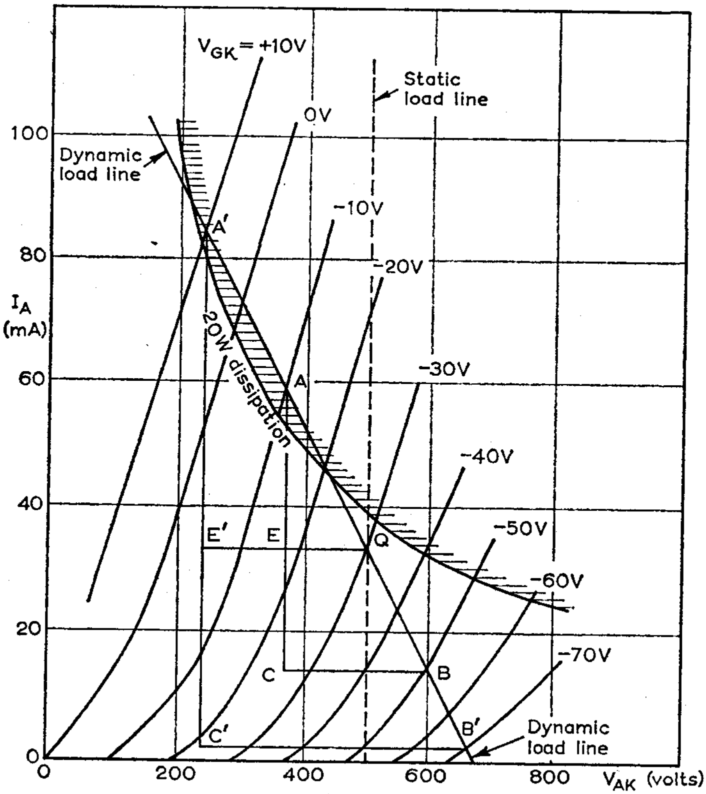


Fig. 24.2

(ii) To give operation in the linear region without exceeding the permissible anode power, a grid bias of -30V will be seen from Fig. 24.2 to be suitable with a supply voltage of 530V. The anode voltage will thus be $530 - 30 = 500$V. Q represents the quiescent operating point and quiescent current. Therefore $I_Q = 33$mA, and

$$\text{Cathode bias resistance} = \frac{V_{GK}}{I_Q} = \frac{30}{33} \times 1{,}000 = \underline{\underline{910\Omega}}$$

(iii) The impedance in the anode circuit to alternating currents is $k_t^2 R$, i.e. $10^2 \times 50 = 5{,}000\Omega$. (This is about twice r_a for the valve.) Thus the anode load line is a line through Q with a slope of $-1/5{,}000$. (Although at some points the load line is above the 20W permissible limit, the mean and quiescent power will still be less than 20W.)

(iv) (*a*) For a peak grid voltage of 20V the operation is between points A and B on the dynamic load line.

Peak-to-peak current swing = AC = 46 mA

Peak-to-peak voltage swing = BC = 222 V

$$\text{Power output (assuming sine waves)} = \frac{46 \times 222}{8 \times 1{,}000} = \underline{\underline{1 \cdot 3\,\text{W}}}$$

$$\text{Power input} \approx V_B I_Q = \frac{530 \times 33}{1{,}000} = 17 \cdot 5\,\text{W}$$

$$\text{Efficiency} = \frac{1 \cdot 3}{17 \cdot 5} \times 100 = 7 \cdot 5 \text{ per cent}$$

(*b*) For a peak grid voltage of 40 V the operation is between A′ and B′ on the dynamic load line. Thus

$$\text{Power output (assuming sine waves)} = \frac{84 \times 420}{8{,}000} = \underline{\underline{4 \cdot 4\,\text{W}}}$$

and

$$\text{Efficiency} = \frac{4 \cdot 4}{17 \cdot 5} \times 100 = \underline{\underline{25}} \text{ per cent}$$

The following point may be noted: when the power is calculated from the peak-to-peak current and voltage swings, the assumption of sine waves gives negligible error. Indeed it gives the fundamental output power without error if only even-order curvature is present.

24.3 Choice of Transformer Ratio in Class A

When a power amplifier is designed the following will normally be specified (*a*) load resistance; (*b*) load power (which, after assuming a reasonable anode or collector efficiency, will determine the rating of the valve or transistor used); (*c*) permissible distortion. Sometimes the supply voltage is also specified, but more usually this is chosen to suit the active devices. A suitable output transformer ratio must then be selected. From the output characteristics of the active device the required bias and signal input are determined, and power output and distortion may then be checked graphically.

If the output characteristics are reasonably linear, and the excursion of the input is not too great, then the maximum power transfer theorem shows that the maximum power transfer occurs when the reflected load resistance, $k_t^2 R$, is equal to the slope of the output characteristics (i.e. the output impedance of the active device).

In practice, in order to obtain the best utilization of the active device, it is usual to operate into the non-linear region of the characteristics as shown in Fig. 24.3. The output characteristics of transistors, tetrodes and pentodes are generally closer together at both ends of the dynamic load line, and hence for a large output signal which extends into these regions there will be some flattening of both its positive and negative peaks. Such a flattening has been seen in Chapter 5 to introduce third-harmonic distortion. If only the negative or only the positive peaks are flattened the resulting distortion is mainly second harmonic—this is the most usual form of distortion in triode power amplifiers.

If the dynamic load line represents too high a reflected load impedance (e.g. the line marked (1) in Fig. 24.3), intolerable distortion occurs at the positive peak of the signal. On the other hand, if the reflected impedance is too low (line (2)), the maximum power output is small. By choosing the transformer ratio so that the dynamic load line passes through the knee of the output characteristic

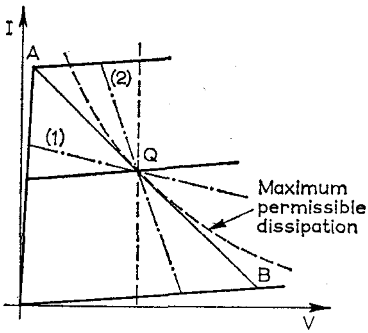


Fig. 24.3 PERTAINING TO THE CHOICE OF A TURNS RATIO TO GIVE A DYNAMIC LOAD LINE THROUGH THE KNEE OF THE OUTPUT CHARACTERISTIC

as shown by line AB on the diagram, a high maximum power output with minimum distortion is achieved. With transistors this knee voltage may be quite small (less than 1 V), but in valve circuits it is generally several tens of volts. This is a further important reason why transistors are so much more efficient power amplifiers than valves, since they enable a much larger range of the characteristic to be used. In all cases care must be taken not to exceed the maximum power dissipation of the transistor or valve.

It can be shown that in triode amplifiers, limited to a negative grid voltage, the optimum ratio is that which makes $k_t^2R = 2r_a$.

24.4 Square-law Distortion

In many cases the relation between input and output signals (particularly in triode amplifiers) may be simply approximated by a square law. Thus let the relation between anode current, i_a, and grid voltage, v_g, for an output valve be given by

$$i_a = a_0 + a_1v_g + a_2v_g^2 \qquad (24.5)$$

Let the grid voltage be given by

$$v_g = V_{gb} + V_m \cos \omega t$$

i.e. a steady bias of V_{gb} in series with a sinusoidal signal. Then

$$
\begin{aligned}
i_a = {} & a_0 + a_1 V_{gb} + a_2 V_{gb}^2 && (1)\\
& + a_1 V_m \cos \omega t + a_2 \,.\, 2V_{gb} \,.\, V_m \cos \omega t && (2)\\
& + \tfrac{1}{2} a_2 V_m^2 && (3)\\
& + \tfrac{1}{2} a_2 V_m^2 \cos 2\omega t && (4)
\end{aligned}
$$

since $\cos^2 \omega t = \frac{1}{2}(1 + \cos 2\omega t)$.

Part 1 of the anode current is the normal quiescent current I_g for the grid bias V_{gb}.

Part 2 is the anode current variation at the frequency of the input, i.e. the desired anode current variation.

Part 3 is an extra direct current introduced owing to the curvature of the characteristics and proportional to the square of the alternating input voltage. It also represents the rectified signal which is characteristic of a square-law transfer relation.

Part 4 is an unwanted double-frequency, or second harmonic, component in the output waveform which is not in the input waveform. This is the result of non-linear distortion. It should be noted that the amplitude of the second-harmonic term is the same as the magnitude of the direct current due to the curvature.

The component and total current waveforms are shown in Fig. 24.4. It will be seen that from quiescent current to peak positive current corresponds to peak fundamental current plus twice peak

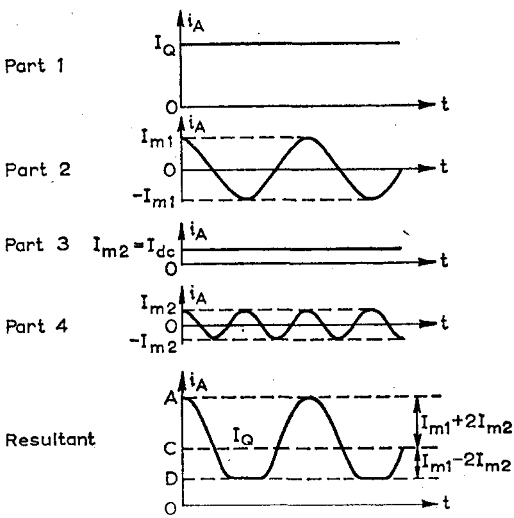


Fig. 24.4 COMPONENTS OF A WAVE WITH SQUARE-LAW DISTORTION

second-harmonic current, and that from quiescent current to peak negative current corresponds to peak fundamental current minus twice peak second-harmonic current.

Referring to the resultant waveform,

$$\mathrm{AC} = I_{m1} + 2I_{m2}$$
$$\mathrm{CD} = I_{m1} - 2I_{m2}$$

so that

$$I_{m1} = \frac{\mathrm{AC} + \mathrm{CD}}{2}$$

and

$$I_{m2} = \frac{\mathrm{AC} - \mathrm{CD}}{4}$$

Therefore

Percentage second-harmonic distortion

$$= \frac{I_{m2}}{I_{m1}} \times 100 = \frac{1}{2}\frac{\mathrm{AC} - \mathrm{CD}}{\mathrm{AC} + \mathrm{CD}} \times 100 \qquad (24.6)$$

Thus the second-harmonic distortion may be simply estimated from the dynamic characteristic. This is actually the distortion of the current waveform—the percentage distortion of the voltage waveform will be the same since the voltage and current are related by the dynamic resistance of the anode load. Though it is actually the second-harmonic power which is undesirable, it is always the percentage second-harmonic current or voltage which is referred to in specifications. "A limiting distortion of 5 per cent" means that the harmonic current should not exceed 5 per cent.

With triodes only the second harmonic is of importance. The distortion in pentodes and transistors has already been mentioned, in Section 24.3.

24.5 Class B and Class C Amplifiers

Operation in Class B is obtained if the active device is biased in the quiescent condition to the cut-off point. Current can therefore flow only during one half-cycle of the input signal. In a "single-ended" wide-band amplifier as shown in Fig. 24.1, this would, of course, give an intolerably distorted output. It may, however, be used in tuned power output stages, where the output frequency is determined by the resonant frequency of the tuned circuit, and hence distortion is negligible if the Q-factor of the tuned circuit is reasonable.

The advantage of this class of operation is a higher theoretical efficiency. This may be derived as follows.

Assume that the output current of the active device is a series of half sine-waves of peak value I_m. Then

$$\text{Mean output current} = \frac{1}{\pi} I_m$$

Since the device is cut off for one half-cycle, the power taken from the supply is

$$\text{Supply power} = \frac{1}{\pi} I_m V_{CC}$$

when V_{CC} is the supply voltage. There is an a.c. power output only during the conducting half-cycle, and this has a value given by

$$\text{A.C. power} = \frac{1}{2} \frac{V_m}{\sqrt{2}} \frac{I_m}{\sqrt{2}}$$

where V_m is the peak alternating output voltage. From this,

$$\text{Conversion efficiency} = \frac{V_m I_m}{4} \frac{\pi}{I_m V_{CC}} = \frac{\pi}{4} \frac{V_m}{V_{CC}} \tag{24.7}$$

Since the maximum theoretical value of the peak alternating output voltage must be equal to the supply voltage, V_{CC}, if the full output characteristic is used, it follows that

$$\begin{aligned}\text{Maximum possible Class B efficiency} &= \frac{\pi}{4} \times 100\\ &= 78{\cdot}5 \text{ per cent}\end{aligned}$$

Actual efficiencies of around 50 per cent may be achieved in tuned Class B amplifiers.

In Class C operation, the active device is biased well beyond the cut-off point, so that conduction occurs only during a fraction of a half-cycle of the signal. The output efficiency is higher than for Class B, but applications are limited, owing to the obvious disadvantage that there will be no output whatever if the input voltage falls below a critical value. Class C operation is, however, much used in tuned high-power valve amplifiers where the high efficiency is a great advantage.

Sometimes, particularly in push-pull circuits as described in the next section, an intermediate class of amplification designated as Class AB is employed. In this case the bias is set so that the active device conducts for more than half a cycle of the input. It is intermediate between Class A and Class B. Very frequently transistor amplifiers which are designated Class B are in fact Class AB.

The various classes of power amplification are illustrated for valve circuits on the mutual dynamic characteristics shown in Fig. 24.5,

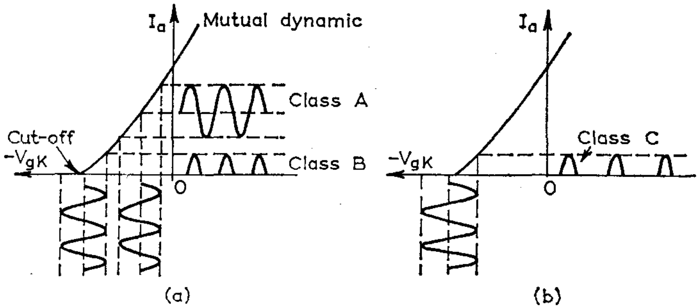


Fig. 24.5 CLASS A, B AND C OPERATION

where output current waveforms are shown for a sinusoidal input and resistive load superimposed on the appropriate grid bias. Notice that it is more usual to employ fixed (external) bias or grid current bias for Class B or C amplifiers, since the devices are cut-off in the quiescent state.

24.6 Class B Push-pull Operation

In Class B push-pull circuits the valves or transistors are arranged in such a way that each conducts for alternate half-cycles of the input. The outputs are combined to give a composite output which is the sum of those due to each active device alone.

A simple transistor output stage is shown in Fig. 24.6(*a*) using centre-tapped input and output transformers. The bases are driven in antiphase by the input transformer. During the positive half-cycles of the input, Tr_1 conducts and Tr_2 is cut off, so that the dotted end of the output transformer secondary goes negative. During the negative half-cycles of the input, Tr_1 is off and Tr_2 conducts, so that the dotted end of the output transformer secondary goes positive. Hence, theoretically, the output is an amplified version of the input.

In practice the base–emitter voltage, V_{BE}, required for conduction means that a transistor will not conduct until the input voltage exceeds V_{BE}. The output wave may therefore be considerably flattened near the zero current values. This flattening, which is shown in the output waveform at (*b*), gives rise to a type of distortion known as *crossover distortion*, which has a large third-harmonic component and is normally quite unacceptable. There is, however, almost no second-harmonic distortion, since positive and negative half-cycles are identical. If the transistors are not well matched, however, the positive and negative half-cycles of the output will not be identical so that second-harmonic distortion will appear. Cross-over distortion is minimized by forward biasing of the base by an amount which is sufficient to overcome V_{BE}. This is shown schematically at (*a*).

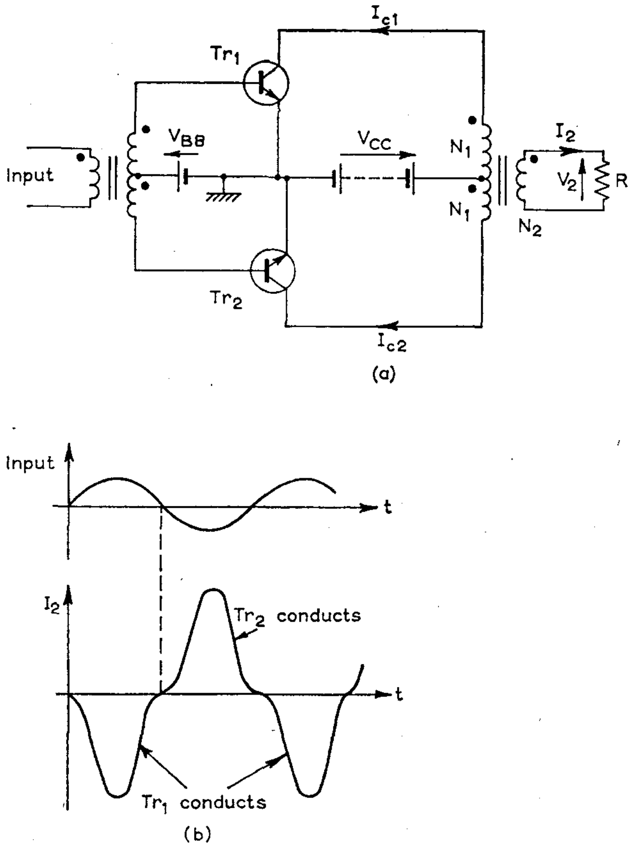


Fig. 24.6 CLASS B PUSH-PULL OUTPUT STAGE

The advantages of push-pull operation include the following:

1. Class B or Class AB operation is possible without the introduction of second-harmonic distortion (which generally has the most unpleasant effects in audio amplifiers). Hence high output efficiencies can be achieved, since larger inputs can be tolerated.

2. There is no resultant d.c. magnetization in the output transformer, since the mean currents in each half of the primary winding flow in opposite directions. It follows that no specially large iron cross-section need be used to reduce the second-harmonic distortion which would arise due to asymmetrical magnetization (and hence asymmetrical saturation) of the core. Any saturation effects will be symmetrical, and will therefore introduce odd harmonics only.
3. Since the fundamental signal currents are in anti-phase in each transistor, the signal current which flows through the d.c. supply (i.e. the sum of I_{c1} and I_{c2}) can contain no fundamental component.
4. Conversely, hum or noise in the d.c. supply will affect each transistor equally, and since the output is proportional to the *difference* between I_{c1} and I_{c2}, any such variations will cancel out.

24.7 Power Output—Class B Push Pull

For the circuit of Fig. 24.6(a), the load resistance reflected into each half of the primary of the output transformer is

$$R' = \left(\frac{N_1}{N_2}\right)^2 R \tag{24.8}$$

For a supply voltage of V_{CC}, the maximum possible collector current occurs if the whole output characteristic is used, i.e. if the transistor is driven up to saturation. The collector–emitter voltage will then be almost zero, so that the peak collector current is instantaneously

$$I_{cm} = \frac{V_{CC}}{R'}$$

and the voltage across the relevant half of the output transformer is V_{CC}.

Note that at this point the collector–emitter voltage of the cut-off transistor will be the supply voltage plus the voltage induced in its own half of the primary (which will have a peak value of V_{CC} also), i.e. it will have a possible maximum value of $2V_{CC}$. This helps to determine the output transistor to be used in a design (or the supply voltage for a given transistor), since it means that the transistor should have a peak voltage rating of at least twice the proposed supply voltage.

The maximum fundamental-frequency power output to the load R, assuming 100 per cent efficiency of the output transformer, is given by

$$P_m = \left(\frac{I_{cm}}{\sqrt{2}}\right)^2 R' = \left(\frac{V_{CC}}{\sqrt{2}R'}\right)^2 R' = \frac{V_{CC}^2}{2R}\left(\frac{N_2}{N_1}\right)^2 \tag{24.9}$$

The collector dissipation can now be found. It is simply the power, P_S, taken from the d.c. supply less the a.c. load power, P_m. Thus assuming half sine-waves of collector current through each transistor, the mean supply current is $(2/\pi)I_{cm}$. Hence

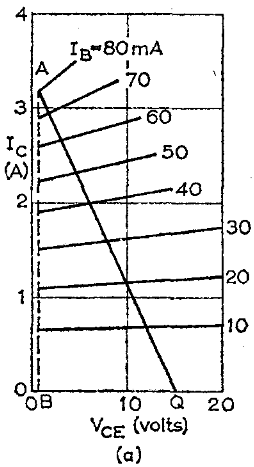
$$P_S = \frac{2}{\pi} I_{cm} V_{CC}$$

so that

$$\text{Collector dissipation per transistor} = \frac{P_S - P_m}{2} \tag{24.10}$$

The output current waveform for one half-cycle can readily be obtained from a load line drawn on the output characteristic of one transistor, the load line having a slope of $-R'$ and passing through the point $I_C = 0$, $V_{CE} = V_{CC}$. For identical transistors the other half-cycle of output current will be the same.

EXAMPLE 24.2 Two transistors whose characteristics are shown in Fig. 24.7(*a*) are used in Class B push-pull, to supply a 15 Ω load through an output transformer. The supply voltage is 15 V. Determine (*a*) a suitable transformer ratio, (*b*) the power output, and (*c*) the conversion efficiency. Sketch one half-cycle of the output waveform, assuming a sinusoidal base-current drive.



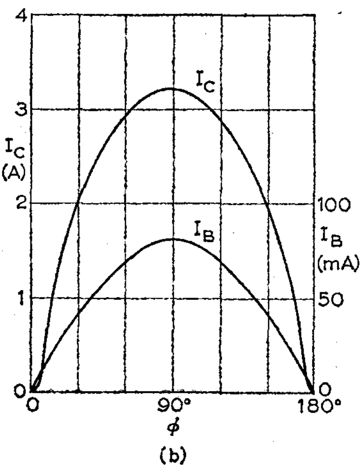


Fig. 24.7

(*a*) For maximum efficiency the dynamic load line (which starts at point Q in Fig. 24.7(*a*)) passes through the knee of the characteristic. From the diagram the effective primary resistance, R', is

$$R' = \frac{\text{BQ}}{\text{AB}} = \frac{14}{3 \cdot 2} = 4 \cdot 4\,\Omega$$

Hence, from eqn. (24.8), the turns ratio of the transformer is

$$\frac{N_1}{N_2} = \sqrt{\frac{R'}{R_L}} = \sqrt{\frac{4 \cdot 4}{15}} = \frac{1}{1 \cdot 86}$$

(*b*) With this turns ratio, the maximum power output to the load is

$$P_{cm} = \frac{I_{cm}^2}{2} R' = \frac{3 \cdot 2^2}{2} \times 4 \cdot 4 = \underline{\underline{22 \cdot 5\,\text{W}}}$$

(*c*) Assuming that the output current consists of a sine wave of peak value $3 \cdot 2$ A, the mean current from the supply is $(2/\pi) \times 3 \cdot 2 = 2 \cdot 03$ A, so that the supply power is $2 \cdot 03 \times 15 = 30 \cdot 5$ W, and

$$\text{Conversion efficiency} = \frac{22 \cdot 5}{30 \cdot 5} = \underline{\underline{0 \cdot 74\,\text{p.u.}}}$$

The collector-current waveform is obtained by sketching the sinusoidal base-current waveform over one half-cycle (i.e. to a peak of 80 mA) against phase angle as at (*b*). At any phase angle the collector current corresponding to the instantaneous base current is obtained from the intersection of the dynamic load line with the given base current.

Note that practical Class B amplifiers are normally arranged to have a small quiescent current in order to reduce crossover distortion.

24.8 Class AB Push-pull Amplifiers

In order to reduce crossover distortion, and at the same time to maintain a reasonably high conversion efficiency, it is usual to bias the output transistors (or valves) so that under quiescent conditions they are not quite cut off. This is Class AB operation. A typical circuit is shown in Fig. 24.8(*a*). Resistors R_1 and R_2 provide the base bias, and resistors R_3 and R_4 in the emitter leads help to stabilize the circuit against changes in temperature.

The graphical solution for the output may be obtained by drawing the composite characteristics of the push-pull pair. First the reflected load resistance is found from the following relations:

Primary voltage per turn = Secondary voltage per turn

i.e.

$$\frac{V_1}{N_1} = \frac{V_2}{N_2} \tag{24.11}$$

and

Total primary ampere-turns = Secondary ampere-turns

i.e.

$$(I_{c1} - I_{c2})N_1 = I_2 N_2 \tag{24.12}$$

where I_{c1} and I_{c2} are the collector signal currents.

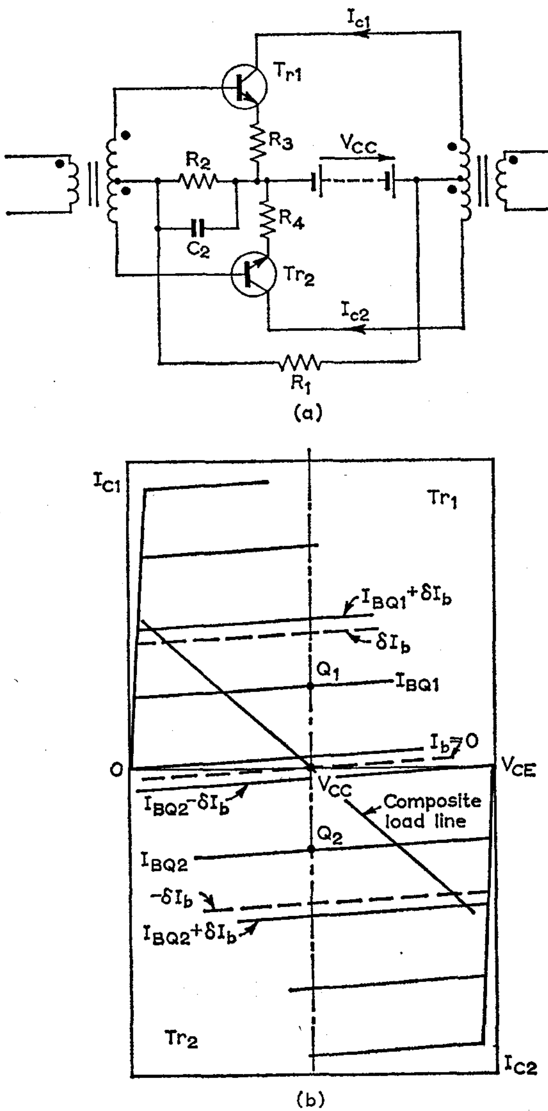


Fig. 24.8 CLASS AB PUSH-PULL OUTPUT STAGE

Dividing eqn. (24.11) by eqn. (24.12),

$$\frac{V_1}{(I_{c1} - I_{c2})N_1^2} = \frac{V_2}{I_2 N_2^2}$$

or

$$\frac{V_1}{(I_{c1} - I_{c2})} = \frac{V_2 N_1^2}{I_2 N_2^2} = \left(\frac{N_1}{N_2}\right)^2 R = R' \qquad (24.13)$$

This equation shows that the reflected load resistance, R', may be represented by a line of slope $-1/R'$ on a characteristic of the difference current $(I_{c1} - I_{c2})$ plotted to a base of collector voltage, V_{CE}. To obtain this characteristic it is convenient to plot two sets of collector characteristics, one of which has both current and voltage axes reversed. The voltage axes are arranged so that the points corresponding to the supply voltage, V_{CC}, coincide as shown in Fig. 24.8(*b*).

The composite characteristic for the quiescent base current (the same in each transistor) is then obtained by adding the currents shown by the static curves I_{BQ1} and I_{BQ2} at each value of voltage to give the broken line labelled $I_b = 0$ (i.e. no signal). Similarly for a signal current of δI_b the composite curve is obtained by adding the currents shown by the static curves $I_{BQ1} + \delta I_b$ and $I_{BQ2} - \delta I_b$ (corresponding to a signal current of $+\delta I_b$ in one transistor and $-\delta I_b$ in the other), to give the broken line labelled $+\delta I_b$; and for a signal current of $-\delta I_b$ by adding the currents shown by the static curves $I_{BQ1} - \delta I_b$ and $I_{BQ2} + \delta I_b$ to give the broken line labelled $-\delta I_b$. This process may be repeated for other values of δI_b as required.

The composite load line then has a slope of $-1/R'$ $(=N_2^2/N_1^2 R)$ and passes through the point $I_C = 0$, $V_{CE} = V_{CC}$.

A similar construction will apply in the case of valve push-pull amplifiers.

Class A push-pull operation is sometimes used in valve amplifiers where a large output with low distortion is required. This class of operation is unusual in transistor circuits because of the large quiescent dissipation and the resulting difficulty in achieving adequate thermal stability. Also in Class A the efficiency is below 50 per cent, and for a given power output the cost of transistors is higher than that of valves.

EXAMPLE 24.3 The valve push-pull amplifier shown in Fig. 24.9(*a*) is operated with a supply voltage of 400 V and a grid bias of -12 V to give Class AB operation. The valves used are identical Mullard EL31 pentodes whose characteristics are shown at (*b*). The output transformer has an overall ratio of 20:1 and the load resistance is 10 Ω. At an anode voltage of 400 V, the valve is cut off if the grid voltage is less than -20 V.

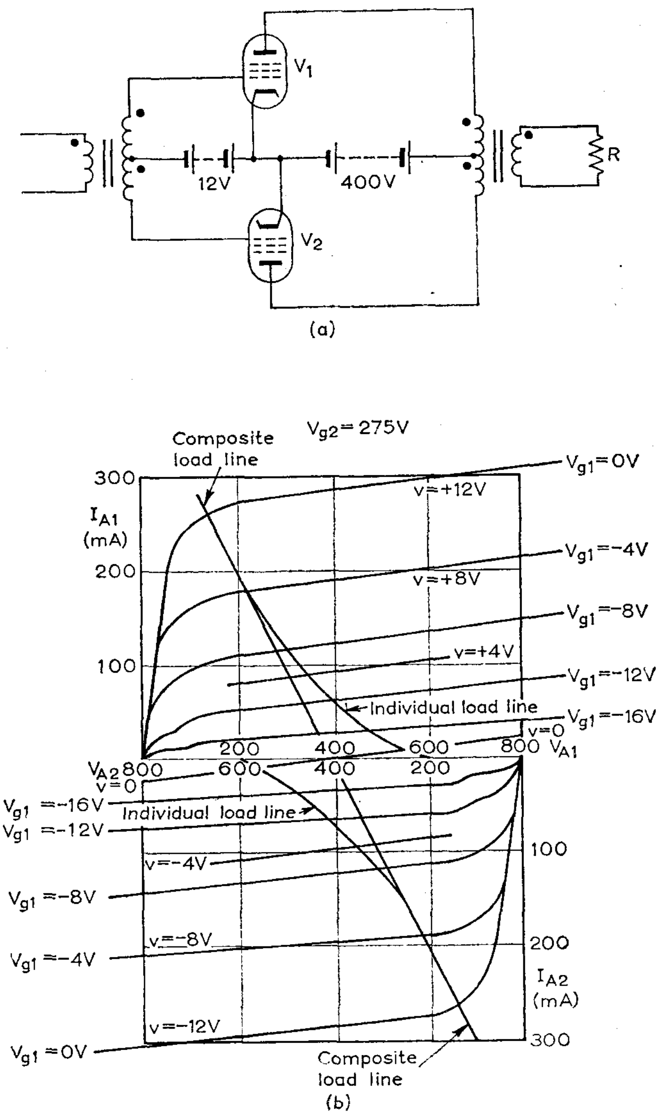

V1
V2
12V
400V
R
(a)
Vg2 = 275V
Composite load line
300
IA1 (mA)
200
100
Vg1 = 0V
v = +12V
Vg1 = −4V
v = +8V
Vg1 = −8V
v = +4V
Vg1 = −12V
Individual load line
Vg1 = −16V
v = 0
200
400
600
800 VA1
VA2 800
600
400
200
v = 0
Vg1 = −16V
Individual load line
Vg1 = −12V
v = −4V
100
Vg1 = −8V
v = −8V
200
Vg1 = −4V
IA2 (mA)
v = −12V
300
Vg1 = 0V
Composite load line
(b)

Fig. 24.9

Draw the composite characteristics, and determine the power output, and estimate the conversion efficiency if the sinusoidal signal input voltage is 12 V peak at each grid.

The composite characteristic is drawn by inverting one set of individual characteristics, and setting the supply voltage points (400 V) together.

For a zero signal ($v = 0$) the composite characteristic is obtained by adding the currents shown by the static curves for $V_{g1} = -12$ V at each value of anode voltage. This is the line marked $v = 0$ at (*b*), passing through the 400 V zero-current point. For a signal of +4 V on valve V_1 (and hence −4 V on valve V_2) the composite curve is obtained by adding the currents of the static characteristic for $V_{g1} = -12 + 4 = -8$ V for V_1 and that for $V_{g1} = -12 - 4 = -16$ V for V_2. This is the line marked $v = +4$ V at (*b*).

When the signal exceeds ±8 V one or other valve becomes cut off and the composite characteristics coincide with the static characteristics for each valve alone.

The composite load line has a slope of $-1/R'$ (where $R' = R(N_1/N_2)^2 = 10 \times (10)^2 = 1{,}000$, since the ratio for each half of the output transformer is 10:1). This load line passes through the 400 V zero-current point as shown. Notice that the individual load lines are curved—these represent the current/voltage relations for the individual valves. Thus, for zero resultant output current, the anode voltages are each 400 V and the grid voltages each −12 V, giving quiescent anode currents of 60 mA, so that (400 V, 60 mA) is a point on the individual load characteristics. When the signal voltage is $v = +8$ V, the anode voltage of the conducting valve is 200 V, and the current is 180 mA, while the anode voltage of the valve which is cut off is 600 V and its anode current is zero. Hence points (600 V, 0 mA) and (200 V, 180 mA) are points on each individual load line.

The d.c. power input is found by estimating the mean anode current when the ±12 V signal is applied. This is done by drawing the anode-current waveform for one valve, using the individual load lines. In this problem, it is estimated that the mean anode current is 100 mA. Hence

$$\text{D.C. power input} = 2 \times 400 \times 0{\cdot}1 = 80\,\text{W}$$

The a.c. power output is obtained from the composite characteristics. Thus

$$\text{Peak difference current} = 263\,\text{mA (at } v = \pm 12\,\text{V)}$$

Hence

$$\text{R.M.S. difference current} = 186\,\text{mA}$$

$$\text{Peak alternating voltage on one side} = 260\,\text{V}$$

Hence

$$\text{R.M.S. alternating voltage on one side} = 184\,\text{V}$$

It follows that

$$\text{A.C. power output} = 184 \times 0{\cdot}186 = \underline{\underline{34\,\text{W}}}$$

Therefore

$$\text{Conversion efficiency} = \tfrac{34}{80} = \underline{\underline{0{\cdot}43\,\text{p.u.}}}$$

24.9 Push-pull Complementary-symmetry Output Stages

The fact that both *p–n–p* and *n–p–n* transistors are available enables push-pull circuits to be designed without transformers—i.e. with single-ended inputs and outputs. Thus if a *p–n–p* and an *n–p–n* transistor are fed from the same drive, a given input swing will cause one transistor to conduct more while the other conducts less, giving a push-pull operation.

Consider the circuit of Fig. 24.10. Tr_1 acts as the driver, the bias being designed so that the voltages at A and B are just sufficient to cause both Tr_2 and Tr_3 to conduct slightly. The voltage at D is set by adjustment of R_1 to be approximately $V_{CC}/2$. Diode D_1 will be a silicon diode if Tr_2 and Tr_3 are silicon power transistors, in order to

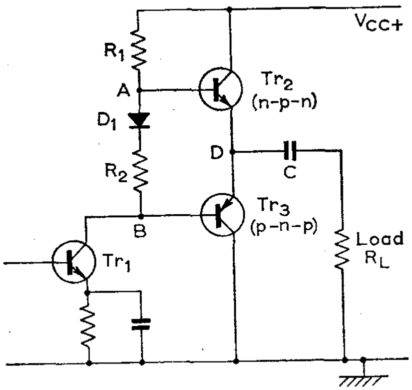


Fig. 24.10 TRANSFORMERLESS OUTPUT STAGE

improve the thermal stability of the circuit; i.e. the change of voltage between A and B with temperature will match the corresponding changes at the emitters of Tr_2 and Tr_3. R_2 is of low value so that A and B are approximately at the same voltage so far as the signal is concerned.

When the current in Tr_1 falls, the voltages at A and B both rise, causing Tr_3 to cut off and Tr_2 to take more current—the current flow being through the load resistor R_L. Under maximum power conditions the voltage at D approaches V_{CC}. When the current in Tr_1 rises above its quiescent value, the voltages at A and B fall, and Tr_2 becomes cut off while Tr_3 takes more current (by discharging capacitor C). The load current through R_L therefore reverses as C discharges, and for maximum power the voltage at D falls to almost zero. Positive and negative half-cycles will be identical if the transistors Tr_2 and Tr_3 have identical characteristics.

For high power outputs the complementary pair of transistors is used to drive two high-power transistors of the same type—added to a circuit such as that of Fig. 24.10, these would be two *n–p–n* power transistors as in Fig. 24.11.

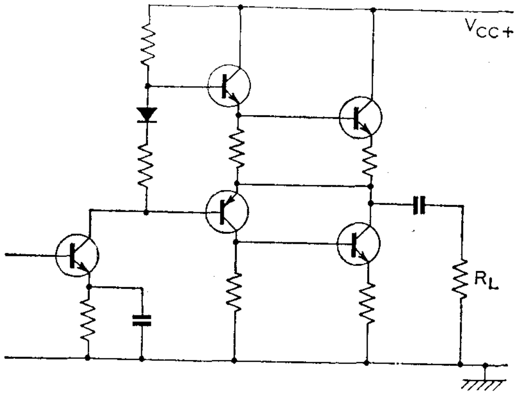


Fig. 24.11 HIGH-POWER COMPLEMENTARY-SYMMETRY OUTPUT STAGE

EXAMPLE 24.4 For the circuit of Fig. 24.10, the complementary transistors have the characteristics shown in Fig. 24.12. The supply voltage is 40 V. Assuming practical Class B operation (i.e. in the quiescent state both output transistors are almost cut off), determine the optimum load resistance, the approximate power output and the conversion efficiency.

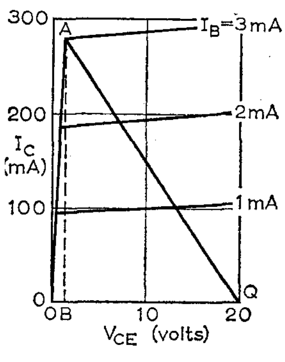


Fig. 24.12

In the quiescent state assume that $V_{CE} = 20$, $I_C = 0$ (point Q). The optimum load will give a load line from Q to the knee of the characteristic as shown. Then the optimum load resistance is

$$R_L = \frac{QB}{AB} = \frac{18{\cdot}6}{0{\cdot}28} = \underline{\underline{66\Omega}}$$

The peak voltage swings across the load will be $\pm 18{\cdot}6$V, so that the maximum power output will be approximately

$$P_m = \frac{18{\cdot}6^2}{2R_L} = \frac{18{\cdot}6^2}{2 \times 66} = \underline{\underline{2{\cdot}6\,\mathrm{W}}}$$

Assuming sine waves, the peak output current is 280mA, and hence the current from the supply is $(1/\pi) \times 0{\cdot}28 = 0{\cdot}09$A. The conversion efficiency is thus

$$\eta = \frac{2{\cdot}6}{0{\cdot}09 \times 40} = \underline{\underline{0{\cdot}72\,\mathrm{p.u.}}}$$

Note that, since Tr_2 is cut off for half a cycle of the signal, current is drawn from the d.c. supply for only one half-cycle and hence the mean current is $(1/\pi)$ of the peak signal current.

PROBLEMS

24.1 A certain pentode has the following characteristics when operated at a screen voltage of 400V:

V_g	V_A	200	300	400	800 volts
−8V		100	110	116	120mA
−16V	I_A	51	59	63	67mA
−24V		19	21	22	23mA

Find the amplification factor, r_a, and g_m, when the grid bias is −16V and the anode voltage is 400V. If the peak grid signal is 8V, calculate the output power if the load is 4,000Ω. (*L.U. part question*)

Ans. 170; 30kΩ; 5·8mA/V; 3·3W.

24.2 Draw the anode characteristics for the triode from which the following test figures were obtained:

$V_g = 0$V	V_A	0	50	100	150	200	250V
	I_A	0	18	53	110	183	260mA
$V_g = -40$V	V_A	100	150	200	250	300	350V
	I_A	3	20	50	100	170	240mA
$V_g = -80$V	V_A	200	250	300	350	400V	
	I_A	5	22	50	100	160mA	
$V_g = -120$V	V_A	300	350	400	450	500V	
	I_A	7	25	52	100	150mA	
$V_g = -160$V	V_A	350	400	450	500	550V	
	I_A	1	10	30	58	95mA	
$V_g = -200$V	V_A	450	500	550	600	650V	
	I_A	2	10	22	58	100mA	
$V_g = -240$V	V_A	550	600	650	700	750V	
	I_A	3	12	30	62	105mA	
$V_g = -280$V	V_A	650	700	750	800V		
	I_A	3	15	30	60mA		

The triode is required to deliver power to a purely resistive load of 300Ω. An output transformer of ratio 3:1 is available. A d.c. supply voltage of 500V is to be used. Considering a mean anode current of 100mA and a peak grid signal of 120V, calculate (*a*) the power delivered to the load, (*b*) the anode efficiency, (*c*) the percentage second-harmonic distortion.

Ans. 7·5W; 17 per cent; 6 per cent approx.

24.3 Draw the anode characteristics for the tetrode for which the following test figures were obtained when the screen voltage was 250V.

	V_A	40	80	120	200	300	400V
$V_g = 0$	I_A	62	72	76	79	81	83mA
$V_g = -5$V	I_A	50	57	61	64	67	70mA
$V_g = -10$V	I_A	37	43	46	48	50	52mA
$V_g = -15$V	I_A	27	32	34	37	37·5	37·5mA
$V_g = -20$V	I_A	17·5	22	24	26	27	27mA
$V_g = -25$V	I_A	10	13	15	17·5	18	18mA
$V_g = -30$V	I_A	4	7	8	10	11	11mA

The tetrode is required to deliver power to a purely resistive load of 70Ω. An output transformer of ratio 10:1 is available A d.c. supply voltage of 250V is to be used. Considering a mean anode current of 30mA and a peak grid signal of 10V, calculate (*a*) the power delivered to the load, (*b*) the anode efficiency, (*c*) the percentage second-harmonic distortion.

Ans. 1·2W; 15 per cent; 4 per cent.

24.4 A push-pull power amplifier utilizes two identical triodes, each of which has the anode characteristics given in Problem 24.2. If the anode supply is 500V and the grid bias is −160V for both valves, determine the power output and efficiency when the peak grid voltage is 160V and the valves are coupled to a load of 15Ω by a transformer of total primary turns to secondary turns ratio of 16·7:1.

Ans. 32W; 30 per cent approx.

24.5 A push-pull power amplifier utilizes two identical tetrodes, each of which has the anode characteristics given in Problem 24.3. If the anode supply is 300V and the grid bias is −25V for both valves, determine the power output and efficiency when the peak grid voltage is 25V and the valves are coupled to a load of 15Ω by a transformer of total primary turns to secondary turns ratio of 31·6:1.

Ans. 8W; 45 per cent approx.

24.6 Discuss, with the aid of sketches, the causes and effects of distortion in amplifiers.

The anode current flowing in an amplifier is given by

$$I_A = 0{\cdot}05(15 + V_g)^2 \quad \text{milliamperes}$$

where V_g is the grid voltage. Calculate (*a*) the steady anode current with a grid bias of 8V, (*b*) the mean anode current when an alternating signal of 6 sin ωt is superimposed on this grid bias, (*c*) the amplitude of the fundamental anode current in (*b*), and (*d*) the amplitude of the second-harmonic current in (*b*).

(*H.N.C.*)

Ans. 2·4mA; 3·3mA; 4·2mA; 0·9mA.

24.7 A power transistor is used in the circuit of Fig. 24.1(*c*). The operating region on the collector characteristic is bounded by (i) $I_{C\max} = 2$ A; (ii) maximum collector dissipation = 10 W; (iii) maximum collector voltage = 30 V. The emitter resistor has a value of 1 Ω. Sketch the limits of the permissible operating region, and determine suitable values for the effective collector load resistance, supply voltage, quiescent collector current and maximum power output for a sinusoidal signal.

Ans. 10Ω; 11 V; 1 A; 5 W.

24.8 Two transistors with the ratings given in Problem 24.7 are to be used in Class B push-pull. The supply voltage is 15 V. Determine a suitable value for the effective resistance presented to each collector, the corresponding maximum power output for a sinusoidal signal, and the maximum collector dissipation.

Ans. 7·5Ω; 15 W; 2·1 W.

24.9 Fig. 24.13 shows the circuit for a popular linear power amplifier using a complementary pair of transistors Tr_1 and Tr_2. Describe the operation of this

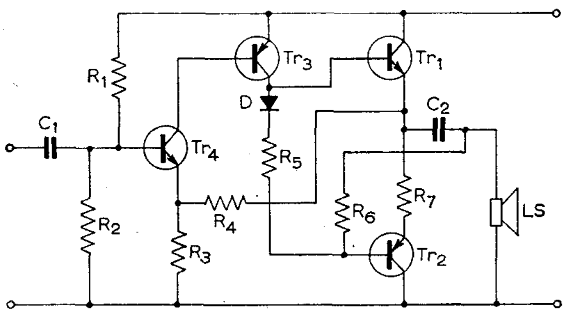


Fig. 24.13

circuit, and comment on the factors that affect the linearity. Discuss the use of the diode D.

Chapter 25

POWER RECTIFICATION

Since electrical distribution almost universally uses alternating currents, applications which require direct current necessitate the provision of either a d.c. generator or a rectifier equipment. For fixed voltage d.c. supplies, half-wave, full-wave, bridge or polyphase connected rectifiers may be used, employing vacuum or gas-filled valves, or semiconductor diodes. In these cases the required output voltage is obtained by a suitable choice of ratio for the transformer which supplies the rectifier circuit. Diode rectification is used for power and communication circuits over the range of frequencies extending from power frequencies up to microwave frequencies,* and including signal-detection circuits. Semiconductor devices can conveniently be used for most applications, ranging from high-power mains-frequency installations right through to applications at frequencies in the microwave region.

Powers involved range from microwatts in signal-detection circuits to megawatts in, for example, the rectifiers employed in d.c. power system interconnectors. Only power rectification will be considered in this chapter.

In many applications a variable d.c. supply is required (e.g. control of d.c. machines). This may involve the use of devices which can be controlled so that conduction takes place for a predetermined fraction of the a.c. cycle. Such devices include the thyratron, mercury-arc rectifier and silicon controlled rectifier (or thyristor).

* Electromagnetic waves whose frequencies are higher than 1 gigahertz are called *microwaves* (1 GHz $= 10^9$ Hz).

25.1 Conduction in a Gas

Gas-filled valves contain a gas at a pressure between 10^{-4} and 10 metres of mercury depending on the design and application of the valve. In these valves there are frequent collisions between gas atoms and electrons during the passage of electron current from cathode to anode. The most common gas filling is mercury vapour, but argon, neon, hydrogen and other gases are also used.

There are three types of collision which may occur between an electron and a gas atom depending on the speed or kinetic energy of the bombarding electron. If the speed is low (corresponding to an energy of 2 or 3 eV), the electrostatic fields of the bombarding electron and the outer electrons of the atom interact with one another and the electron is repelled. The "collision" follows the ordinary laws of mechanics, and the electron rebounds with only a small transference of energy to the very much more massive atom. This is called an *elastic collision.* If the bombarding electron has a somewhat greater energy, it may, on collision, cause a temporary disturbance of the electron orbits within the atom. In the disturbance the atom will, internally, absorb some of the energy of the bombarding electron so that the electron rebounds with reduced energy. The disturbance within the atom consists of the temporary raising of the energy associated with an internal electron followed by the return of that electron from its abnormal energy state to its normal energy state. During the return the excess energy, which has been absorbed from the bombarding electron, is emitted from the atom as an electromagnetic wave with a characteristic wavelength depending on the atom. For some gases the wavelength is in the visible band (e.g. sodium vapour gives a characteristic yellow light), and this is the basis of operation of some electric discharge lamps. During the disturbance the atom is said to be excited and the collision is known as an *excitation collision.*

The third type of collision is known as an *ionizing collision.* The energy of the bombarding electron is then from 10 to 20 eV, depending on the gas. With this energy an internal electron may be completely freed from the atom so that there are two free electrons and a positively charged atom, or *positive ion.* The atom will have absorbed from the bombarding electron an amount of energy equal to the amount required to free an electron from the atom.

The ionizing collision is by far the most important from an electrical circuit point of view, and this type of collision will be implied hereafter unless otherwise stated. The potential difference through which an electron must pass to gain sufficient energy to give an ionizing collision with a particular gas atom is called the *ionization potential* for the gas.

The above are the fundamental concepts of gaseous conduction upon which the operation of gas valves depend. It will be evident that a gas will be a good insulator unless ionizing collisions (ionization) occur. It will then be a fairly good conductor.

25.2 Gas Diodes

The gas diode has an indirectly heated cathode which provides a copious supply of primary electrons. The gas is almost invariably mercury vapour, since (i) this gas has a low ionization potential (10·6 V), which leads to efficient operation, and (ii) the mercury

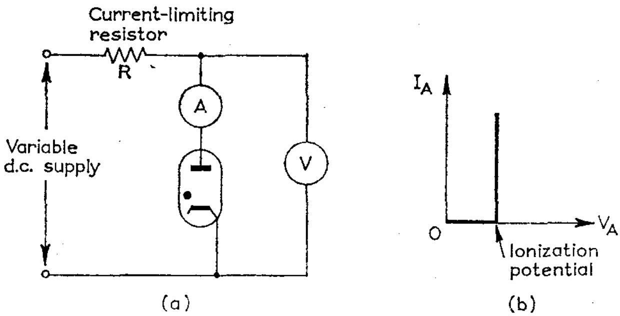


Fig. 25.1 CHARACTERISTIC OF A GAS DIODE

vapour does not react chemically with the oxide coating on the cathode. The principal disadvantage of mercury vapour is the large pressure variation which occurs with a moderate temperature variation (0·0001 mm Hg at 0°C to 0·1 mm Hg at 80°C). This makes the operation satisfactory over only a narrow temperature range (40–80°C). The anode is either solid carbon or carbon coated to give a good thermal emissivity.

The circuit of Fig. 25.1(*a*) may be used to determine the valve characteristics. If the supply voltage is raised from zero, the current will be found to be negligible until a critical voltage is reached. At this voltage ionization occurs and the valve becomes highly conducting. The current-limiting resistor *R* must be included in the circuit to limit the current when ionization occurs. For all values of current up to the permissible limit, the voltage drop across the valve remains almost constant at approximately the ionization potential. The current conduction continues until the applied potential falls to too low a value to maintain the ionization. The anode characteristic is shown in Fig. 25.1(*b*).

The action of the gas diode depends on the greater mobility of the electron compared with the positive ions (due to the much smaller mass of the electron). The electrons move much more quickly towards the anode than do the positive ions towards the cathode, so that a resultant positive space charge develops; this tends to increase the electric stress towards the cathode, and to decrease the electric stress towards the anode. This is depicted in Fig. 25.2(*a*), where it

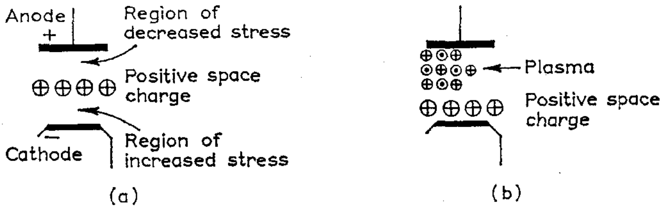


Fig. 25.2 CONDITIONS IN CATHODE–ANODE SPACE OF A GAS DIODE

has been assumed that the positive space charge first builds up in the middle of the interspace. Owing to the high stress, electrons will move quickly from the negative space charge around the thermionic cathode to the positive space charge, and at the same time electrons will move slowly from the positive space charge to the anode, owing to the decreased electric field strength in this region. With the increased flow of electrons from the cathode space charge more positive ions will be formed and the positive space charge will continue to increase, the limit being the value of the anode potential, for if the space charge potential exceeded the anode potential, electrons would move back into the positive space charge and neutralize the effect of positive ions until the space charge potential again fell below the anode potential.

Between the positive space charge and the anode there is a region in which the electric stress becomes very low; this region is filled with about equal numbers of positive ions and negative electrons; it is highly conducting and is called the *plasma*. This is illustrated in Fig. 25.2(*b*). It will be seen that, in effect, the positive space charge has become the effective anode joined to the actual anode by the highly conducting plasma. Since the plasma contains about equal numbers of electrons and positive ions at any given instant, and since, even in the plasma, the electrons will be moving with a much greater velocity than the positive ions, there must be many more electrons passing through a given section in any given time than there are positive ions. Therefore by far the larger portion of the current through the valve is due to electrons thermionically emitted from the cathode. The proportions have been estimated to be of the order of 100 to 1. It should be clearly realized that the function of the positive ions is to assist the thermionically emitted electron current rather than to contribute to the current themselves.

The constant-voltage nature of the conduction results from the tendency of the positive space charge to approach the cathode surface. The closer the positive space charge is to the cathode the higher will be the electric stress. However, provided that saturation current for the given cathode temperature is not reached, the electric stress at the cathode will determine the electron current from the cathode space charge. For a given external current the positive space charge will only approach the cathode surface to a distance that will correspond with the required current. If the space charge approached more closely than this, there would be an excess electron emission which would cancel the positive space charge. If the external current is increased the space charge will move correspondingly nearer to the cathode and so give the additional current through the valve with no additional voltage drop across the valve. In most cases it is sufficiently accurate to assume that the voltage drop is constant for all normal currents and is equal to the voltage at which conduction will first start (*ignition voltage*) or the voltage at which conduction ceases (collapse or *extinction voltage*).

The most severe limitation on the use of gas valves is the peak-current limitation—this, usually, must not exceed two or three times the mean current even momentarily (bombardment of the coated cathode by ions may permanently destroy the surface if excess currents are permitted). Reservoir capacitors should not be used in gas-valve rectifier circuits as these tend to produce high peak currents.

EXAMPLE 25.1 A gas diode is connected in series with the output of a single-phase transformer to supply the following circuits in turn:

(*a*) a pure resistance of 50Ω in series with a choke of 0·2H inductance and negligible resistance,
(*b*) a battery of constant e.m.f. 80V (connected for charging) in series with a limiting resistance of 10Ω.

If the gas diode has an effective voltage drop of 10V and the transformer gives a sinusoidal output of 100V (r.m.s.) at 50Hz, calculate the mean and peak currents in each case.

(*a*) The circuit is represented in Fig. 25.3(*a*) and the wave diagram is shown in Fig. 25.3(*b*). V_A is the potential of the anode with respect to earth potential. The instant at which conduction starts (i.e. when $V_A = 10$V) is taken as the zero reference for the time base. Therefore

$$V_A = 141 \sin(\omega t + \phi)$$

where $141 \sin\phi = 10$, i.e. $\phi = \sin^{-1}(10/141) = 4°$.

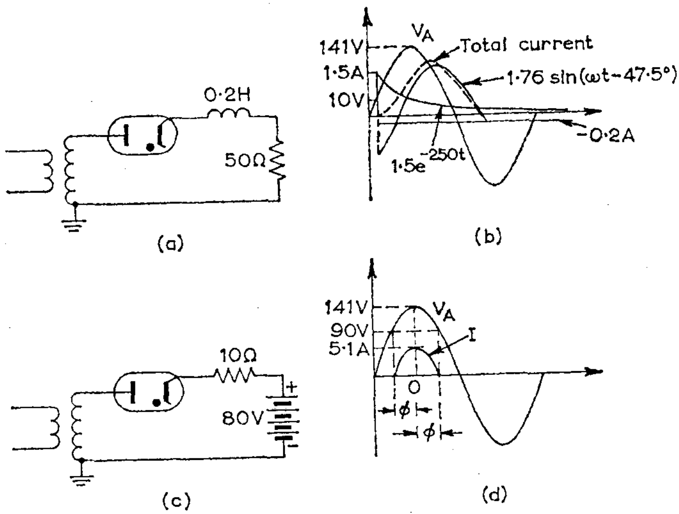


Fig. 25.3

Thus the valve may be taken as conducting between the instant $t = 0$ and the instant at which the cathode–anode voltage is again 10 V, i.e. the instant at which the current becomes zero. For this range,

$$L\frac{di}{dt} + Ri = 141 \sin(\omega t + \phi) - 10$$

This is a transient equation which will have a solution of the form i = steady-state current (i_s) + transient current (i_t) (by Section 6.1), where

$$i_s = \left\{\frac{141}{\sqrt{R^2 + \omega^2 L^2}} \sin\left(\omega t + \phi - \tan^{-1}\frac{\omega L}{R}\right) - \frac{10}{R}\right\}$$

$$= \{1{\cdot}76 \sin(\omega t - 47{\cdot}5°) - 0{\cdot}2\}\text{A}$$

and

$$i_t = A e^{-\frac{R}{L}t}$$

where A is a constant.

At $t = 0$, $i = 0$, due to the inductor effect. Thus

$$i = i_t + i_s = 0 = A + 1{\cdot}76 \sin(-47{\cdot}5°) - 0{\cdot}2$$

so that $A = 1{\cdot}5$, and the total current is

$$i = 1{\cdot}5\, e^{-250t} + 1{\cdot}76 \sin(\omega t - 47{\cdot}5°) - 0{\cdot}2$$

The three components of this current wave are drawn and added graphically in Fig. 25.3(*b*). It will be seen that the conduction is spread over more than half a cycle. From the graphs

$$\frac{\text{Peak current}}{\text{Mean current}} = \frac{0{\cdot}8}{0{\cdot}33} = \underline{\underline{2{\cdot}4}}$$

Inductive smoothing is very useful with gas valves for it helps to overcome the peak current limitations. Capacitor smoothing has the opposite effect.

(*b*) The circuit is shown in Fig. 25.3(*c*). It will be seen that the potential of the cathode is 80 V above earth when no current flows. Thus the anode potential must exceed 90 V above earth to start conduction. If the peak positive voltage is taken as the zero reference of time, then conduction will commence when

$$\phi = -\cos^{-1}\frac{90}{141} = -50{\cdot}4^\circ$$

and conduction will cease when

$$\phi = \cos^{-1}\frac{90}{141} = +50{\cdot}4^\circ$$

Between $-\phi$ and $+\phi$ the voltage across the current-limiting resistor is

$$v_R = 141\cos\omega t - 90$$

Therefore

$$\text{Charging current} = \frac{v_R}{10} = \frac{141}{10}\cos\omega t - \frac{90}{10}$$

and

$$\text{Peak current} = 14{\cdot}1 - 9 = \underline{\underline{5{\cdot}1\,\text{A}}}$$

$$\text{Mean current} = \frac{1}{10}\frac{1}{2\pi}\int_{-\phi}^{+\phi}[141\cos\omega t - 90]d(\omega t)$$

$$= \frac{141\sin\phi}{10\pi} - \frac{90\phi}{10\pi} = \underline{\underline{0{\cdot}93\,\text{A}}}$$

Note that with a load consisting of a pure resistance the conduction will commence at the striking voltage of the valve and will cease at the extinction voltage.

25.3 Gas Triodes (Thyratrons)

A cross-section of the electrode assembly for a gas triode is shown in Fig. 25.4(*a*). The anode and cathode are screened from each other by a shield or grid electrode. The shield is designed to prevent any electron reaching the anode from the cathode except by way of the hole in the grid electrode. The circuit for testing a gas triode is shown in Fig. 25.4(*b*).

If the grid potential is held negative with respect to the cathode, then until the anode voltage reaches a critical potential (the *striking* voltage) no electrons will leave the cathode space charge and there will be neither ionization nor conduction. This corresponds to the cut-off conditions in a vacuum valve. If the striking voltage is exceeded, electrons will pass from the cathode space charge and give ionizing collisions with gas molecules.

When ionization commences, positive ions immediately collect around the negative grid and completely neutralize its effect. The conditions in the triode are then identical with those in the gas diode, i.e. the anode voltage falls to an almost constant value approximately equal to the ionization potential and nearly independent of the current through the valve. This current is limited only by the external impedances and will cease only when the supply voltage is reduced below the ionization or *maintaining voltage*. Any change in grid voltage after ionization has occurred will not affect the conduction through the valve, and the grid cannot be used to stop

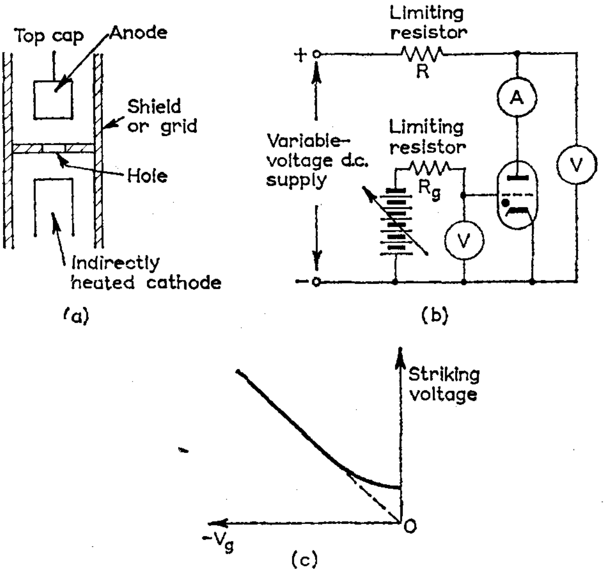


Fig. 25.4 THE GAS TRIODE

conduction. The grid of a gas triode only controls the value of anode voltage at which the valve will start to conduct.

Fig. 25.4(*c*) shows the grid-control characteristic for a gas triode (i.e. the anode striking voltage for a range of negative grid potentials). Above about 3 V the ratio of striking voltage to control voltage is constant at a value called the *control ratio*. This ratio is usually about 30.

Gas triodes are not of great use in circuits with direct current supplies, since once conduction has commenced it is difficult to interrupt the anode current except by interrupting the supply. The principal use of gas triodes is as grid-controlled rectifiers, in which case the grid regains control of the valve every negative half-cycle.

25.4 Grid-controlled Single-phase Rectification

Simple negative bias control of a thyratron is not normally employed, since it permits control only over the first quarter of the a.c. waveform applied to the anode.

PULSE CONTROL

The circuit is shown in Fig. 25.5(*a*). The grid has a large negative bias so that even the peak supply voltage will not be sufficient to

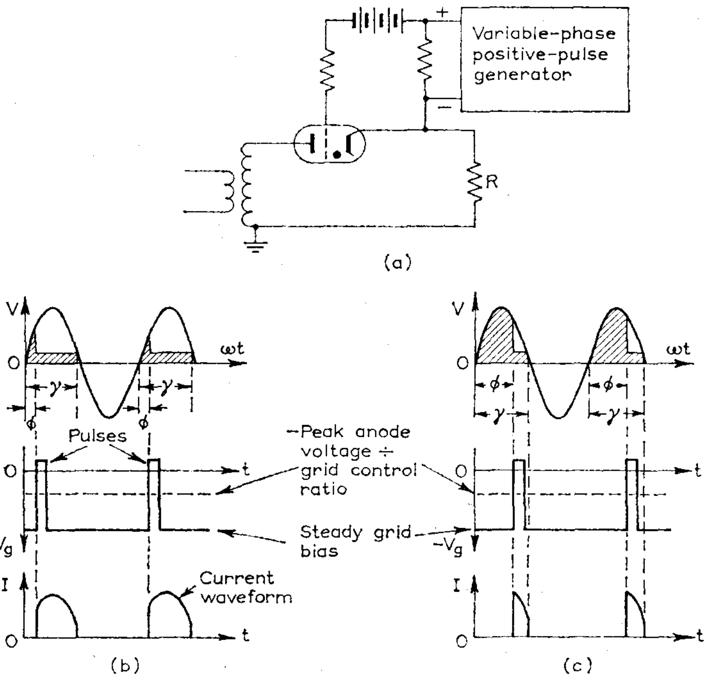


Fig. 25.5 PULSE CONTROL OF A GAS TRIODE

cause conduction through the valve in the absence of the positive pulses injected into the grid circuit. The amplitude of the positive grid pulses is at least equal to the magnitude of the grid bias voltage so that conduction will always commence at the instant in the cycle at which the positive pulse is applied, provided only that the supply voltage is positive and greater than the ionization potential at the instant at which the pulse occurs. Fig. 25.5(*b*) shows the anode voltage and current waveforms when the pulse occurs in the first positive quarter-cycle. Fig. 25.5(*c*) shows the same waveforms when the pulse occurs in the second positive quarter-cycle. It will be clear that this method has the advantages of (i) controlling the current from maximum mean value to zero mean value, and (ii) starting the conduction at a definite instant in the cycle independent of any supply voltage variations. The mean current for a resistive load may be determined from the following equation:

$$\text{Mean output current} = \frac{1}{2\pi R}\int_{\phi}^{\gamma}[V_m \sin \omega t - \text{volt drop}]d(\omega t) \tag{25.1}$$

where R = Load resistance
ϕ = Ignition angle
γ = Extinction angle

The pulse unit may be a mechanical contact driven by a synchronous motor, a peaking transformer supplied from a phase-shifting circuit, or an electronic pulse generator with a phase-shifting circuit giving pulses synchronized with the supply frequency.

PHASE-SHIFT CONTROL

This method has been developed to preserve the advantages of the pulse control circuit without the necessity of having a pulse generator. A sinusoidal voltage with or without a negative grid bias is applied to the grid of the gas triode. The sinusoidal voltage has the same frequency as the supply voltage and usually a constant magnitude. It will be found that, by altering the phase of the grid voltage relative to that of the supply voltage, the instant at which conduction commences may be controlled.

Fig. 25.6(*a*) shows a phase-controlled gas-triode rectifier circuit incorporating one possible method of deriving a constant-magnitude variable-phase voltage. By Section 1.7 an *RC* series circuit with variable resistance has a semicircular complexor locus diagram as shown in Fig. 25.6(*b*). It will be seen that the potential difference between the centre point of the winding an the junction of the resistor and capacitor has a constant magnitude and a phase which is variable over almost 180° as R is varied from zero to infinity. An additional 180° phase shift may be obtained by reversing the connexions to the auxiliary transformer winding.

The value of the mean output current may be obtained from eqn. (25.1), where the ignition angle ϕ is found in the following manner. In Fig. 25.6(*c*) the supply voltage waveform is drawn and also the waveform of critical grid voltage corresponding to the supply voltage, i.e. the negative grid voltage which would be just sufficient to prevent the valve from striking at the instantaneous value of supply voltage. The external grid-voltage wave is now superimposed

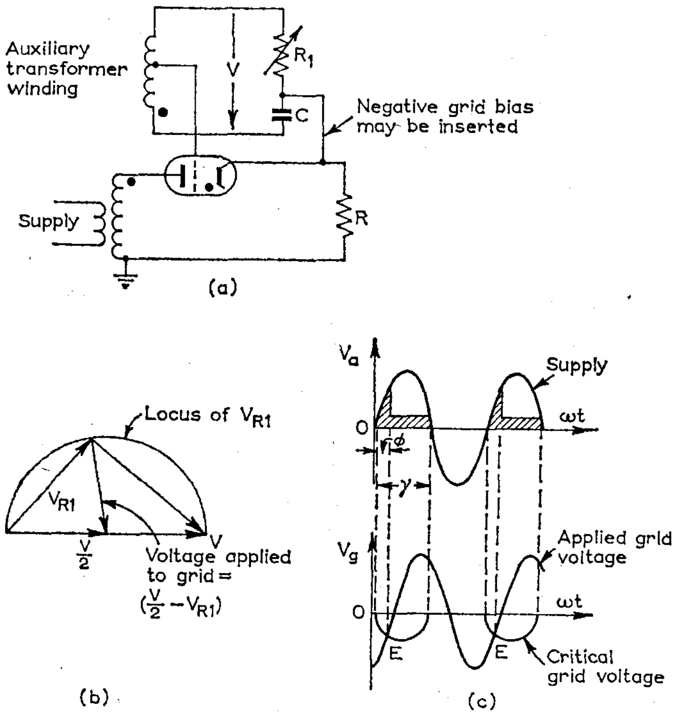


Fig. 25.6 PHASE-SHIFT CONTROL OF A GAS DIODE

on the critical grid-voltage wave in the correct phase. The point, E, of intersection of the external grid-voltage waveform and the critical grid-voltage waveform where the external voltage is going positive gives the striking point in the cycle. It is evident that by altering the phase of the external voltage the striking point and hence the mean current may be controlled.

EXAMPLE 25.2 A gas-filled triode, used as a phase-controlled rectifier on a 250 V supply, has a control ratio of 20 and a negligible tube drop. If the voltage applied to the grid is 100 V (r.m.s.) lagging behind the anode supply voltage by 60°, find the mean anode current in a load of 1,500 Ω. (*H.N.C.*)

The supply voltage may be represented as 354 sin ωt volts. Therefore

$$\text{Critical grid voltage} = -\frac{354}{20}\sin\omega t = -17{\cdot}7\sin\omega t$$

$$\text{Voltage applied to grid} = 141\sin(\omega t - 60^\circ)$$

These waves may be drawn as in Fig. 25.6(c) or, since in this case the control ratio is assumed to be constant, the ignition angle ϕ may be found by solving the equation

$$\begin{aligned} -17{\cdot}7\sin\phi &= 141\sin(\phi - 60^\circ) \\ &= 141\sin\phi\cos 60^\circ - 141\cos\phi\sin 60^\circ \\ &= 70{\cdot}7\sin\phi - 122\cos\phi \end{aligned}$$

whence $\tan\phi = 122/88{\cdot}4 = 1{\cdot}38$; therefore the ignition angle, ϕ, is 54°.

Since the voltage drop is to be neglected, the extinction angle, γ, may be taken to be π.

$$\text{Mean load voltage} = \frac{1}{2\pi}\int_{54^\circ}^{180^\circ} 354\sin\omega t\, d(\omega t) = 89{\cdot}5\,\text{V}$$

and

$$\text{Mean anode current} = \frac{89{\cdot}5}{1{,}500} = \underline{\underline{59{\cdot}5\,\text{mA}}}$$

25.5 Mercury-arc Rectifiers

In its single-phase form the mercury-arc rectifier consists of a graphite or carbon-coated iron anode and a mercury-pool cathode enclosed in an envelope from which all air has been removed. Mercury vapour fills the space between anode and cathode, the pressure of the vapour varying widely with temperature. Unlike the gas diode and triode, there is not an external power source to provide a supply of primary electrons. Thus conduction will not commence by merely raising the anode–cathode voltage above the ionizing potential for mercury vapour (10·6 V).

In one type of single-phase rectifier (called an *ignitron*) an auxiliary pointed electrode (*ignitor*) dips into the mercury pool. At the instant during the positive voltage cycle at which conduction to the anode should commence, a heavy short-duration current is passed through the ignitor to the mercury pool. This current raises the temperature at the point to a white heat so that some primary electrons are emitted. The primary electrons ionize the mercury vapour, whereupon positive ions travel back to the cathode. The action of the positive ions at the cathode is such that copious primary-electron emission develops at a particular spot, which becomes white hot. The mechanism by which electrons are emitted at the cathode spot is not clearly understood. It would seem that thermionic, secondary and field emission effects are all present. The conduction through the mercury vapour is of the same nature as in the gas diode, i.e. a positive space charge forms in front of the cathode and is joined to the anode by the plasma. The voltage drop between anode and cathode when the rectifier is conducting is somewhat greater than the ionization potential alone, since with the greater currents and greater anode–cathode separation there is an appreciable voltage drop across the plasma, and a voltage drop is also found to exist at the anode surface. The total voltage drop is usually between 20 and 30 V. Ignitrons capable of carrying currents of several thousand amperes have been constructed.

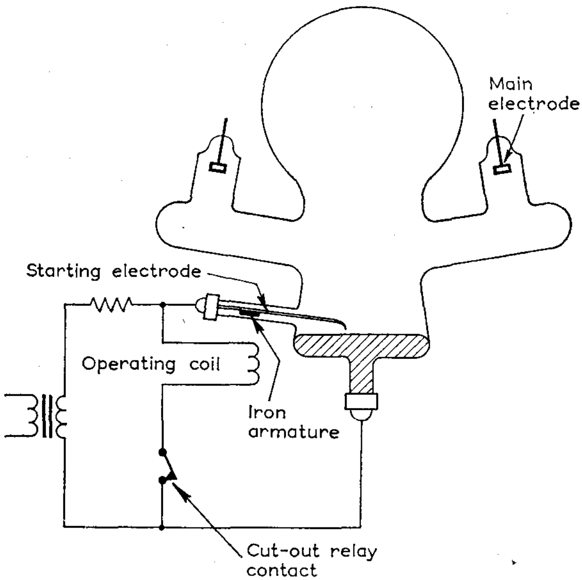


Fig. 25.7 BASIC MERCURY-ARC RECTIFIER

The mercury-arc rectifier has the advantage over the thermionic gas diode in that, though the mean current will naturally be limited by temperature rise as a whole, the peak current is almost unlimited and very high overloads may be tolerated for short periods. There is also the considerable advantage that, since the mercury cathode is almost indestructible, the life of the rectifier is exceedingly great.

The single-phase form of the rectifier is particularly useful for the control of welding currents. It will be noted that the mean current may be controlled by varying the instant at which conduction starts.

A glass-bulb permanently-evacuated mercury-arc rectifier is illustrated in Fig. 25.7. Two anodes only are shown, although rectifiers of this kind usually operate from a 3-phase supply and have

3, 6 or 12 main anodes with two auxiliary anodes. In this case the arc is continuous and does not require to be restruck each cycle. After conduction starts a relay cuts out the ignition system, which consists of a flexible electrode which is pulled into contact with the mercury pool by an electromagnet. When the flexible electrode contacts the mercury pool the electromagnet is short-circuited so that the flexible electrode springs back giving a spark as it breaks contact with the pool. This spark gives the initial electrons which are attracted by an anode at a positive potential and give rise to ionization of the mercury vapour.

In the glass-bulb rectifier the anodes are housed in arms which leave the bulb at acute angles. This is to prevent mercury globules forming on the anode surfaces. If these form, then *backfire* (reverse conduction) or *crossfire* (anode-to-anode conduction) may occur. Steel-tank rectifiers are also manufactured; in these the anodes are protected by insulating barriers. Another method of protection against crossfire and backfire is the use of grids controlling the instant at which an anode may start to conduct. The grids are in the form of a perforated metal sheet or a metal grill cutting off the anode from the rest of the rectifier. The grids are held at a negative potential except at the instant when the anode should start to conduct. The grids may also be used to reduce the conduction angle and hence the output voltage.

In addition to the main anodes which carry the load current there are usually two auxiliary anodes arranged to form a separate full-wave rectifier supplying a small dummy load. The purpose of this is to maintain the cathode hot spot should the main load current temporarily fall to zero. When the load is reimposed the main anodes will again conduct without the ignition system requiring to be separately operated.

Glass-bulb rectifiers are generally used for loads up to 500 A at 500 V. Permanently-evacuated steel-tank rectifiers are used for loads above this and up to 750 A at 750 V. Continuously-pumped steel-tank rectifiers are used for loads up to 3 MW at voltages exceeding 20 kV.

The internal efficiency of a rectifier largely depends on the voltage at which it operates since the voltage drop along the arc is constant.

Let V_{dc} and I_{dc} be the direct voltage (including arc drop) and current generated when the arc drop is V_{arc}. Then

$$\text{Efficiency, } \eta = \frac{V_{dc}I_{dc} - V_{arc}I_{dc}}{V_{dc}I_{dc}} \times \frac{100}{1} \text{ per cent}$$

$$= 1 - \frac{V_{arc}}{V_{dc}} \text{ p.u.} \qquad (25.2)$$

e.g. if the arc drop is 25 V the efficiency at 100 V output is only 75 per cent, but at 1,000 V it is 97·5 per cent.

A rectifier will, however, always require a transformer and various auxiliary circuits and the losses in these must also be taken into account.

25.6 Polyphase Rectification

The semiconductor diode rectifier is now replacing the mercury-arc rectifier for polyphase rectification in all applications except those involving the highest voltages. The forward voltage drop across a

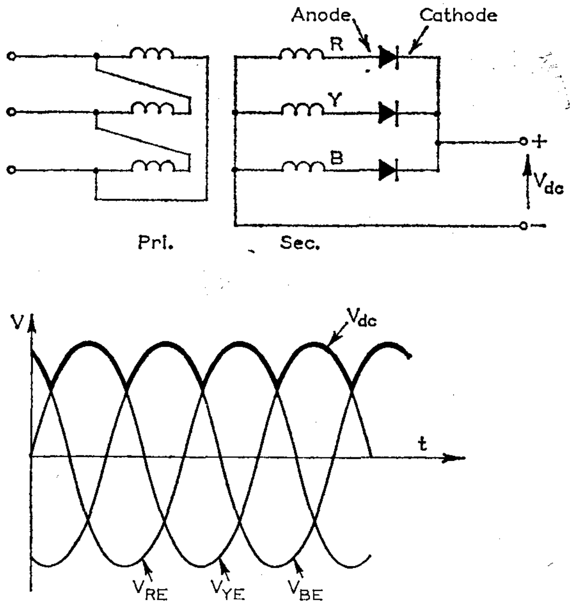


Fig. 25.8(a) SIMPLE 3-PHASE RECTIFIER CONNEXIONS

semiconductor diode is only a fraction of a volt, and hence the efficiency as defined by eqn. (25.2) will be almost unity (replacing V_{arc} by the diode forward voltage). Germanium diodes are used for high-current and silicon diodes for high-voltage applications.

By using polyphase rectification, the ripple voltage in the d.c. output may be made small without the necessity of using smoothing capacitors—this is a great advantage in power rectification.

The secondary of the transformer supplying a polyphase rectifier must have a neutral point when a mercury-arc rectifier (which has only one cathode) is used in order to give the negative output connexion. The primary winding is generally delta-connected to avoid the difficulties which arise with star-star transformers. The primary is fed from a 3-phase supply, and the secondary is connected to give 3, 6 or 12 phase operation as required. A delta-connected primary is shown in Fig. 25.8(*a*) and is assumed in Figs. 25.8(*b*) and (*c*). In these diagrams diodes are shown, but the connexions apply

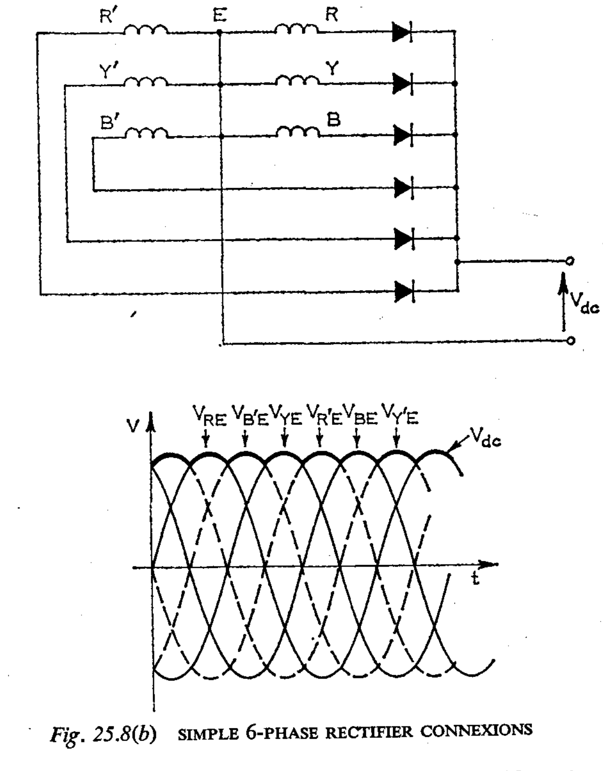


Fig. 25.8(b) SIMPLE 6-PHASE RECTIFIER CONNEXIONS

equally to 3- and 6-anode mercury-arc rectifiers in which case all the "cathodes" are common and are in fact the mercury pool.

A simple 3-phase circuit is shown in Fig. 25.8(*a*), together with the corresponding wave diagrams. Conduction takes place through whichever diode (or to whichever anode in a mercury-arc rectifier) has its anode at the highest positive potential. The common cathode potential will then be equal to the potential of the most positive anode less any internal voltage drop. (Remember that this voltage drop may be 20 to 30 V in a mercury-arc rectifier). The d.c. output is shown by the heavy line in the wave diagram. The ripple is seen to be considerably less than that in a full-wave rectifier.

A simple 6-phase connexion is shown in Fig. 25.8(c). The supply transformer has three centre-tapped secondary windings, the centre taps forming the neutral. The e.m.f.s of each side of one phase winding will be in antiphase, and the wave diagram shows that the

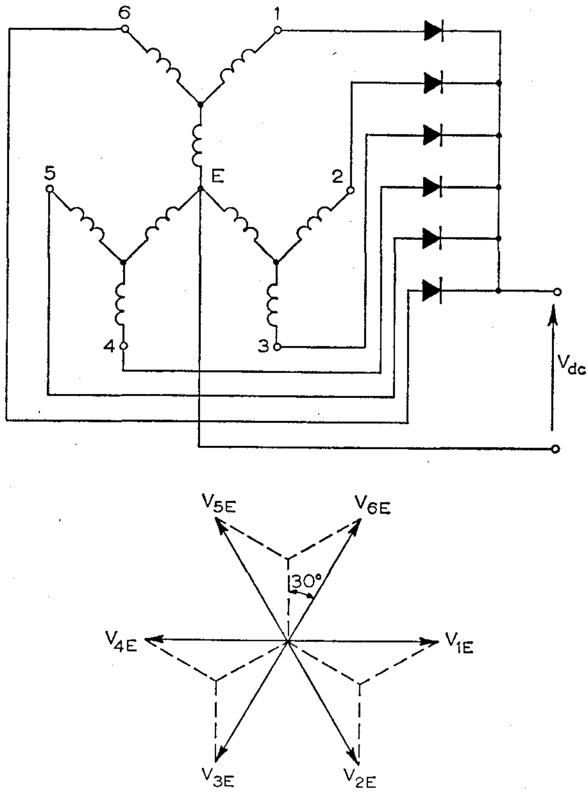


Fig. 25.8(c) ZIGZAG RECTIFIER CONNEXIONS

e.m.f.s on the diodes will reach their positive maxima in the sequence R, B′, Y, R′, B, Y′. Each diode conducts for one-sixth of a cycle. The primary currents will not be sinusoidal, since each diode passes a block of current for only one-sixth of a cycle. The output d.c. ripple is less than for the 3-phase connexion.

An alternative connexion, shown in Fig. 25.8(c), is the zigzag connexion. This gives a more sinusoidal form of transformer primary current than the simple 6-phase circuit, but maintains 6-phase smoothing. The complexor diagram shows how the symmetrical 6-phase system is developed. Note that the transformer primary current is distributed over two phases no matter which diode is conducting.

In Fig. 25.9 is shown a 3-phase bridge circuit which is commonly used with diode rectifiers. It cannot be used with a mercury-arc

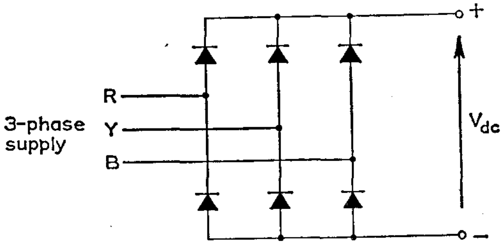


Fig. 25.9 THREE-PHASE BRIDGE RECTIFIER

rectifier since the cathodes are not common. This circuit does not require a star-connected secondary. The peak output direct voltage is the peak line voltage of the supply. Smoothing is 6-phase.

25.7 Voltage and Current Ratios

In polyphase rectifiers, conduction always occurs to the most positive anode. The "cathode" potential will therefore be the potential of the most positive anode minus the voltage drop V_a, across the rectifier

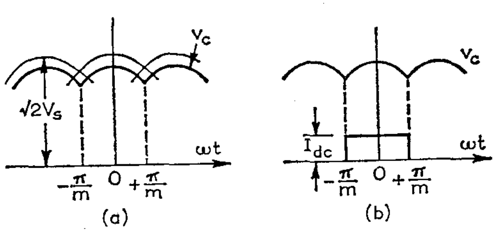


Fig. 25.10 WAVEFORMS IN POLYPHASE-RECTIFIER OPERATION

at any instant. Thus, as conduction commutates from anode to anode, the output voltage waveform for a 3-phase connexion will be as shown in Fig. 25.8(*a*). A 6-phase system could be dealt with similarly, and it will be seen that increasing the number of anodes reduces the ripple in the output voltage waveform.

Fig. 25.10 shows the general waveform for an *m*-phase rectifier, with the origin of the angular base at a peak value. Conduction to one anode will commence at $-\pi/m$ and be complete at $+\pi/m$.

Let V_s be the r.m.s. secondary output voltage per phase; then

$$\text{Output voltage} = \sqrt{2}V_s \cos \omega t - V_d \quad \text{for} \quad -\frac{\pi}{m} < \omega t < +\frac{\pi}{m}$$

Therefore

$$\text{Mean output voltage} = \frac{1}{2\pi/m}\int_{-\pi/m}^{+\pi/m} (\sqrt{2}V_s \cos \omega t - V_d)\,d(\omega t)$$

so that

$$\text{Mean output voltage, } V_{dc} = \sqrt{2}V_s \frac{\sin \pi/m}{\pi/m} - V_d \tag{25.3}$$

$$= E_{dc} - V_d$$

where

$$E_{dc} = \frac{\sqrt{2}V_s \sin \pi/m}{\pi/m} \tag{25.4}$$

From eqn. (25.4) it will be seen that E_{dc} increases with the number of anodes and tends towards $\sqrt{2}V_s$ as $m \to \infty$.

The current ratio is usually calculated on the assumption that the load circuit has a smoothing choke or at least sufficient inherent inductance to eliminate any ripple in the current waveform as shown at (*b*). It is also assumed that the full current instantaneously commutates from one diode to the next (see Section 25.8).

Let I_{dc} be the output current and I_s be the r.m.s. current at a transformer terminal; then

$$I_s = \sqrt{(\text{mean square of diode current})} = \sqrt{\left(\frac{1}{2\pi} I_{dc}^2 \frac{2\pi}{m}\right)}$$

whence

$$I_s = \frac{I_{dc}}{\sqrt{m}} \tag{25.5}$$

There is only one pulse of current through each diode per cycle. The rating of the transformer secondary winding is then given by

$$\text{Rating of secondary} = mV_sI_s \quad \text{volt-amperes} \tag{25.6}$$

The primary winding may have a rating somewhat less than this owing to the improved current waveforms, but this depends on the method of connexion.

Substituting for V_s and I_s from eqns. (25.4) and (25.5),

$$\text{Secondary rating} = m \frac{E_{dc}}{\sqrt{2}\sin \pi/m} \frac{\pi}{m} \frac{I_{dc}}{\sqrt{m}}$$

$$= \frac{\pi/m}{\sin \pi/m} E_{dc} I_{dc} \sqrt{\frac{m}{2}} \tag{25.7}$$

$E_{dc}I_{dc}$ is the actual power output (neglecting rectifier losses). Therefore

$$\frac{\text{Actual power}}{\text{Full-load rating}} = \frac{\sin \pi/m}{\pi/m} \sqrt{\frac{2}{m}} \tag{25.8}$$

This is termed the *utilization coefficient* of the transformer secondary. It may be shown that the utilization coefficient has its maximum value of 0·675 for a 3-anode rectifier and is only 0·4 for a 12-anode rectifier. A low utilization coefficient will increase the size and cost of the transformer. Thus, though a large number of diodes gives a smooth output, it also leads to an expensive transformer.

25.8 Overlap Effects

In deriving the previous voltage and current relationships it has been assumed that the commutation of the current from one anode to the next is accomplished in zero time, i.e. the full current instantaneously stops flowing to one anode and starts flowing to the next. This is, in fact, impossible since the transformer windings in the anode circuits must have some leakage inductance, through which it is not possible to have an instantaneous change of current.

Fig. 25.11(*a*) shows two successive phases of a secondary winding with the equivalent reactance represented as external to the windings. The resistance may usually be neglected. Suppose that diode 1 is first conducting and that the full current I_{dc} flows to it. When the potential of diode 2 reaches that of diode 1 the current to diode 1 cannot immediately cease but must decay over a finite time while the current to diode 2 increases from zero to I_{dc}. Thus for a short period the current must split between diodes 1 and 2, and for this period the anodes of these diodes must have the same potential. This is called *overlap*. During the overlap period the output e.m.f. is the mean of the e.m.f.s of the two secondary phases alone as shown at (*b*). The length of time during which overlap takes place and the reduction of the mean output voltage due to it are proportional to the load current.

Since the current through the transformer secondary windings has no longer the rectangular waveform previously assumed, the relationship between the r.m.s. anode current and the direct load current will be modified from that given by eqn. (25.5). For 3-phase rectification the modification is negligible, and may often be neglected

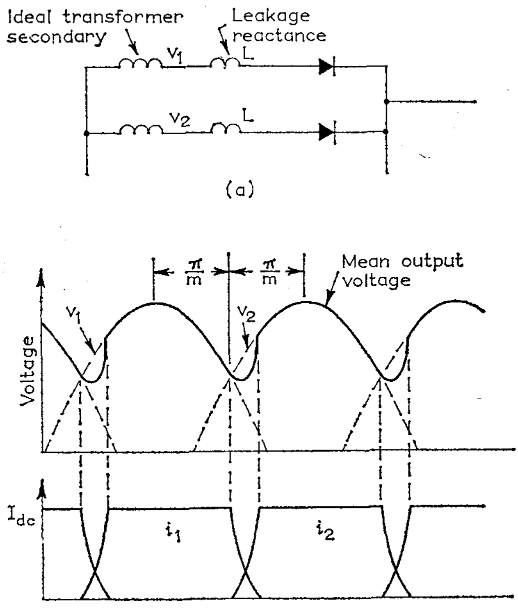


Fig. 25.11 EFFECT OF OVERLAP

for 6-phase rectification. For 12-phase rectification the modification is considerable.

25.9 Double Three-phase Connexion

By the addition of an *interphase reactor** to the simple 6-phase connexion it is possible to obtain rectifier action with the smoothness of normal 6-phase rectification and the utilization factor of 3-phase rectification. The principal characteristic of the double 3-phase connexion is that conduction is made to split into two parts at all instants of the cycle. The basic principle is similar to that of overlap where the transformer leakage inductance causes the current to split into two parts at each commutating point.

* Also called *interphase transformer*.

A centre-tapped iron-cored choke is used as shown in the circuit diagram of Fig. 25.12. Since the sides of the interphase winding have equal numbers of turns and are wound round the same core in the same direction, the e.m.f.s induced in the two halves of the winding will be equal in magnitude and in direction at all instants. Neglecting resistance and leakage reactance, the potential differences across the two halves will then be equal in magnitude and direction at all instants. It is to be particularly noted that successive anodes are connected on opposite sides of the interphase reactor, so that

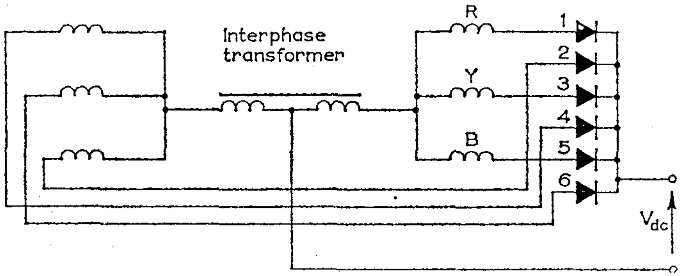


Fig. 25.12 INTERPHASE REACTOR OPERATION

if ordinary 6-phase operation occurred, the direct current would alternate between the sides of the interphase reactor.

Suppose that the whole current flows to anode 1 and attempts to commutate naturally to anode 2. This means that the direct current should cease to flow in the right-hand side of the interphase reactor and should suddenly start to flow in the left-hand side (Fig. 25.12). The inductance of the interphase reactor will react against these changes of current: the right-hand terminal of the reactor will be driven positive with respect to earth potential while the left-hand terminal becomes negative. This raises the potential of anodes 1, 3 and 5 while lowering that of anodes 2, 4 and 6. Provided that the inductance of the interphase reactor is sufficient, the current will not wholly transfer from anode 1 to anode 2, for the raising of the potential of the right-hand side and the lowering of the potential of the left-hand side will maintain anodes 1 and 2 at the same instantaneous potential with respect to earth. The instantaneous potentials are illustrated in Fig. 25.13(*a*). Eventually the potential of anode 3 will exceed that of anode 1 and the current which flows to anode 1 will transfer to anode 3, since this does not involve a change of current in the interphase reactor (there may be some delay due to simple overlap). The current is now shared between anodes 2 and 3. The potential of anode 3 will eventually tend to exceed that of anode 2 and this will cause the interphase reactor e.m.f.s to reverse, so that the potential of the left-hand side is raised and that of the right-hand side is lowered.

The current continues to be shared between anodes 2 and 3 until the current flowing to anode 2 transfers to anode 4. This is illustrated in Figs. 25.13(*b*)–(*d*). The process will continue in a similar manner round all the anodes.

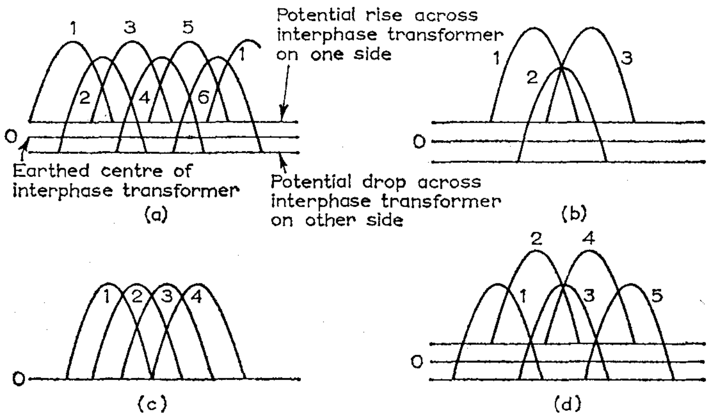


Fig. 25.13 WAVEFORMS WITH INTERPHASE REACTOR

It will be seen that there are always two successive anodes conducting and that the mean current in each half of the interphase reactor will be half the output current, i.e. $I_{dc}/2$. It should also be noted that the mean currents in each part of the interphase reactor are in opposite directions and will produce no net magnetizing effect in the iron core.

The output voltage at any instant is the mean of the e.m.f.s of the conducting secondary windings at the same instant. Fig. 25.14

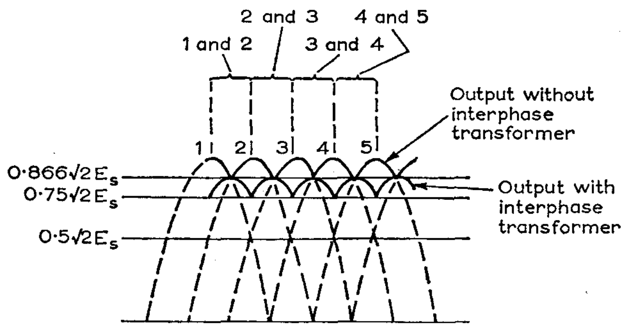


Fig. 25.14 OUTPUT VOLTAGE WITH INTERPHASE CONNEXION

shows the e.m.f.s of the secondary windings, and also the waveform of the average voltage of successive pairs of anodes as these pairs share the conduction. This latter waveform is also the waveform of the effective output voltage; the output voltage ripple obviously corresponds to 6-phase rectification.

If the load current is less than a critical value which depends on the inductance of the interphase reactor, then the operation will be almost normal 6-phase operation. As the load is increased the transition between normal 6-phase and double 3-phase operation is marked by a sudden sharp decrease in the output voltage from the 6- to the 3-phase value. Often a small permanent load is connected across the output of the rectifier so that I_{dc} does not fall below the critical value.

Exact analysis shows that the mean output voltage (neglecting overlap) is $0{\cdot}826\sqrt{2}E_s$, which happens to be exactly the output expected for a 3-phase rectifier where the r.m.s. secondary e.m.f. is E_s. Therefore

$$\text{Mean output voltage} = \sqrt{2}E_s\frac{\sin\pi/3}{\pi/3} - V_d \tag{25.9}$$

It will also be seen that each anode and each secondary winding carries a current $I_{dc}/2$ for one-third of a cycle. Thus

$$\text{R.M.S. current per secondary phase} = \frac{I_{dc}}{2}\frac{1}{\sqrt{3}} \tag{25.10}$$

This gives the arrangement the same utilization coefficient as is normally obtained for 3-phase rectification.

25.10 Primary Current Waveforms

In the following examination of waveforms, the effect of overlap on the current waveform will be neglected for the sake of clarity. In general, the overlap effect will give rise to smoother primary current waveforms than are obtained when the effect of overlap is neglected.

DELTA PRIMARY, SIX-PHASE SIMPLE SECONDARY (Fig. 25.15(*a*))

Fig. 25.15(*b*) shows the current waveforms for the primary phases and lines. It will be realized that current flowing in the secondary to anode 1 will have a corresponding current in the primary phase 1, and current flowing in the secondary to anode 4 will have a

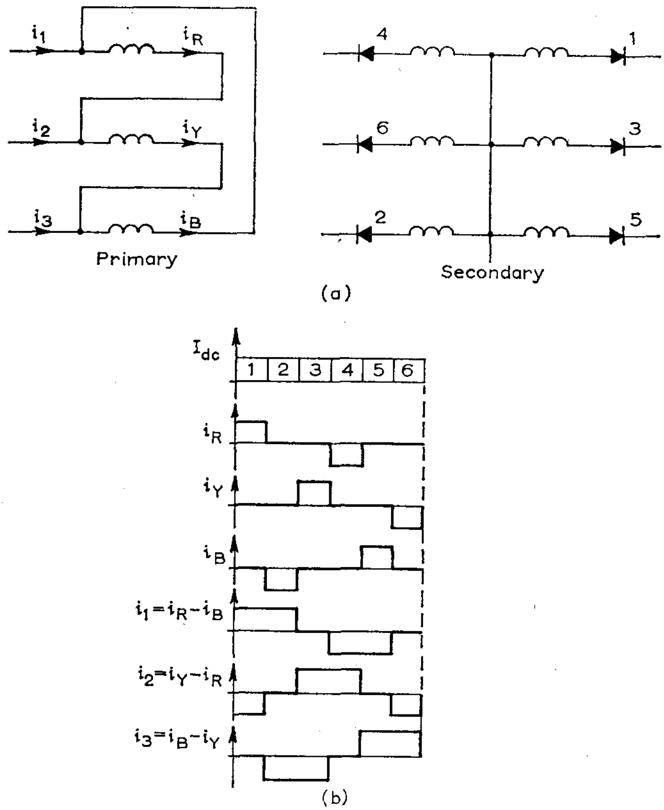


Fig. 25.15 PRIMARY CURRENTS FOR A DELTA-CONNECTED PRIMARY AND SIMPLE 6-PHASE OPERATION

corresponding current in the same primary phase, but in the reverse direction. The currents in the other phases will have similar effects.

The primary phases each carry current for one-third of a cycle, and hence the primary utilization factor will be better than the secondary utilization factor.

DELTA PRIMARY, DOUBLE THREE-PHASE SECONDARY

Each secondary phase carries current over one-third of a cycle (see Section 25.9). The primary waveforms are shown in Fig. 25.16. It will be seen that each primary phase now carries current over two-thirds of a cycle so that the utilization factor is still further improved.

The primary line currents are also more nearly sinusoidal so that there will be less harmonic current in the lines.

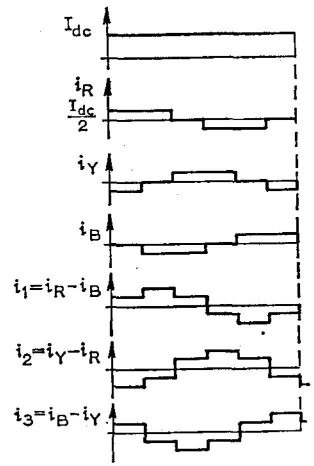


Fig. 25.16 PRIMARY CURRENTS WITH INTERPHASE CONNEXION

25.11 The Thyristor

The thyristor is a four-layer silicon semiconductor junction device, which has characteristics similar to those of a gas-filled triode. It can be triggered into the conducting state, and conducts unidirectionally until the voltage across it is reduced to almost zero. The voltage drop across it in the conducting state is small. This device is also known by the term silicon controlled rectifier (s.c.r.). Important characteristics of the thyristor are the extreme rapidity with which the device can be turned on and the very fast switch-off time.

Fig. 25.17(*a*) is a schematic drawing of the thyristor together with its circuit symbol. The operation can be deduced by considering that the four layers form two transistors and a diode as shown at (*b*). If reverse bias is applied (i.e. if the "anode" is connected to the negative of the supply), then junctions 1 and 3 are reverse biased, and only the small reverse-bias saturation current will flow until reverse breakdown occurs. This is almost independent of any voltage on the *gate* terminal.

Consider the gate connexion open-circuited, and a positive bias applied (i.e. the anode positive). Then junctions 1 and 3 will be forward biased and junction 2 will be reverse biased. The first three layers form a p–n–p transistor, with junction 1 forming an emitter–base junction. Hence holes injected across junction 1 will be collected across the reverse-biased junction 2 by normal transistor action to give a hole current $\alpha_1 I_h$ as shown at (*b*). In the same way, the last three layers form an n–p–n transistor, in which junction 3 forms the emitter–base junction. Electrons injected from right to left across junction 3 are collected across the reverse-biased junction 2. At the same time junction 2 acts as a reverse-biased diode junction, and

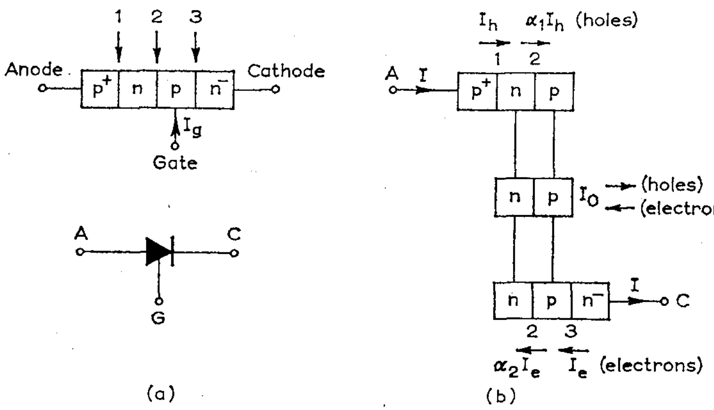


Fig. 25.17 THE TWO-TRANSISTOR REPRESENTATION OF A THYRISTOR

the reverse bias saturation current I_0 (consisting of holes moving from n to p and electrons from p to n) flows.

If I_h is the injected hole current across junction 1 (assuming that the p-layer is heavily doped), then $\alpha_1 I_h$ is the fraction of this flowing across junction 2. Also, if I_e is the electron current flowing from right to left across junction 3, then a fraction $\alpha_2 I_e$ is collected across junction 2. The total current across 2 is therefore

$$I_0 + \alpha_1 I_h + \alpha_2 I_e$$

flowing conventionally left to right. But by Kirchhoff's law, there is the same current across each junction, and hence across 1 and 3, $I_h = I_e = I$ and across 2, $I = I_0 + \alpha_1 I + \alpha_2 I$, or

$$I = \frac{I_0}{1 - (\alpha_1 + \alpha_2)} \tag{25.11}$$

By suitable choice of doping, the current gain, α, of silicon transistors can be made to have a low value at very small emitter currents, the gain rising towards unity as the current increases.

For small values of applied voltage α_1 and α_2 can both be well below 0·5 so that the resultant current is small. As the applied voltage increases $(\alpha_1 + \alpha_2)$ can become unity, and eqn. (25.11) shows that in this condition I increases without limit (hence a current-limiting resistor must be included in the circuit). When this happens junction 2 breaks down, and the voltage across the thyristor falls to the small constant breakdown voltage across junction 2 plus the even smaller voltage drops across the bulk resistance of the layers and the forward-biased junctions 1 and 3.

The characteristic is shown in Fig. 25.18(*a*). Between 0 and A the term $(\alpha_1 + \alpha_2)$ is less than unity, but is increasing so that I

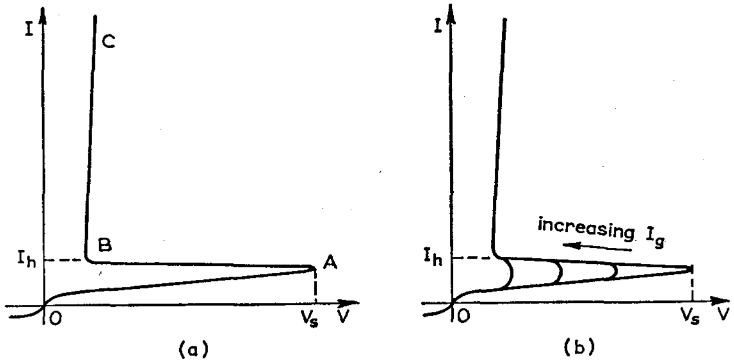


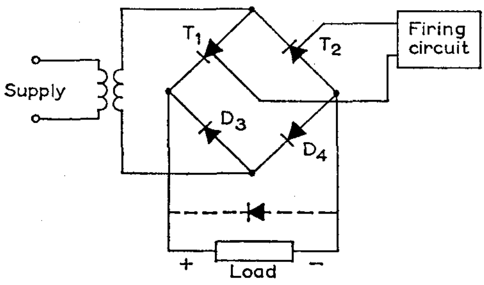
Fig. 25.18 CURRENT/VOLTAGE CHARACTERISTIC OF A THYRISTOR
(*a*) Gate current zero
(*b*) Effect of increasing gate current

increases. At A $(\alpha_1 + \alpha_2) = 1$ and breakdown occurs—the voltage falls abruptly (AB). Breakdown conditions will persist so long as the anode current is above the holding value I_h, giving an almost constant voltage across the device, irrespective of the current (BC on the characteristic).

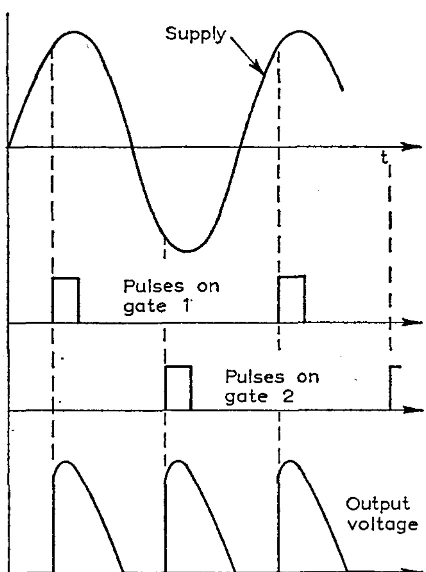
If a current is now fed into the gate connexion, this increases the current across junction 3, so that α_2 increases and $(\alpha_1 + \alpha_2)$ now becomes equal to unity at a lower value of applied anode–cathode voltage, as shown at (*b*). If the applied voltage is less than the switching voltage, V_s, at which breakdown occurs when there is no gate current, then breakdown will not occur until a gate current flows. In this way the thyristor can be made to act as a switch.

In order to avoid undue heat dissipation at the gate it is usual to feed the gate with a pulse waveform. When an a.c. supply is connected across the thyristor, the instant in the positive half-cycle at which the device "fires" is controlled by the time at which the gate pulse is applied. A typical thyristor bridge rectifier circuit and waveforms are shown in Fig. 25.19.

Care must be taken in thyristor circuits when the load is inductive since high reverse voltages may then appear across the rectifier when the load is switched off. Sometimes an avalanche diode is connected with reverse polarity across the load, as shown by the broken line in Fig. 25.19, in order to minimize this effect.



(a)



(b)

Fig. 25.19 THYRISTOR-CONTROLLED BRIDGE RECTIFIER

25.12 Grid-controlled Mercury-arc Rectification

Grid control may be applied to polyphase mercury-arc rectification in the same way as it was applied to single-phase rectifier circuits (Section 25.3). The electron path to an anode may be blocked by a negatively biased grid so that conduction to a given anode may not

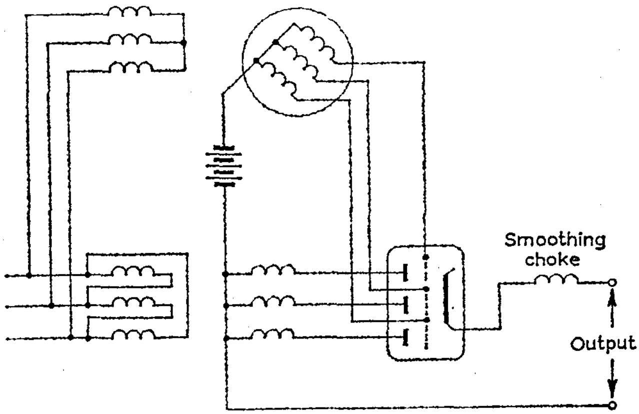


Fig. 25.20 GRID-CONTROLLED MERCURY-ARC RECTIFIER

commence until the negative grid potential is reduced. Usually phase shift control with an additional steady negative bias is used. A variable-phase voltage for grid control may be obtained from a phase-shifting transformer (which is constructed like an induction motor), with a 3-phase stator winding and a stationary rotor wound for the number of phases corresponding to the number of anodes. By varying the angular position of the rotor the grid voltage may be given any desired phase relationship to the supply voltage.

A connexion diagram for 3-phase grid-controlled rectification is shown in Fig. 25.20, a battery being shown as providing the steady negative grid bias where normally a single-phase metal rectifier circuit would be employed.

25.13 Controlled Polyphase Rectification–Output Voltage

The output voltage waveform for an m-phase rectifier in which grid or gate control has delayed the commutation by a phase angle α is shown in Fig. 25.21. Taking a voltage maximum as the reference zero, the angle of conduction per phase will be seen to be from $(-\pi/m + \alpha)$ to $(+\pi/m + \alpha)$. Therefore

$$\text{Mean output voltage} = \frac{1}{2\pi/m} \int_{-\frac{\pi}{m}+\alpha}^{\frac{\pi}{m}+\alpha} (\sqrt{2}E_s \cos \omega t - V_d)d(\omega t)$$

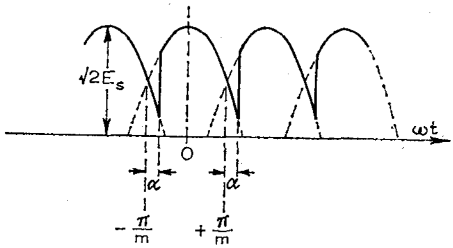


Fig. 25.21 CONTROLLED POLYPHASE RECTIFICATION WAVEFORMS

where V_d is the drop across the rectifier.

$$V_{out} = \frac{m}{2\pi}\sqrt{2}E_s\left[\sin\left(\frac{\pi}{m}+\alpha\right) - \sin\left(-\frac{\pi}{m}+\alpha\right)\right] - V_d$$

$$= \sqrt{2}E_s \frac{m}{\pi} \sin \frac{\pi}{m} \cos \alpha - V_d$$

$$= \sqrt{2}E_s \frac{\sin \pi/m}{\pi/m} \cos \alpha - V_d \qquad (25.12)$$

This equation, however, ceases to apply if the delay angle α is greater than $(\pi/2 - \pi/m)$, for in this case the current conduction becomes discontinuous. The waveform is shown in Fig. 25.22.

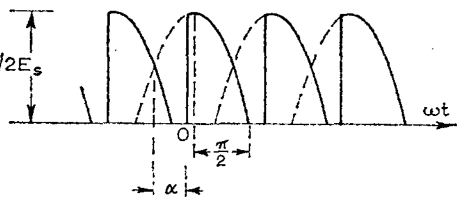


Fig. 25.22 DISCONTINUOUS CONDUCTION WITH GRID OR GATE CONTROL

For this case,

$$\text{Mean output voltage} = \frac{1}{2\pi/m}\int_{-\frac{\pi}{m}+\alpha}^{\frac{\pi}{2}} (\sqrt{2}E_s \cos \omega t - V_d)d(\omega t) \qquad (25.13)$$

25.14 Invertor Operation

A controlled polyphase rectifier may be used to link a d.c. system to an a.c. system so that energy flows from the d.c. system to the a.c. system. In this case the rectifier is called an *invertor*. This operation will only be considered briefly.

The a.c. system must have at least one synchronous generator connected to it to determine the frequency of operation. The d.c.

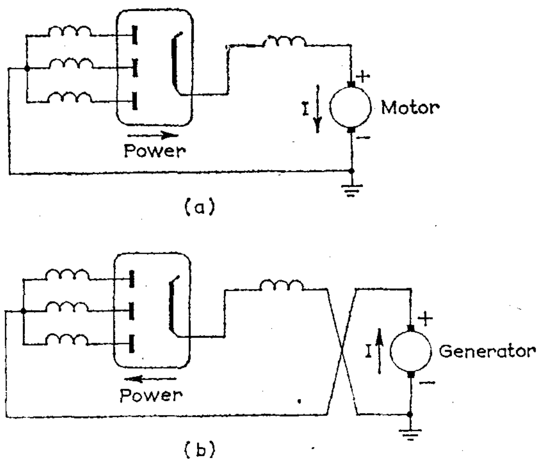


Fig. 25.23 INVERTED OPERATION OF A MERCURY-ARC RECTIFIER
(*a*) Rectifier (*b*) Inversion

system must, naturally, include a d.c. source of energy. It should be first noted that, when a rotating machine changes from motoring to generating operation, the current direction rather than the e.m.f. direction changes. With a rectifier, however, it is not possible for the current direction to change and thus the e.m.f. direction should be changed. (This illustrated in Fig. 25.23.) With respect to the a.c. system, it will be realized that, if current conduction during a positive voltage half-cycle led to the transmission of power from the a.c. system to the d.c. system, then for power transmission in the opposite sense and with the same direction of current conduction it will be necessary to have the current conduction during a negative voltage half-cycle. This is arranged by the connexion of the generator and by using grid control to delay conduction to a negative half-cycle.

As shown in Fig. 25.23 the generator voltage raises the potential of the transformer neutral to a positive value with respect to earth potential. The anode potentials are then the d.c. potential plus the e.m.f. of the corresponding secondary phase. The anode potentials are shown in Fig. 25.24. The direct voltage of the generator is assumed to be approximately equal to the output voltage of the same rectifier if used to deliver power from the same a.c. system. This is usually approximately the case but is not necessarily so. Except for the p.d. across the choke the cathode of the rectifier is at

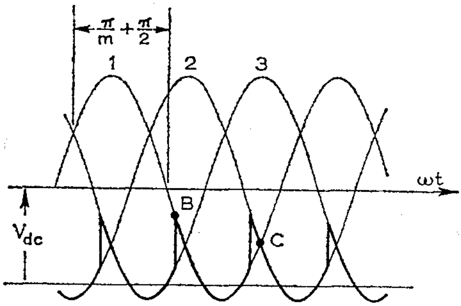


Fig. 25.24 WAVEFORMS FOR INVERTED OPERATION

earth potential, and the anodes are positive with respect to the cathode over most of the cycle.

It is essential that commutation be delayed over at least $(\pi/m + \pi/2)$ radians so that conduction to a given anode will only occur when the secondary e.m.f. is negative. If conduction occurs when the secondary e.m.f. is positive, there will be rectifier action and short-circuit conditions will ensue. Conduction may be permitted to start at a point such as B in the alternating voltage cycle of each anode. It is essential also that commutation occurs before the point C is reached; otherwise it will not occur at all as the anode "carrying" the arc will become more positive than the succeeding anode after the point C is reached. If commutation to the succeeding anode does not occur, the arc will remain with an anode whose alternating potential will become positive, giving rectification and short-circuit conditions. Thus the phase angle of the grid may only be varied over a limited range.

Neglecting the arc drop, the cathode potential becomes the potential of the conducting anode at each instant. The difference between the cathode potential and earth potential is the p.d. across the smoothing choke.

It should be remembered that with a constant d.c. system voltage and a constant a.c. system voltage the grid potential will control the current flow between the two systems.

Exactly similar considerations apply in the case of thyristors.

D.C. links (with a rectifier at one end and an invertor at the other end of the link) are now commonly used to connect two a.c. power distribution systems together. By doing this the following advantages are obtained:

(*a*) There is no need to synchronize the two a.c. systems.

(*b*) Only *power* flow takes place in the d.c. link (i.e. there is no reactive flow through the link).

Such links are used to connect the British grid to the European power network, and are also used to interconnect parts of the British grid system.

PROBLEMS

25.1 A single gas diode is to be used to charge a 100 V battery of negligible internal resistance from a 250 V 50 Hz supply. The voltage drop and ignition voltage across the gas diode may be taken as 20 V. Find the series resistance so that the average charging current is 2 A.

Ans. 28 Ω.

25.2 An a.c. supply at 20 V is connected in series with a full-wave rectifier to a 12 V battery. The total resistance of the circuit is 5 Ω in the conducting direction and infinity in the reverse direction. Plot the current waveform and determine the mean value of the current in the circuit. (*L.U.*)

Ans. 0·7 A.

25.3 A thyratron with a control ratio of 20 is to be used as a half-wave rectifier from a 250 V 50 Hz supply. Plot a curve of mean output voltage to a base of grid bias voltage for a variation of grid voltage between 0 and −20 V. The ignition voltage and voltage drop may be taken as 20 V (assume a pure resistance load).

25.4 A thyratron used in a grid-controlled rectifier has a voltage drop of 25 V when conducting, this being also the extinction voltage. Its control ratio is 30. An alternating p.d. of 212 V (r.m.s.) is applied, via a load resistance between anode and cathode, and a p.d. of 10·6 V (r.m.s.) lagging a quarter of a cycle behind the anode p.d. is applied between the grid and cathode. During what fraction of the alternating voltage cycle does the valve conduct? What is the mean anode current if the resistance is 100 Ω?

Ans. 0·34; 0·69 A.

25.5 A gas-filled triode, having a control ratio of 15 and negligible tube drop, is used as a controlled rectifier on a 180 V 500 Hz supply. Find the mean current in a load of 1,000 Ω resistance when a bias voltage of −10 V is applied to the valve grid. (*H.N.C.*)

Ans. 73 mA.

25.6 A gas-filled triode has a control ratio of 14·1 and operates from a 250 V a.c. supply. The load resistance is 150 Ω. Calculate the mean current in the load if the grid voltage consists of a steady voltage of −20 V superimposed on an alternating voltage of r.m.s. value 70·7 V which lags behind the anode voltage by 90°. Neglect valve voltage drop. (*H.N.C.*)

Ans. 0·395 A.

25.7 A mercury-arc rectifier has an arc drop of 25 V. What will be its internal efficiency when giving (*a*) 250 V, (*b*) 2,000 V on the d.c. side?

Ans. 91 per cent; 98·8 per cent.

25.8 A 250 V 400 A 6-anode mercury-arc rectifier operates from a transformer star-connected on the secondary side. Ignoring arc drop and impedance, calculate (*a*) the transformer secondary voltage, (*b*) the r.m.s. anode current, and (*c*) the rating of the secondary winding.

Ans. 185 V; 163 A; 181 kVA.

25.9 A 6-anode mercury-arc rectifier is to be supplied from 3-phase mains. Discuss the possible transformer arrangements and compare their relative advantages.

In a particular case the 800-turn primary windings are connected in delta and supplied at 6·6 kV; the direct voltage is 480 V and the arc drop is 25 V. Neglecting other voltage drops, determine the requisite number of secondary turns for six-phase star connexion. (*L.U.*)

Ans. 45.

25.10 Explain the operation of a six-anode mercury-arc rectifier with an interphase reactor. If the output is 300 V and 100 A, determine the secondary current and voltage for the supply transformer. Neglect overlap and assume an arc voltage drop of 20 V. If the transformer is delta connected, sketch and explain the theoretical primary current waveform. (*H.N.C.*)

Ans. 28·9; 278 V.

25.11 In the 3-phase bridge rectifier of Fig. 25.8, the three top diodes are replaced by thyristors. The 3-phase line voltage is 400 V, and pulses are applied to the gates to delay conduction by an angle of 25°. Determine the mean output voltage (d.c.), and sketch the line current waveforms. The thyristor voltage drop may be neglected.

Ans. 425 V.

25.12 Repeat Problem 25.11 for a delay angle, α, of 35°.

Ans. 384 V.

25.13 A thyristor bridge rectifier, as shown in Fig. 25.18 is used to supply a load of 100 Ω resistance. The a.c. supply voltage is 40 V r.m.s., and the holding current of the thyristor is 35 mA. Determine (*a*) the extinction angle, and (*b*) the mean load voltage when the gate firing pulses are adjusted to give a firing delay of 90°. The thyristor voltage drop may be assumed constant at 1·2 V.

Ans. 175°; 17·3 V.

Hint. Remember that conduction is discontinuous.

Chapter 26

FIELD-EFFECT TRANSISTORS

Field-effect transistors, or FETS, are a group of semiconductor devices that complement, and in some cases may replace, bipolar transistors. Although they have been known in principle for some time, production on a commercial scale has become possible only because of recent advances in metallurgical technology. The FET is so called because it uses a transverse electric field to modulate the longitudinal current.

In some ways FETS have characteristics that closely resemble those of thermionic valves—for example, both valves and FETS have very high input resistances. However, in common with bipolar transistors and unlike valves, FETS require no heater current and can be obtained in complementary forms. They may have a very high frequency response. In addition, at zero bias, they can be either cut off or conducting depending on type. They have very high current and power gains, and because they are majority carrier devices (bipolar transistors are, of course, basically minority carrier devices) they have a high resistance to radiation damage and are less temperature dependent. In particular FETS are very suited to integrated circuit technology.

This chapter will briefly review the two main types of FET and will give some circuit applications.

26.1 The Junction-gate Field-effect Transistor

The junction-gate field-effect transistor, or JUGFET, was the first commercially available type of FET. It can be fabricated by either alloy junction or planar techniques. An n-channel alloy device is shown schematically in Fig. 26.1(a). It consists of a thin wafer of n-type

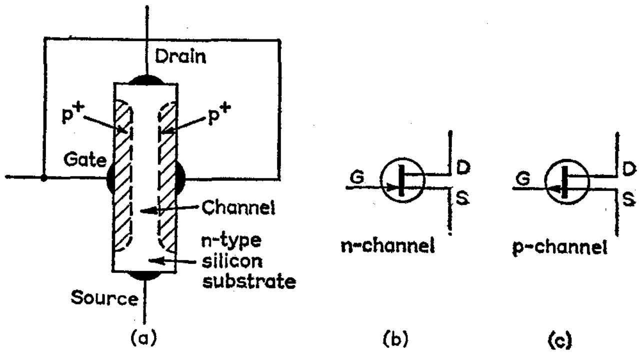


Fig. 26.1 THE JUGFET, WITH GRAPHICAL SYMBOLS

silicon into the opposite faces of which are diffused two very heavily doped p-type regions which are normally connected together to form the *gate* or control electrode. Ohmic connexions called the *drain* and *source* are made at the ends of the wafer for the load current. The gate, source and drain correspond to the base, emitter and collector of a bipolar transistor, or to the grid, cathode and anode of a thermionic triode.

With no external voltages the usual potential barrier builds up across the two p–n junctions and depletion layers are formed. Because the doping of the p-regions is much greater than that of the n-type wafer the depletion layers are mainly in the n-region. The drain and source are connected by the relatively narrow n-type *channel* and are separated from the gate connexions by the depletion layers. A similar device is possible using a p-type wafer and n-type gate connexions, giving a p-channel type. The graphical symbols are shown at (b) and (c) for an n-channel and a p-channel FET respectively.

In operation the gates are reverse biased with respect to source and drain, so that the gate current is only the very small reverse-bias saturation current. The reverse bias voltage extends the width of the depletion layers and hence narrows the conducting channel and so controls the drain-to-source current. This current is, of course, mainly a majority carrier current for the basic silicon wafer (electrons in the n-channel and holes in the p-channel type). The shape of the depletion layer is shown in Fig. 26.2 for a biased n-channel JUGFET.

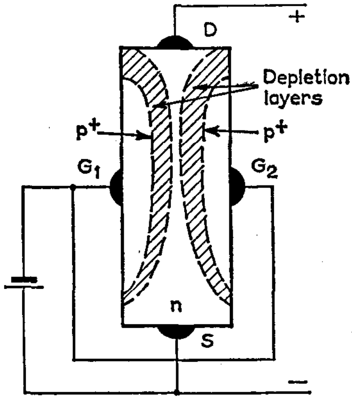


Fig. 26.2 AN n-CHANNEL JUGFET WITH BIAS

Since D is more positive than S the reverse bias across the gate–channel junctions is greater towards the drain end than at the source end. This explains why the channel is narrower at the drain end than at the source end, as shown.

The drain characteristics of an n-channel JUGFET are shown in Fig. 26.3(a) where the drain current I_D is drawn as a function of

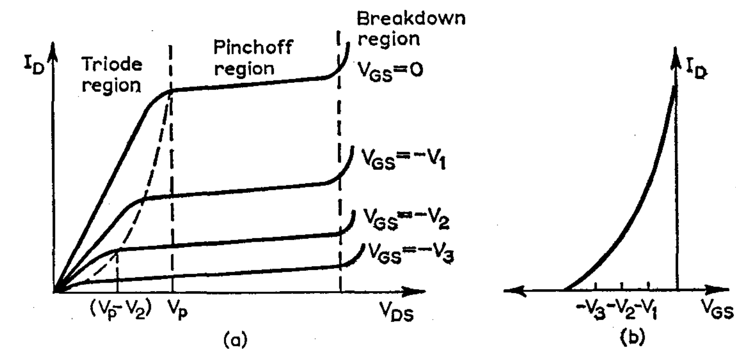


Fig. 26.3 CHARACTERISTICS OF AN n-CHANNEL JUGFET

(a) Drain characteristic
(b) Transfer characteristic

drain–source voltage, V_{DS}, for various fixed values of gate–source voltage, V_{GS}. The characteristics can be divided into three regions—the triode region, the pinchoff region and the breakdown region.

Consider $V_{GS} = 0$. As V_{DS} is increased the channel current, I_D, increases almost linearly. However, since the drain is now positive, while the gate remains at zero potential, a reverse bias builds up between the gate and the drain and the depletion layer takes the form shown in Fig. 26.2. As V_{DS} is further increased a point is reached when the two depletion layers almost meet. The voltage at which this occurs is called the *pinchoff voltage*, V_p. Further increase in V_{DS} causes practically no further increase in I_D—all that happens is that the depletion layers almost meet over a longer distance and this gives rise to an increase in the channel resistance in proportion to the increase in voltage. If V_{DS} continues to increase a point is reached where breakdown of the channel occurs and the current may rise to destructively high values.

At any negative value of V_{GS}, the depletion layers will already be closer together than for $V_{GS} = 0$, so that the initial slope of the curve in the triode region is less and pinchoff will occur at a value of V_{DS} lower than V_p by about the value of gate bias applied.

The size of the pinchoff voltage depends on the thickness of the initial channel. For relatively thick channels V_p will be large and drain voltages will be of the same order as in pentode valve circuits. For thin channels V_p will be low, and operating voltages will correspond to those of bipolar transistors. Minimum values of V_p are around 0·5 V, but normally JUGFETS have much higher values of V_p. The transfer characteristic of a JUGFET at a fixed value of V_{DS} in the pinchoff region is shown in Fig. 26.3(*b*).

In the triode region of the JUGFET characteristics the curves of I_D against V_{DS} are reasonably linear, with a slope that depends on the applied gate–source voltage. The characteristics apply for negative as well as for positive values of V_{DS} (Fig. 26.4) as long as the negative

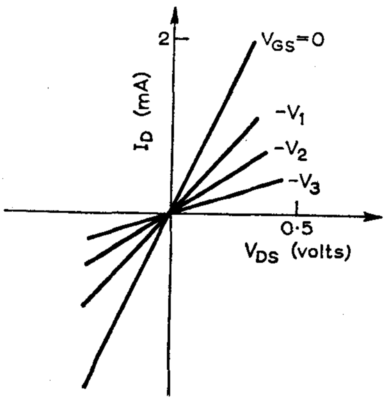


Fig. 26.4 THE FET CAN BE USED AS A VOLTAGE-CONTROLLED RESISTOR

value is not large enough to cause appreciable conduction of the gate–drain diode. In silicon JUGFETs this negative voltage will be approximately 0·6 V. Over the range of voltages of ±0·6 V, therefore, a JUGFET may be used as a linear resistor, and this is often done in integrated circuits. In addition, the resistance will vary with the bias voltage. This latter effect can be used in automatic gain control circuits in amplifiers, and amplitude control circuits in oscillators.

26.2 The Insulated-gate Field-effect Transistor

As with the JUGFET, the insulated-gate field-effect transistor, or IGFET, can be constructed with either an n-type or a p-type channel. In addition, it can be designed so that it conducts only when forward gate–source bias is applied (*enhancement mode*) or conducts with zero bias (*depletion mode*). Because of its construction, this type of FET is often called a MOST (metal-oxide-semiconductor transistor).

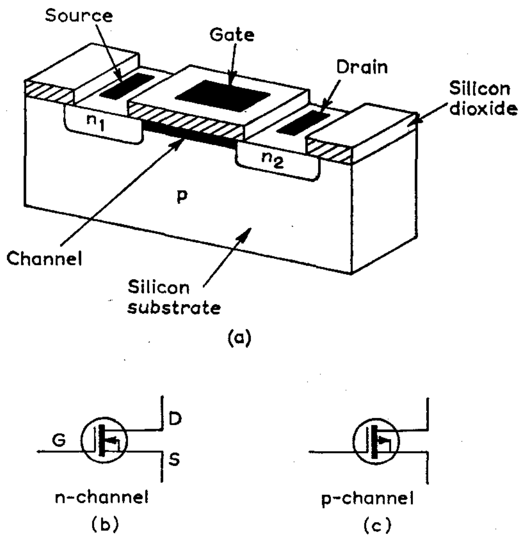


Fig. 26.5 TYPICAL CONSTRUCTION AND GRAPHICAL SYMBOLS FOR AN IGFET

An enhancement-mode n-channel IGFET is shown in Fig. 26.5(a). Two heavily doped n-regions, n_1 and n_2, are diffused into a lightly doped p-type substrate. A thin ($\sim$ 0·2 μm) *insulating* layer of silicon dioxide is grown over the surface, with gaps to allow the ohmic connexions to be made to n_1 and n_2. These source and drain "windows" are etched by photolithographic techniques. The gate connexion is a metallic layer (often aluminium) on the silicon dioxide that lies between the two *n*-regions. Similar metallic layers on n_1 and n_2 form the source and drain connexions.

With no gate voltage ($V_{GS} = 0$) the two *n*-regions are separated by a *p*-type region, typically 5 μm long. Hence for any voltage applied between drain and source there will always be two opposing *p–n* junctions in series and negligible current will flow no matter what polarity is applied between D and S. If the gate is made positive with

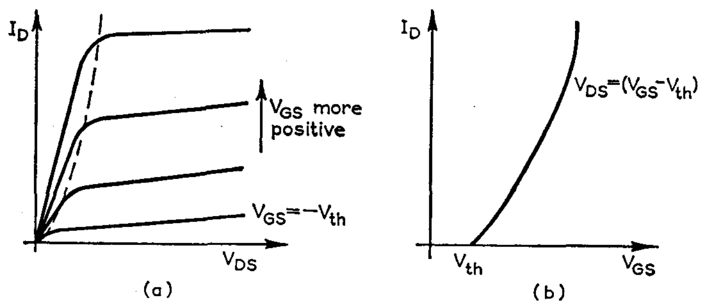


Fig. 26.6 CHARACTERISTICS OF AN *n*-CHANNEL ENHANCEMENT-TYPE IGFET

(*a*) Drain characteristics
(*b*) Transfer characteristic

respect to the source, then an electric field is set up beneath the gate electrode. Since the silicon-dioxide layer is thin and since the silicon substrate is not a very good conductor this field will extend into the *p*-type silicon and will cause an attraction of electrons to just beneath the dioxide layer. Eventually as the gate–source voltage, V_{GS}, is increased it reaches a sufficiently positive value (called the *threshold voltage*, V_{th}) for the movement of electrons to the surface to be large enough for a predominately *n*-type region to form near the surface. This then constitutes a conducting channel between n_1 and n_2 as shown in Fig. 26.5(*a*).

With V_{GS} more positive than the threshold value any voltage on the drain will cause conduction between drain and source. This drain current will be almost linearly related to both V_{DS} and V_{GS}. However, as the drain voltage is made more positive, the voltage between gate and drain falls, until eventually the pinchoff voltage is reached and the conducting channel becomes pinched off as in the JUGFET. The drain current cannot increase further and becomes almost constant until breakdown occurs, at which point the current rises very rapidly. The family of drain characteristics are shown in Fig. 26.6(*a*)—the corresponding transfer characteristic is shown at (*b*). The transfer characteristic approximates very closely to a square law, giving, in the pinchoff region,

$$I_D = -\frac{\beta}{2}(V_{GS} - V_{th})^2 \tag{26.1}$$

where β is a constant that depends on the geometry of the device. It is because the application of forward bias increases the channel conductivity that this device is said to operate in the enhancement mode.

In the depletion-mode *n*-channel IGFET, an *n*-type surface layer is diffused between the drain and source, giving an initial channel which allows conduction when $V_{GS} = 0$. If the gate is made negative, electrons are repelled from the channel, which becomes a poorer conductor. If V_{GS} is made positive the channel becomes more definitely *n*-type and the drain current increases. For any values of V_{GS} the pinchoff effect described in considering the enhancement-mode IGFET will apply, so that the drain characteristics will be as shown in Fig. 26.7.

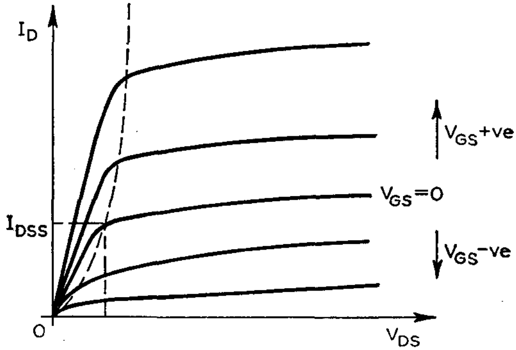


Fig. 26.7 DRAIN CHARACTERISTICS OF AN *n*-CHANNEL DEPLETION-MODE IGFET

One parameter that is often quoted by manufacturers is the saturation drain current that flows when gate and source are connected together. This is given the symbol I_{DSS}.

Exactly similar relations will be obtained using an *n*-type substrate with a *p*-type channel and heavily doped *p*-type drain and source areas, i.e. the IGFET can be manufactured in complementary forms.

Transfer characteristics for p-channel and n-channel depletion- and enhancement-mode IGFETs are shown in Fig. 26.8.

A depletion-mode IGFET is conducting when $V_{GS} = 0$. This can sometimes be a useful property, e.g. in oscillator circuits. An enhancement-mode IGFET must be biased for conduction to take place. As with bipolar transistors, this has the advantages of (i) allowing circuits to be designed for direct interstage coupling without requiring an auxiliary d.c. supply of opposite polarity, and (ii) giving automatic cut-off when $V_{GS} = 0$. Note that depletion-mode devices operate by enhancement when the gate voltage increases. Enhancement-mode devices, on the other hand, can never act in the depletion mode—there is nothing to deplete!

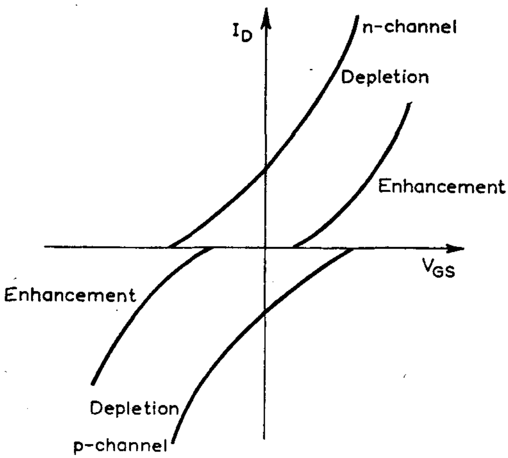


Fig. 26.8 IGFET TRANSFER CHARACTERISTICS

The substrate of an IGFET is normally connected to the source but may be biased to alter the pinchoff voltage. In the pinchoff region the equivalent circuit may be deduced by considering the load line on drain/source characteristics that are assumed to be parallel straight lines of slope $1/r_{DS}$ (r_{DS} is the *drain/source slope resistance*). A simple enhancement-mode common source amplifier circuit is shown in Fig. 26.9(*a*). Here the gate bias is obtained from resistors R_1, R_2 and R_g. With this arrangement R_g can be of a high value to give a high impedance bias voltage source as seen from the gate. The drain/source characteristics with a load line AB are shown at (*b*). This load line is drawn from B (where V_{DS} is equal to the supply voltage V_{DD}) to A (where I_D is equal to V_{DD}/R_L). For any

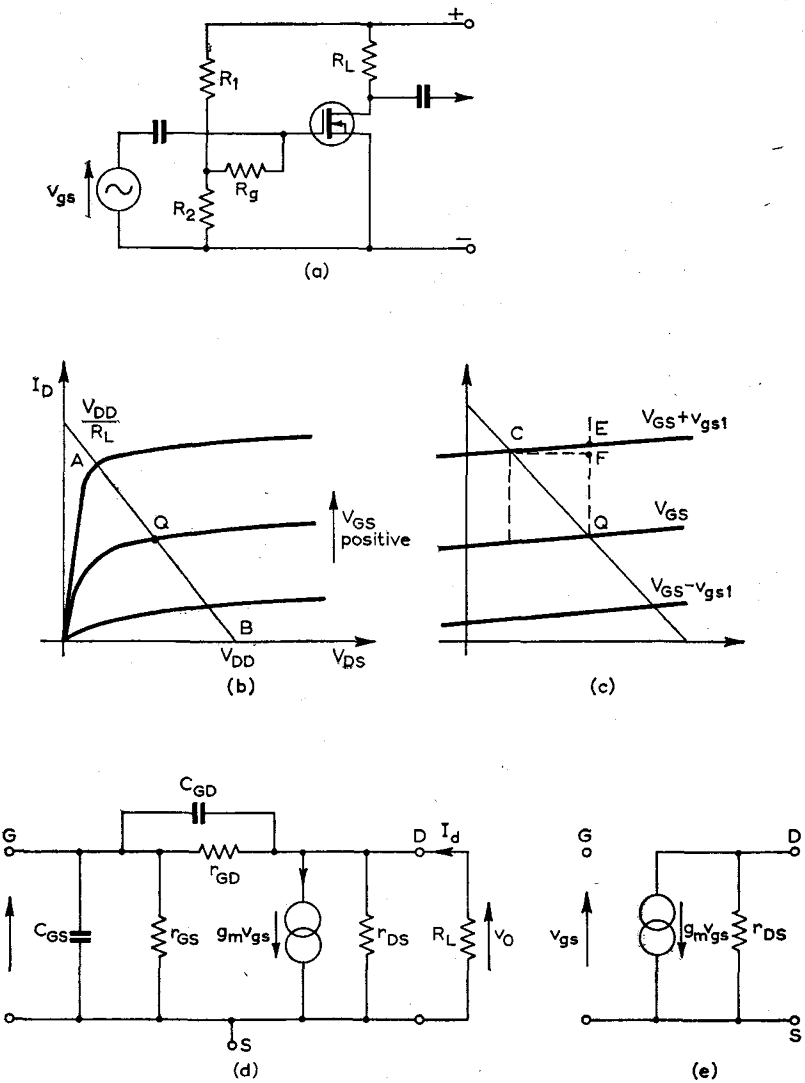


Fig. 26.9 LOAD LINE AND SMALL-SIGNAL EQUIVALENT CIRCUITS

variation in gate voltage about the quiescent value, the operating point moves up or down AB to the corresponding value of V_{GS}.

Consider the drain/source characteristics as at (*c*). With $R_L = 0$ any change in gate–source voltage would cause a proportional change in drain current I_D, given by

$$\delta I_D' = g_m \delta V_{GS} \qquad \text{where} \qquad g_m = \left.\frac{\delta I_D}{\delta V_{GS}}\right|_{V_{DS}=\text{constant}}$$

g_m is known as the *mutual conductance* of the FET. The change in current, $\delta I_D'$, is represented on the characteristics by QE in Fig. 26.9(*c*). With R_L connected the actual drain current change, δI_D, will be QF (the operating point will actually move to C). Hence

$$\delta I_D = g_m \delta V_{GS} - \text{EF} = g_m \delta V_{GS} - \frac{\text{CF}}{r_{DS}}$$

But CF is the change in output voltage δV_0, where $\delta V_0 = -\delta I_D R_L$ (an increase in I_D causes a decrease in V_0). Hence

$$\delta I_D = g_m \delta V_{GS} - \frac{\delta I_D R_L}{r_{DS}}$$

or

$$\delta I_D = g_m \delta V_{GS} \frac{r_{DS}}{r_{DS} + R_L} \tag{26.2}$$

This is the current that would flow in a load resistance R_L connected across a constant-current source of value $g_m V_{GS}$ with internal impedance r_{DS} as shown at (*d*).

The above analysis applies to both JUGFETS and IGFETS, and to both depletion and enhancement modes.

The other components of the equivalent circuit represent the leakage resistances, r_{GS} and r_{GD}, between gate and source, and gate and drain, respectively, and in JUGFETS the depletion-layer capacitances, or in IGFETS the insulation-layer capacitances, C_{GS} and C_{GD}. Typical values for these quantities are 100 MΩ and 10 pF for JUGFETS, and 10^8 MΩ and 2 pF for IGFETS. Typical values of mutual conductance, g_m, are 2 mA/V for JUGFETS and 1·5 mA/V for IGFETS, though in some devices much higher values are obtained. These values are considerably lower than the 40 mA/V common in bipolar transistors. Note that the gate input resistances are generally so high that they may often be neglected, giving the low-frequency equivalent circuit shown in Fig. 26.9(*e*). At high frequencies, however, the gate input capacitance cannot be neglected.

The phase relationships in the common-source circuit which has been described (and which is equivalent to the bipolar common-emitter circuit) are obtained from the fact that any increase in gate–source voltage causes an increase in drain current. This in turn increases the voltage drop across the load resistance and causes a decrease in output voltage—i.e. there is an inherent 180° phase shift. Hence in complexor r.m.s. signal values

$$I_d = g_m V_{gs} \frac{r_{DS}}{r_{DS} + R_L} \tag{26.3}$$

and

$$V_o = -I_d R_L = -\frac{g_m r_{DS} R_L}{r_{DS} + R_L} V_{gs} \tag{26.4}$$

The voltage gain is thus

$$A_v = \frac{V_0}{V_{gs}} = -\frac{g_m r_{DS} R_L}{r_{DS} + R_L} \tag{26.5}$$

The product $g_m r_{DS}$ is sometimes called the FET *amplification factor* μ. Note that subscripts in small letters are used to denote r.m.s. quantities.

EXAMPLE 26.1 An *n*-channel IGFET is connected as a simple amplifier as in Fig. 26.9(*a*). If $g_m = 1{\cdot}5$ mA/V, $r_{DS} = 45$ kΩ and $R_L = 47$ kΩ, determine the overall mid-frequency voltage gain. If $R_1 = 100$ kΩ, $R_2 = 10$ kΩ and $R_g = 10$ MΩ find the current and power gains.

From eqn. (26.4),

$$\frac{V_o}{V_{gs}} = A_v = -\frac{g_m r_{DS} R_L}{r_{DS} + R_L}$$

$$= -\frac{1{\cdot}5 \times 10^{-3} \times 45 \times 10^3 \times 47 \times 10^3}{92 \times 10^3}$$

$$= \underline{\underline{-34{\cdot}5}}$$

The input resistance of the IGFET can be considered to be very much greater than the 10 MΩ of R_g. Hence the input current I_i is essentially that which flows through R_g and the parallel combination of R_1 and R_2—i.e. it is almost equal to V_{gs}/R_g, since $R_1R_2/(R_1 + R_2) \ll R_g$. Since the output current $I_o = V_o/R_L$ the current gain is

$$\frac{I_o}{I_i} = A_i = \frac{V_o}{V_{gs}}\frac{R_g}{R_L} = -34{\cdot}5 \times \frac{10^7}{47 \times 10^3} = \underline{\underline{-7{\cdot}34 \times 10^3}}$$

and the power gain is $A_v A_i = \underline{\underline{253 \times 10^3}}$

These results should be compared with those for bipolar transistors.

26.4 Temperature Effects

In JUGFETS the reverse biased gate diode exhibits the increase in leakage current with temperature that is common to all junction diodes; hence the input resistance falls as the temperature rises. This effect is much less pronounced in IGFETS because the input resistance is determined by the temperature coefficient of the insulating silicon-dioxide layer. In both types the channel resistance increases with temperature (due to a decrease in carrier mobility) giving a negative temperature coefficient of drain current. In JUGFETS the gate diode contact voltage drops with temperature (by about 2 mV/K at around 300 K). This causes a positive temperature coefficient of drain current which will be more pronounced the lower the value of pinchoff voltage, V_p. Thus JUGFETS with low values of V_p ($\sim 0{\cdot}5$ V) have normally an overall positive temperature coefficient, while those with $V_p > 1{\cdot}5$ V have normally an overall negative temperature coefficient of drain current.

Since the carriers in the conducting channel do not tend to "freeze out", IGFETS can be operated down to liquid helium temperatures (~ 4 K), and temperature effects are generally small even up to 400 K (~ 130°C). (In bipolar transistors the impurity centres tend to fix or "freeze" the excess holes or electrons at low temperatures so that they are no longer mobile. In FETS the channel field prevents this effect).

Because of the small changes that occur in drain current with increase in temperature, FETS are not subject to the thermal runaway that can occur in bipolar transistors.

26.5 High-frequency Performance

As with bipolar transistors, the high-frequency performance of FETS is basically limited by the source-to-drain transit time of the majority carriers. In addition, parasitic channel impedances and gate capacitance can limit high-frequency performance. In general the h.f. performance can be specified in the same manner as was developed in Chapter 21 for bipolar transistors. Operating frequencies extend well into the 100 MHz region, and are generally limited by the external circuit strays rather than the device itself. This is where the higher mutual conductance of the bipolar transistor can be of advantage.

One advantage of the FET over the bipolar transistor arises because it is a majority-carrier device. For this reason, on step inputs the turn-off time will be small due to the absence of minority-carrier storage effects. Turn-on and turn-off times in the nano-second region are possible.

26.6 Bias Arrangements

Some typical small-signal common-source circuits are shown in Fig. 26.10. At (*a*) the bias resistors, R_1 and R_2, can be of high value to give the required bias for operation in the enhancement mode

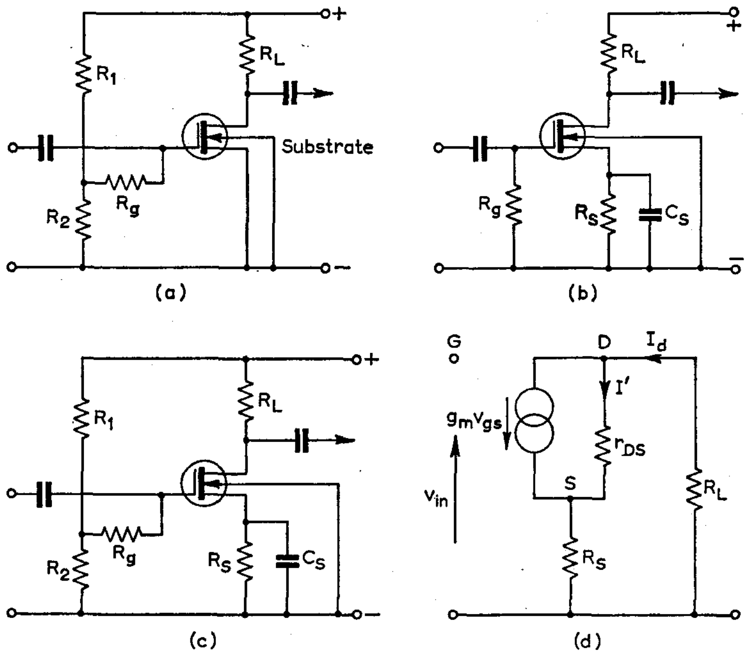


Fig. 26.10 TYPICAL BIASING ARRANGEMENTS FOR AN *n*-CHANNEL IGFET

(*a*) Enhancement type
(*b*) Depletion type
(*c*) Either type
(*d*) Equivalent circuit

but taking very little power from the supply. A possible bias arrangement for the depletion-mode IGFET is shown at (*b*). The circuits at (*a*) and (*b*) may be combined to give that at (*c*), which can be applied to both modes. The source resistance R_s gives negative feedback on direct current, and this increases the temperature stability. The voltage gain for these circuits is given by eqn. (26.5):

$$A_v = -\frac{g_m r_{DS} R_L}{r_{DS} + R_L}$$

If the bypass capacitor C_S is omitted from (*b*) or (*c*) then signal negative current feedback takes place. The equivalent circuit for both connexions is then as shown at (*d*). For this

$$V_{gs} = V_{in} - I_d R_S \tag{i}$$

$$I_d = I' + g_m V_{gs}$$

and

$$I_d(R_L + R_S) + I' r_{DS} = 0 = I_d(R_L + R_S + r_{DS}) - g_m r_{DS} V_{gs} \tag{ii}$$

Hence, substituting in (ii) from (i) to eliminate V_{gs},

$$I_d(R_L + R_S + r_{DS}) = g_m r_{DS}(V_{in} - I_d R_s)$$

so that

$$I_d = \frac{g_m r_{DS} V_{in}}{R_L + (1 + g_m r_{DS})R_S + r_{DS}}$$

and the voltage gain A_{vf} with feedback is

$$A_{vf} = \frac{V_o}{V_{in}} = -\frac{I_d R_L}{V_{in}} = -\frac{g_m r_{DS} R_L}{r_{DS} + R_L + (1 + g_m r_{DS})R_S} \tag{26.6}$$

This expression should be compared with the corresponding result in Chapter 22, i.e.

$$A_{vf} \approx \frac{A_v}{1 - \beta A_v} \tag{26.7}$$

which can be derived from the expression for A_v and eqn. (26.6) if $\beta = R_S/R_L$.

Note that for all of these circuits the input resistance of the FET itself is very high (hundreds of megohms). The circuit input impedance depends on the bias circuit arrangements. The output impedance is r_{DS} without feedback. This impedance is increased by negative current feedback and reduced by negative voltage feedback.

26.7 Common-drain Connexion (Source Follower)

This connexion is shown in Fig. 26.11. The circuit is one with 100 per cent negative voltage feedback. From the equivalent circuit at (*b*),

$$V_{gs} = V_{in} - V_o \tag{i}$$

$$I_d = g_m V_{gs} + I' \tag{ii}$$

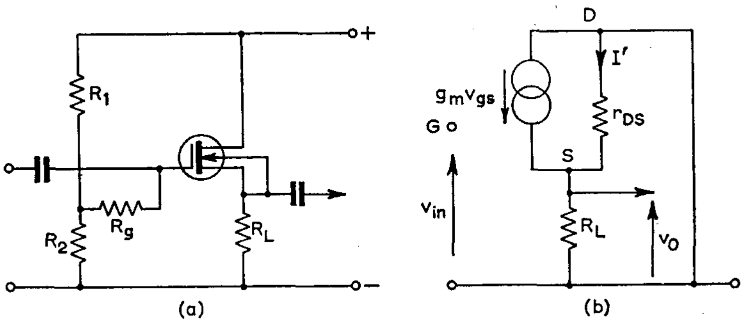


Fig. 26.11 SOURCE FOLLOWER AND EQUIVALENT CIRCUIT

and

$$I'r_{DS} = -V_o = -(I_d - g_m V_{gs})r_{DS} \tag{iii}$$

Hence

$$-\frac{V_o}{r_{DS}} = I_d - g_m(V_{in} - V_o)$$

so that

$V_o = I_d R_L$ (Note that there is *no* phase reversal with this connexion)

$$= -\frac{V_o}{r_{DS}} R_L + g_m V_{in} R_L - g_m V_o R_L \tag{26.8}$$

and the voltage gain is

$$A_{vf} = \frac{V_o}{V_{in}} = \frac{g_m R_L}{1 + \dfrac{R_L}{r_{DS}} + g_m R_L} \approx 1 \tag{26.9}$$

provided that $g_m R_L \gg 1 + R_L/r_{DS}$.

Rewriting eqn. (26.8),

$$A_{vf} = \frac{g_m R_L}{\left(\dfrac{1}{r_{DS}} + g_m\right)\left(\dfrac{r_{DS}}{1 + g_m r_{DS}} + R_L\right)}$$

so that (comparing with eqn. (22.16)) the output impedance is

$$Z_{out} = \frac{r_{DS}}{1 + g_m r_{DS}} \tag{26.10}$$

$$\approx \frac{1}{g_m} \qquad \text{if } g_m r_{DS} \gg 1 \tag{26.11}$$

i.e. the output impedance is very low.

The input impedance of the circuit will depend on the bias arrangement rather than the FET.

The source follower may be used as a buffer stage in the same way as the emitter follower.

26.8 Common-gate Circuit

The common-gate circuit corresponds to the common-base circuit of the bipolar transistor. It may be used as an impedance convertor from low input to high output values.

26.9 Cascaded Stages

FETS may be cascaded in the same way as bipolar transistors. Couplings may be *CR* or direct. A typical circuit is shown in Fig. 26.12.

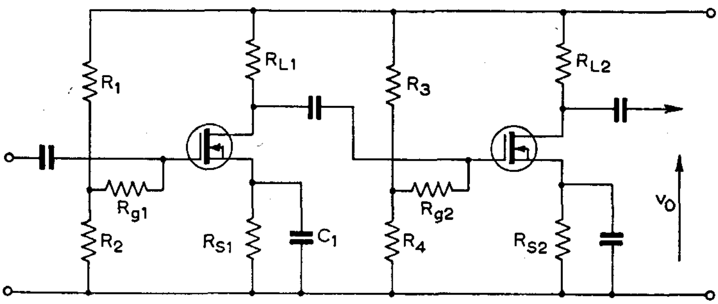


Fig. 26.12 CASCADED FET AMPLIFIER

Frequently FET circuits are cascaded with bipolar power transistors using the high-impedance FET to feed the low-impedance bipolar transistor. High power gains can be achieved, but the low input impedance of the bipolar transistor acting as the load on the FET reduces the voltage gain to a low value.

FETS are particularly suited to incorporation in integrated circuits. They may also be used in digital circuits and logic circuits. In integrated circuits the ability of the FET to act as a resistor is of great value. Normally in integrated circuits *p*-channel operation is used. In general FETS have been used only in small-signal or low-power applications due to difficulties with power dissipation. However, they have been employed from audio to ultra-high frequencies (>400 MHz).

PROBLEMS

The *n*-channel depletion-mode IGFET used in these problems has the following characteristics. I_D is in milliamperes, and all voltages are in volts.

$V_{GS} = 3$	V_{DS}	10	20	30		
	I_D	1·2	1·7	1·85		
$V_{GS} = 2$	V_{DS}	10	20	30		
	I_D	0·8	1·2	1·28		
$V_{GS} = 1$	V_{DS}	5	10	20	30	40
	I_D	0·38	0·65	0·72	0·77	0·82
$V_{GS} = 0$	V_{DS}	5	10	20		40
	I_D	0·20	0·31	0·35		0·38
$V_{GS} = -1$	V_{DS}	5	10			50
	I_D	0·06	0·08			0·09

26.1 Draw the FET characteristics. From these determine for a quiescent operating point at $V_{DS} = 25$ V and $V_{GS} = 1$ V approximate values for r_{DS}, g_m and μ.

Ans. 400 kΩ; 0·45 mA/V; 180.

26.2 A common-source amplifier uses the above FET. The supply voltage is 50 V and the load resistance is 25 kΩ. Draw the load line. If the quiescent gate voltage is 1 V determine from the characteristics the approximate small-signal voltage gain. Sketch the output voltage waveform for a sinusoidal signal input voltage of peak value 1 V.

Ans. 10.

26.3 The FET is used in the amplifier shown in Fig. 26.10(*b*). If the supply voltage is 40 V, $R_L = 24$ kΩ and $R_S = 1$ kΩ, determine the quiescent operating point.

Ans. $V_{DS} = 33$ V; $I_D = 0{\cdot}66$ mA.

Chapter 27

LOGIC

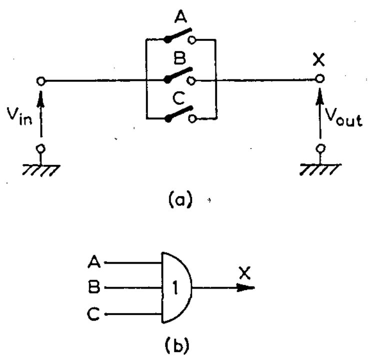
Logic circuits are used in digital computers, data processors, and many forms of control and sequencing systems. The individual elements (or *gates*) of a logic system are normally obtained direct from the manufacturers as "black boxes", and it is the interconnexion of these elements that gives rise to a particular logic system. Hence it is very important to understand the interconnexions, and such overall systems will be dealt with before the individual electronic elements are considered. This "systems approach" applies whether the logic elements are pneumatic (fluidic) or electrical, electronic, etc. The chapter will close with a brief outline of some electronic means of achieving logic functions. It is emphasized, however, that the chapter as a whole attempts only to introduce the subject.

Essentially logic circuits operate on a digital or two-state basis. In positive electronic logic circuits the presence of a positive voltage at a point is conventionally designated the logic "1" state at that point (or ON state). The absence of a voltage is described as the logic "0" state (or OFF state). Note that the symbols "1" and "0" do not have their normal mathematical meaning, but indicate simply the presence or absence of voltage. Various voltage levels may be chosen for the "1" state, e.g. +1 V, +5 V, + 20 V, etc. In negative logic the "1" state is represented by the more negative voltage of the two voltage levels chosen. Generally the output of any logic element or gate will be either almost the circuit supply voltage or almost zero. This output will depend on the states ("0" or "1") of the various inputs to the gate. (Note that in some instances the "1" state may be represented by zero volts—e.g. in a positive logic system with reference voltages of -3 V and 0 V, or a negative logic system with reference voltages $+5$ V and 0 V—this however is not common.)

Logic circuits are so called because the output quantity depends on the logic inputs on a "yes-no" (or "true-false") basis, just as the results of classic logic depend on the initial statements.

27.1 The Logic OR Gate

This gate gives a logic "1" output if any one of its various inputs is in the "1" state. The number of inputs possible for the gate depends on the design and output loading of the element and is referred to as the *fan in*. The OR gate may be considered symbolically as resulting from the parallel connexion of switch contacts as shown for a 3-input gate in Fig. 27.1(*a*), where the switches are taken to be in the



Inputs			Output
A	B	C	X
0	0	0	0
1	0	0	1
0	1	0	1
1	1	0	1
0	0	1	1
1	0	1	1
0	1	1	1
1	1	1	1

Fig. 27.1 THREE-INPUT "OR" GATE

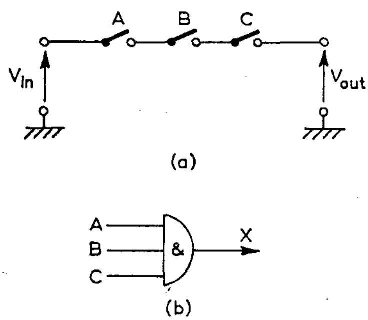
"1" state when closed. The logic symbol is shown at (*b*); the figure 1 inside the semicircle represents the minimum number of inputs that must be in the logic "1" state in order to give a logic "1" output—in this case one. The *truth table* on the right of the diagram relates the state of the output, X, to that of the three inputs, A, B and C. Symbolically, using the notation of the algebra developed by George Boole, and called *Boolean algebra*,

$$X = A + B + C \tag{27.1}$$

where the + sign is taken to represent the logic function OR. This equation means that X takes the logic value "1" if, and only if, *at least one of* the inputs A OR B OR C has the logic value "1". The truth table shows how X varies with A, B and C.

27.2 The Logic AND Gate

The AND gate gives a logic "1" output only if *all* its inputs are in the "1" state. It may be considered to be the result of connecting a group of switch contacts in series as shown in Fig. 27.2(*a*). There



Inputs			Output
A	B	C	X
0	0	0	0
1	0	0	0
0	1	0	0
1	1	0	0
0	0	1	0
1	0	1	0
0	1	1	0
1	1	1	1

Fig. 27.2 THREE-INPUT "AND" GATE

will be an output only if contacts A AND B AND C are closed. The logic symbol is shown at (*b*), and the truth table on the right represents the Boolean expression

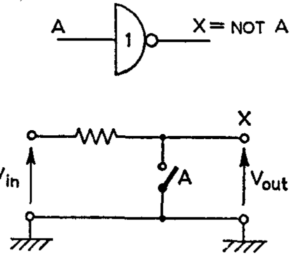
$$X = A \, . \, B \, . \, C \tag{27.2}$$

where the points represent the logic function AND. The equation means that the output X takes the value "1" only if A AND B AND C are *all* in state "1".

Notice that in both the OR and the AND truth tables the conditions of the inputs have been arranged for convenience in a logical form that derives from binary arithmetic. Thus the first input (A) alternates between "0" and "1", the second input (B) is in pairs of zeros and of ones, and the third input (C) is in groups of four zeros and four ones. If this scheme is followed, no possible combinations of inputs will be missed from the truth table. Note that the total number of different combinations of input for an n-input gate will be 2^n.

27.3 The NOT Gate—NOR and NAND

The NOT gate is essentially an invertor with one input. The output, X, is "1" if the input is "0", and is "0" if the input is "1". The logic and circuit symbols are shown in Fig. 27.3. Notice that the



A	X
0	1
1	0

Fig. 27.3 THE "NOT" GATE

small circle in the logic symbol represents the negation or inversion of the output (in the same way a small circle at the input represents negation of that input).

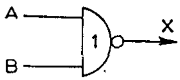
In Boolean notation the output X is NOT A, and the negation is represented by a bar above the symbol, i.e.

$$X = \bar{A} \tag{27.3}$$

The truth table is shown on the right. Obviously two NOT gates in cascade (i.e. with the output of the first NOT gate providing the input to the second) cancel each other, i.e.

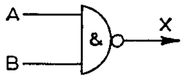
$$\bar{\bar{A}} = A$$

Usually the operations of OR and NOT and of AND and NOT are combined, since this normally results in some simplification of the electronic circuitry involved due to the inherent inversion of a single stage. These combined operations are termed NOR and NAND respectively. Truth tables for two-input NOR and NAND gates are shown in Figs. 27.4(*a*) and (*b*) respectively, the logic symbols being shown on the left.



(a) NOR gate

A	B	$X = \overline{A + B}$
0	0	1
1	0	0
0	1	0
1	1	0



(b) NAND gate

A	B	$X = \overline{A . B}$
0	0	1
1	0	1
0	1	1
1	1	0

Fig. 27.4 TWO-INPUT "NOR" AND "NAND" GATES

The Boolean expressions for the NOR and NAND functions for multiple-input gates are

NOR $$X = \overline{A + B + C +} \ldots \qquad (27.4)$$

NAND $$X = \overline{A . B . C} \ldots . \qquad (27.5)$$

Note that multiple-input NOR or NAND gates in which only one input is used simply give the NOT function. (See Section 27.5 and later).

It is left as an exercise for the reader to draw the truth tables for 3-input NOR and NAND gates.

27.4 Simple Boolean Relations

It is obvious that, for more than three of four inputs, truth tables become cumbersome. In such cases it is usually more convenient to simplify the logic system by using some of the following relationships that apply to Boolean algebra. The logic system required in any given instance is, of course, a function of the action required.

Boolean algebra obeys the three basic laws of ordinary algebra, namely commutation ($A + B = B + A$), association ($((A + B) + C) = (A + (B + C))$ and distribution ($A . (B + C)) = (A . B + A . C)$. It differs, however, from ordinary algebra in two significant respects, namely:

(*a*) "add", "negate" and "multiply" are the only allowed operations,
(*b*) the "answer" must be either "1" or "0".

The following equations illustrate this:

$$0 + 0 = 0 \quad (27.6) \qquad 0.0 = 0 \quad (27.9)$$

$$0 + 1 = 1 \quad (27.7) \qquad 0.1 = 0 \quad (27.10)$$

$$1 + 1 = 1 \quad (27.8) \qquad 1.1 = 1 \quad (27.11)$$

In particular, eqn. (27.8) expresses the fact that two closed switches in parallel are simply equivalent to a single closed switch. If A is taken as a logic gate input which can be either in the "1" or the "0" state, then the following Boolean expressions apply:

$$A + A + A + \ldots = A \tag{27.12}$$

$$A + \bar{A} = 1 \tag{27.13}$$

$$1 + A = 1 \tag{27.14}$$

$$0 + A = A \tag{27.15}$$

$$\overline{A + \bar{A}} = 0 = A . \bar{A} \tag{27.16}$$

$$A . A . A . \ldots = A \tag{27.17}$$

$$A . \bar{A} = 0 \tag{27.18}$$

$$1 . A = A \tag{27.19}$$

$$0 . A = 0 \tag{27.20}$$

$$\overline{A . \bar{A}} = 1 = A + \bar{A} \tag{27.21}$$

For two inputs A and B (both of which can take the logic value "1" or "0") the following can readily be verified by drawing up the appropriate truth tables, or applying the relationships stated above:

$$A + A . B = A(1 + B) = A . 1 = A \tag{27.22}$$

$$A . (A + B) = A . A + A . B = A + A . B = A \tag{27.23}$$

$$A + \bar{A} . B = A + B \tag{27.24}$$

$$A . (\bar{A} + B) = A . \bar{A} + A . B = A . B \tag{27.25}$$

The following important relations, which are known as *de Morgan's theorem*, may also be verified by a truth table:

NOR $$\overline{A + B + C} = \bar{A} \, . \, \bar{B} \, . \, \bar{C} \tag{27.26}$$

NAND $$\overline{A \, . \, B \, . \, C} = \bar{A} + \bar{B} + \bar{C} \tag{27.27}$$

Extending these equations,

$$A + B + C = \overline{\bar{A} \, . \, \bar{B} \, . \, \bar{C}} \tag{27.28}$$

and

$$A \, . \, B \, . \, C = \overline{\bar{A} + \bar{B} + \bar{C}} \tag{27.29}$$

Equation (27.28) shows that a NAND gate may be used to perform the OR function provided that each input is inverted before being applied to the gate. In the same way, by inverting the inputs, a NOR gate will perform the AND function. Such inverted inputs are normally available from the input transducers. Thus, if the input is derived from a changeover microswitch as shown in Fig. 27.5, then both A and $\bar{A}$ inputs are available.

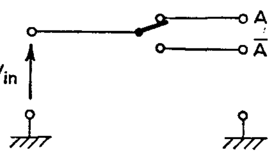


Fig. 27.5 A CHANGEOVER MICROSWITCH PROVIDING BOTH A AND Ā OUTPUTS

EXAMPLE 27.1 A coin-operated hot-drink dispenser will provide a paper cup of tea or coffee under the following conditions:

(*a*) the correct coin is inserted (I)
AND (*b*) a paper cup is in position (P)
AND (*c*) hot water is available (W)
AND (*d*) the selector is set at "tea" (T) OR "coffee" (C)
$\left.\begin{array}{l}\text{AND milk } (M) \text{ OR no milk } (\bar{M}) \\ \text{AND sugar } (S) \text{ OR no sugar } (\bar{S})\end{array}\right\}$ are set

Obtain the logic expression for the conditions under which a drink D, may be obtained. Show how this may be implemented using (*a*) NOR gates only, (*b*) NAND gates only.

The Boolean expression for D is derived from the conditions under which action is required, i.e.

$$\begin{aligned} D &= I \, . \, P \, . \, W \, . \, ((T + C) \, . \, (M + \bar{M}) \, . \, (S + \bar{S})) \\ &= I \, . \, P \, . \, W \, . \, (T + C) \end{aligned}$$

since $(M + \bar{M}) = (S + \bar{S}) = 1$ where a drink is obtained whenever $D = 1$. Note that a drink will be obtained whichever setting of the "milk" and "sugar" selectors is used.

(*a*) From eqn. (27.29),

$$D = \overline{\bar{I} + \bar{P} + \bar{W} + \overline{(T + C)}}$$

i.e. D is the output of a NOR gate whose inputs are $\bar{I}$, $\bar{P}$, $\bar{W}$ and $\overline{(T + C)}$. Note that $\overline{(T + C)}$ is itself the output of a NOR gate whose inputs are T and C. The resultant logic circuit is shown in Fig. 27.6(*a*).

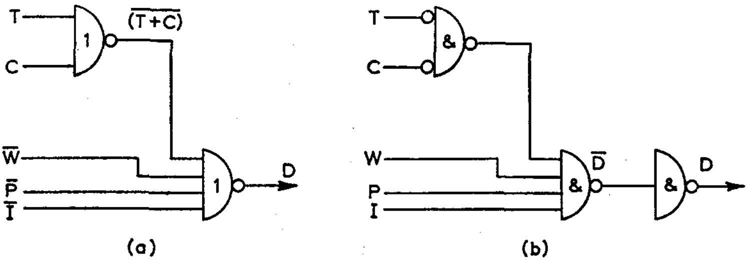


Fig. 27.6

(*b*) In order to be able to realize the expression for D using NAND gates only, the equation $D = I \,.\, P \,.\, W \,.\, (T + C)$ must be manipulated so that only NAND expressions appear on the right-hand side. The part in brackets is readily expressed in terms of a NAND gate using eqn. (27.28). Thus

$$T + C = \overline{\bar{T} \,.\, \bar{C}}$$

i.e. $T + C$ can be obtained using a NAND gate with inputs $\bar{T}$ and $\bar{C}$. It is not, however, possible to express D directly in terms of NAND gates. We must first obtain $\bar{D}$ as

$$\bar{D} = \overline{I \,.\, P \,.\, W \,.\, (\overline{\bar{T} \,.\, \bar{C}})}$$

i.e. $\bar{D}$ is obtained as the output of a NAND gate whose inputs are I, P, W and $\overline{\bar{T} \,.\, \bar{C}}$. Then D itself is readily obtained by adding one further NAND gate as shown in Fig. 27.6(*b*) and recalling that $D = \bar{\bar{D}}$. Note that the inverted inputs $\bar{T}$ and $\bar{C}$ are represented at (*b*) by the small circles at the input of the first NAND gate.

27.5 Electronic Logic Families

Electronic logic circuits are commonly constructed in integrated circuit form, where a complete gate (or indeed several complete gate circuits and their interconnexions) are formed on a single monocrystalline silicon chip. They may also be constructed from discrete components. Such circuits may be obtained in AND, OR, NAND, NOR etc., configurations over a range of positive or negative logic voltages. The various types of circuit may be considered as families that are distinguished from one another on the basis of the following properties.

(*a*) *Cost per gate.* Very frequently this is the most important criterion.

(*b*) *Propagation delay time per gate.* It always takes a finite time for an electronic circuit to change over from one steady state to another.

(*c*) *Threshold voltage.* This is the input voltage level that is required to make the circuit change from one logic state to another. It is important since logic inputs are often the outputs from other logic elements.

(*d*) *Noise margin.* It is undesirable that an unwanted signal (or "noise") should cause malfunction of any logic gate. The *noise margin* of a logic circuit is the difference between the operating voltage and the threshold voltage. If any unwanted disturbance exceeds this noise margin the gate may change state without any change in the true logic input.

(*e*) *Maximum fan-in.* This is the total maximum number of logic inputs with which any particular logic circuit is designed to operate. If the design number is exceeded, faulty logic operation may result.

(*f*) *Maximum fan-out.* This is the total maximum number of logic circuits that any one gate is capable of driving. If it is exceeded, operating voltages may fall below the threshold value.

(*g*) *Power dissipation per gate.* Where there are a great many logic elements, it becomes important to keep the overall power dissipation to as low a value as possible.

Four families of logic circuit will be considered in the next four sections. There are:

RTL. Resistor-transistor logic
DTL. Diode-transistor logic
TTL. Transistor-transistor logic
ECL. Emitter-coupled logic

Generally the cost, power dissipation per gate and fan-out all increase as we go from RTL to ECL, while the propagation delay time decreases. The noise margin is typically highest for the TTL family of circuits, while the fan-in is highest for DTL and TTL. It is left as an exercise for the reader to account for these facts from the following circuit descriptions.

27.6 Resistor-Transistor Logic (RTL)

A typical circuit using RTL to produce the NOR logic function with positive logic is shown in Fig. 27.7. A NAND form of circuit is also possible.

In the circuit shown, the voltage $-V_{BB}$, the input resistors, and the resistor R_4 are so chosen that with input voltages of near zero at A, B and C the base–emitter junction of transistor Tr_1 is reverse biased. The transistor is therefore cut off, and the output at X will be very nearly V_{CC} or logic "1", depending on the external loading of the gate; obviously this loading must be restricted or the voltage at X will be considerably below V_{CC}). If any one of the inputs has a sufficient positive voltage applied to it then the transistor will be turned "on" and will saturate, with the result that the output voltage

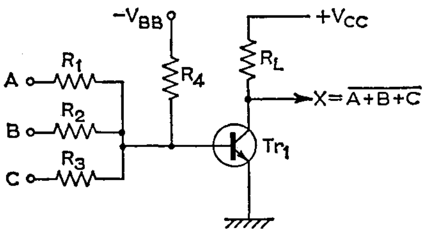


Fig. 27.7 THREE-INPUT RESISTOR-TRANSISTOR-LOGIC "NOR" CIRCUIT

falls to almost zero. Thus, if A or B or C is "up" in voltage (i.e. logic "1") the output at X is "down" (i.e. logic "0") and the NOR function has been produced.

Note that, if only one input is used and the others are left unconnected, then the circuit acts as an invertor or NOT circuit, i.e. the output is "up" if the input is "down" and vice versa. A NOR followed by a NOT gives an OR circuit.

In some RTL circuits each input is applied through a resistor direct to the base of a separate transistor (hence there are as many transistors as inputs). The transistors all have one common-collector load resistor from which the output is taken. This form of circuit has the advantage that it does not require a separate negative supply rail. The added complexity of additional transistors does not increase the relative cost of integrated circuits in the proportion that it would do in discrete circuits.

The relatively high propagation delay time of RTL circuits arises because of (*a*) the combination of input resistors and the transistor base–emitter capacitance, giving rise to a *CR* delay, and (*b*) the fact that the transistor saturates (giving high turn-off times).

RTL circuits, which were the first commercially available, are not now common in new equipments.

27.7 Diode-Transistor Logic (DTL)

A typical 3-input DTL NAND gate is shown in Fig. 27.8. Again positive logic is used. Only if all the inputs, *A*, *B* and *C* are positive together will none of the diodes D_1, D_2, or D_3 conduct. In this case point K will be at a positive potential, D_4 and the base–emitter junction of transistor Tr_1 will conduct, Tr_1 will saturate and the output at X will be almost zero (typically 0·2 V). The voltage between K and

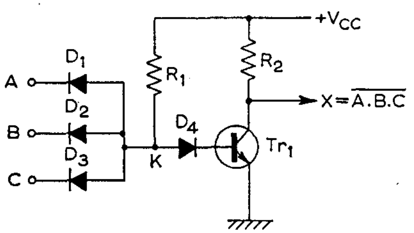


Fig. 27.8 THREE-INPUT DIODE-TRANSISTOR-LOGIC "NAND" GATE

earth will be 1·4 V (twice the mean forward conducting voltage of a silicon junction diode, since two junctions are in series between K and earth). This may be increased in 0·7 V steps by adding diodes in series with D_4 in order to increase the noise margin.

If any input voltage goes "down", the corresponding input diode conducts and the voltage at K becomes approximately 0·7 V. D_4 and the transistor are then both cut off and the output voltage rises to almost $+V_{CC}$ depending on the output loading. This represents the logic NAND operation. If only one input is connected the circuit acts as a simple invertor.

The propagation delay time is less than for RTL on account of the low forward resistance of the input diodes.

In integrated-circuit form, DTL may be cheaper than RTL since diodes are more readily produced than resistors.

21.8 Transistor–Transistor Logic (TTL)

The circuit of a multi-emitter TTL NAND gate is shown in Fig. 27.9. This form of emitter has been developed with integrated circuits and provides considerable flexibility in circuit design. The second transistor, Tr_2, provides a power output stage and permits a high fan-out.

With each input "up" (i.e. at a positive voltage, or logic "1"), the emitter–base junctions of Tr_1 are reverse biased, but the collector junction is forward biased, so permitting the flow of base current to transistor Tr_2, which saturates. The output voltage is therefore nearly zero (typically 0·2 V).

If any of the inputs *A*, *B* or *C* goes "down" (logic "0") to almost zero, then the corresponding emitter–base junction of Tr_1 is forward biased. The base voltage of Tr_1 becomes almost zero, and hence so does the base voltage of Tr_2, which cuts off, so that the output rises (logic "1").

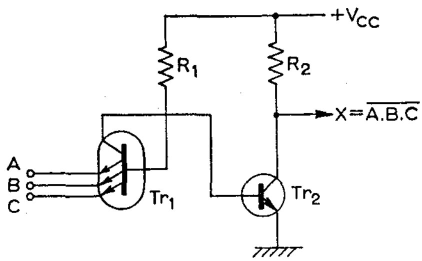


Fig. 27.9 THREE-INPUT TRANSISTOR-TRANSISTOR-LOGIC "NAND" GATE

Thus only if *A* AND *B* AND *C* are all in the logic "1" state together will the output be at logic "0"—this is the NAND gate condition, i.e.

$$X = \overline{A \, . \, B \, . \, C}$$

Alternative connexions allow for TTL NOR gates, while the inclusion of invertors allows AND and OR operations to be performed.

27.9 Emitter-Coupled Logic (ECL)

Because it operates in the non-saturated condition the ECL family of logic circuits provides the lowest propagation delay times achieved so far commercially, but at the expense of higher cost and increased power dissipation per gate. When designing a logic system, the usual engineering compromises are thus seen to be necessary, balancing speed of operation and overall performance, etc., against cost per gate.

A basic ECL NOR gate is shown in Fig. 27.10. The emitter resistor, R_3, is assumed to be large enough to give almost constant-current operation for transistors Tr_1–Tr_4. With inputs *A*, *B* and *C* all at about zero volts (logic "0"), Tr_1, Tr_2 and Tr_3 are all cut off so that their common collector leads rise towards V_{CC} volts. This turns transistor Tr_5 on and gives an output voltage.

Transistor Tr_6 and diodes D_1 and D_2 provide a constant reference voltage, V_{ref}, at the base of Tr_4. If, now, any one of the inputs rises above V_{ref}, that input transistor is turned on (current is still limited to below the saturation value by R_3) and the collector voltage falls so that Tr_5 is cut off and the output falls to nearly zero (logic "0"). This describes the logic NOR function. The reference voltage defines the circuit threshold voltage, but since non-saturated operation is used the on and off voltage levels are less precisely defined. The emitter-follower output gives the possibility of a very large maximum fan-out.

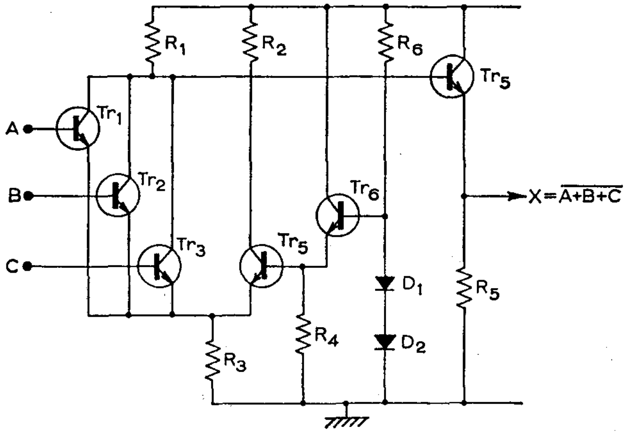


Fig. 27.10 THREE-INPUT EMITTER-COUPLED-LOGIC "NOR" GATE

ECL circuits are sometimes supplied with OR outputs but are not suitable for generating NAND or AND functions.

PROBLEMS

A knowledge of binary arithmetic is NOT required in solving these problems.

27.1 The logic equations of a binary half-adder, which can be used in computer circuits as part of the process of binary addition for two inputs A and B, and a "carry" function C, are as follows.

$$\text{Sum} = S = \bar{A} . B + A . \bar{B} \qquad \text{Carry} = C = A . B$$

Show that S can be represented by $(A + B) . \overline{AB}$. Hence or otherwise construct the logic circuits required to realize S and C using AND, OR and NOT logic elements.

27.2 Repeat Problem 27.1 using NOR logic elements only.

27.3 The truth table for an "exclusive OR" or "modulo 2 adder" is given below. The inputs are A and B, and a logic "1" output is obtained only when A and B

have different logic states. Using any logic gates you wish, construct a suitable logic circuit to realize this function.

A	B	$X = A \, . \, \bar{B} + \bar{A} \, . \, B$
0	0	0
1	0	1
0	1	1
1	1	0

27.4 Repeat Problem 27.3 using NOR elements only.

27.5 In a 4-bit coincidence detector a logic "1" output is required if and only if the inputs A, B, C and D are equal to four reference inputs A_r, B_r, C_r and D_r respectively. Obtain a suitable logic circuit using (*a*) NAND gates, (*b*) NOR gates.

Chapter 28

THE RELIABILITY OF ELECTRICAL AND ELECTRONIC EQUIPMENT

The concept of reliability has always been associated, in a qualitative way, with good design, endurance, consistent quality and dependability. In recent years, however, the much greater complexity of electrical and electronic equipment and the seriousness of a failure in the system have made it necessary to attempt not only to improve the reliability of equipment but also to assess it in quantitative terms.

In order to appreciate some of the difficulties which are involved in the quantitative assessment of reliability, imagine a discussion concerning the relative merits of two types of television receiver. In the first place the specifications are compared objectively and this is followed by, say, comparisons of the picture quality and styling. The discussion may then turn to the likelihood of faults developing in the sets. This is important not only because of the annoyance caused to the viewer by a failure but also because of the cost of repair. The customer should be prepared to pay a higher initial cost for a receiver in return for an assurance that the extra cost will mean smaller maintenance costs.

If the reliability of each type of set is to be compared the number of faults occurring in their operation will have to be measured. For the comparison to be meaningful the measurements will have to be made on a reasonably large sample of each type of set, operated under the same environmental conditions, for the same length of time. The type of fault would also need to be considered since faults vary in their seriousness and in the maintenance costs they cause.

From these considerations alone it will be seen that the assessment of reliability is not a simple matter and it will be appreciated that the achievement of high reliability is an aim in which many people must be involved. The component manufacturer, the designers, the production team, the test and quality control engineers, the installation engineer and the customer must all contribute to this aim, and it is the object of this chapter to indicate some of the considerations which are involved.

28.1 Quantitative Measurement of Reliability

The important factors which must be included in any statement of reliability have already been mentioned, and while it is difficult to evolve a definition satisfactory in all circumstances, the following meets most requirements:

> Reliability is the characteristic of a component or of a system which may be expressed by the *probability* that it will perform a *required function* under *stated conditions* for a *specified period of time*.

There are a number of difficulties which arise when this definition is applied to the assessment of an equipment. For example, it is not always easy to specify precisely its required function, or to determine the environmental conditions in which the equipment must operate reliably.

The reliability characteristic is unlike the other equipment characteristics in that it is based on statistical concepts. Whereas, for example, the gain of an amplifier can be specified either as, say, 55 ± 2 dB or as, say, $\nless 37$ dB, the reliability cannot be expressed as "this equipment will function for $1{,}000 \pm 15$ hours", or even as "this equipment will function for not less than 700 hours".

Since a probability of zero means that the event cannot happen and a probability of unity means that the event certainly will happen, a practical reliability figure will be between 0 and 1, and values which are very close to unity mean that it is very improbable that the equipment will fail.

EXAMPLE 28.1 Explain what is meant by the statement: "the probability that a certain capacitor will not fail within the next 50 hours is 0·9995".

Imagine that a very large number of similar capacitors are tested under the stated conditions for 50 hours each and that the number which have failed in that time is recorded. Then, *on average*, 5 capacitors in every 10,000 tested will have failed.

Although it is very unlikely that a selected capacitor will fail in the specified time, it is important to note that in complex equipment which uses many components it is quite likely that an equipment fault will occur within the 50 hour period. This is because it is assumed that the equipment fault will be produced by a failure in any one capacitor.

28.2 Reliability and Unreliability

If the number of components tested is N_O, the number of components which fail in time t is N_F and the number which survive is N_S, then

$$\text{Reliability, } R(t) = \frac{N_S}{N_O} \tag{28.1}$$

and

$$\text{Unreliability, } Q(t) = \frac{N_F}{N_O} \tag{28.2}$$

provided that N_O is very large.

Since $N_O = N_S + N_F$, then

$$R(t) + Q(t) = 1 \tag{28.3}$$

Thus reliability is a function which varies with time from unity at the beginning of the test ($t = 0$) to zero at a time when all the components have failed. Note that the term "failure" includes a change in a parameter to a value outside the permitted tolerance, as well as complete failures, such as short- and open-circuits. A component or system failure is *any* inability of the item to carry out its specified function.

Failures may be either *partial* or *complete*, *gradual* or *sudden*, and may be caused by an *inherent weakness* or by *misuse*.

Catastrophic failures are both sudden and complete, whereas *degradation* failures are both gradual and partial.

Primary failures are failures in components which are not caused by a failure or failures in another part of the system. *Secondary failures* are those which are caused by the failure of another part of the system.

28.3 Mean Time to Failure (MTTF)

This is a term which is applied to non-repairable parts and is a measure of the average time to failure of a large number of similar parts which operate under specified conditions. In general, the MTTF of a part will be altered by a change in the stress conditions.

For example, an increase in the operating temperature of a capacitor will reduce its MTTF.

MTTF may be calculated from the equation

$$\text{MTTF} = \frac{\text{Sum of time to failure of each component}}{\text{Number of components under test}}$$

In practice, however, the MTTF is often calculated from data taken over a period of time in which not all the components fail. In this case,

$$\text{MTTF} = \frac{\text{Total operating time for all components}}{\text{Number of failures in that time}} \tag{28.4}$$

EXAMPLE 28.2 Five hundred parts are operated under specified stress conditions for a period of 312 h, and the following fault data were recorded.

Time from start of test, t (hours)	No. of failures during time interval, n_f	Cumulative failures, n_c	No. of survivors
0			500
	19		
24		19	481
	15		
48		34	466
	13		
72		47	453
	17		
96		64	436
	12		
120		76	424
	16		
144		92	408
	12		
168		104	396
	14		
192		118	382
	11		
216		129	371
	14		
240		143	357
	12		
264		155	345
	9		
288		164	336
	13		
312		177	323

Estimate the mean time to failure of the part when operated under the specified conditions.

The total number of operating hours for all components may be determined by using the mid-ordinate rule to find the area under the curve of component survivors plotted to a base of time, as set out in the table below.

Time from start of test (hours) (1)	Number of survivors (2)	Average no. of survivors in each 24 hour period (mid-ordinate) (3)	Total operating time for all components in each 24 hour period (4)
0	500		
		490·5	490·5 × 24
24	481		
		473·5	473·5 × 24
48	466		
		459·5	459·5 × 24
72	453		
		444·5	444·5 × 24
96	436		
		430	430 × 24
120	424		
		416	416 × 24
144	408		
		402	402 × 24
168	396		
		389	389 × 24
192	382		
		376·5	376·5 × 24
216	371		
		364	364 × 24
240	357		
		351	351 × 24
264	345		
		340·5	340·5 × 24
288	336		
		329·5	329·5 × 24
312	323		
Total operating hours for all components = Σ Col. (4) = 5266·5 × 24			

From eqn. (28.4),

$$\text{Estimated mean time to failure} = \frac{5266{\cdot}5 \times 24}{177} = \underline{\underline{714 \text{ h}}}$$

28.4 Mean Time Between Failures (MTBF)

This is a term which is applied to repairable items, and is a measure of the average time that a particular equipment will remain in service. The MTBF of an equipment depends on the operating stresses, including the environmental conditions, but it may be reduced by potential defects introduced by poor maintenance procedures.

If the time between failures is long compared with the repair time, and if faults occur at times $t_1, t_2 \ldots t_n$, then

$$\text{MTBF} \approx \frac{1}{n}\sum_{k=1}^{k=n}(t_k - t_{k-1}) = \frac{t_n - t_0}{n} = \frac{t_n}{n}$$

since $t_0 = 0$. Hence

$$\text{MTBF} = \frac{\text{Total operating time}}{\text{Number of failures in that time}} \tag{28.5}$$

Comparing eqns (28.4) and (28.5) it can be seen that

$$\text{MTTF} = \text{MTBF} = \bar{m} = \frac{\text{Total operating time}}{\text{Number of failures in that time}} \tag{28.6}$$

It should be clear that a significant number of faults must be recorded in order to obtain a high confidence that the measured value of the MTBF is close to the true value. If, for example, an equipment which has an MTBF of 4,000 h is tested for only 1,000 h, there is a high probability that no failure will occur.

Instead of testing one equipment for a very long time, which would be impracticable in most cases, it is usual to test a number of equipments simultaneously for a shorter period each, and to determine the total number of faults in the total operating time of all the equipments. It must be understood that this method assumes that failures occur by chance at any time during the test and that wear-out failure has not occurred.

28.5 Failure Rate

Although *failure rate* is related to the number of failures per unit time, it is not defined simply as that number, because the number of items which fail in a given time depends, not only on the quality of the item and on the stresses applied to it, but also on the number of the components which are in operation. If the number of components in operation at the time of a failure is N_s the failure rate $\lambda(t)$ is defined by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{N_s} \frac{\Delta N_F}{\Delta t} = \frac{1}{N_s} \frac{dN_F}{dt} \tag{28.7}$$

where ΔN_F is the number of failures which occur in the time Δt.

The way in which $\lambda(t)$ varies with time depends on weaknesses in the part or equipment and on the operating stresses. In many cases the variation of $\lambda(t)$ takes the form shown in Fig. 28.1, which is often called the *bath-tub* curve. The curve shows three distinct phases in the life of the equipment. The first phase, the *early failure period*, is the time when very weak components fail. The weakness of these components is due to inevitable minor defects in the materials and in the manufacturing processes. These components are commonly removed by testing for a time t_1, with the result that the operational reliability of the equipment is improved. The second phase, the *constant failure rate period*, is the time during which

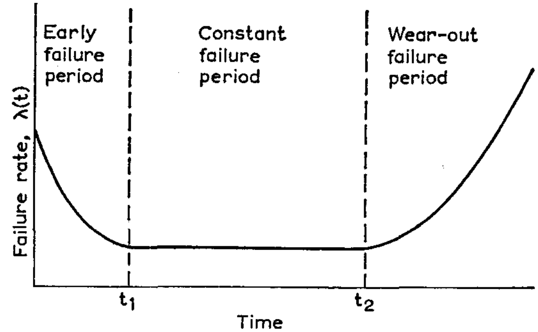


Fig. 28.1 THE BATH-TUB CURVE

the component or equipment is most usefully employed since the failure rate is lowest and the failures which occur are due only to fortuitous combinations of high stresses. This period is terminated by the *wear-out failure period* beginning at time t_2.

Preventative maintenance should be carried out before the time t_2 and the appropriate part replaced.

28.6 Constant Failure Rate Period

It will be shown that, if the failure rate is constant, then there is a simple relationship between the reliability at time t and the failure rate. For the purpose of this analysis, failures occurring in the early failure period have been eliminated and it has been assumed that the constant failure rate régime prevails, so that operational time is measured from t_1; i.e. at time t_1 in Fig. 28.1, $t = 0$. Then

$$\lambda(t) = \text{constant} = \lambda \text{ for } t < t_2 - t_1.$$

From eqn. (28.7), $$\lambda = \frac{1}{N_s}\frac{dN_F}{dt}$$

where N_S is the number which have survived up to time t, N_F is the number which have failed up to time t, and $N_0 = N_S + N_F$, where N_0 is the total number in operation at $t = 0$. Hence

$$\lambda = \left(\frac{1}{N_O - N_F}\right)\frac{dN_F}{dt}$$

or

$$\int_0^t \lambda\, dt = \int_0^{N_F} \frac{dN_F}{N_O - N_F}$$

Therefore

$$-\lambda t = \Big[\log_e (N_O - N_F)\Big]_0^{N_F}$$

and

$$e^{-\lambda t} = 1 - \frac{N_F}{N_O} \tag{28.8}$$

From eqn. (28.1),

$$R(t) = \frac{N_S}{N_O} = \frac{N_O - N_F}{N_O} \tag{29.9}$$

$$= 1 - \frac{N_F}{N_O} \tag{28.10}$$

Comparing eqns. (28.8) and (28.10), evidently

$$R(t) = e^{-\lambda t} \tag{28.11}$$

This equation does not apply to the early failure period nor to the wear-out period; a discussion of the reliability during these periods is outside the scope of this book.

28.7 Relation between MTTF, MTBF and Failure Rate

If failures are due to chance and if the failure rate is constant, then it is immaterial whether one equipment is tested for T hours or N equipments are each tested for T/N hours, since the probability of failure in a specified time will be the same in each case.

Also from eqn. (28.7),

$$\lambda = \frac{1}{N_S}\frac{dN_F}{dt} \qquad \text{so that} \qquad \int_0^t N_S\, dt = \frac{1}{\lambda}\int_0^{N_F} dN_F = \frac{N_F}{\lambda}$$

where N_F failures occur in time t. From eqn. (28.6),

$$\bar{m} = \frac{\text{total operating time}}{\text{number of failures}}$$

Hence

$$\bar{m} = \frac{1}{N_F}\int_0^t N_S \, dt = \frac{1}{\lambda} \tag{28.12}$$

It follows that the constant failure rate λ is

$$\lambda = \frac{1}{\text{MTTF}} \quad \text{for non-repairable parts}$$

or

$$\lambda = \frac{1}{\text{MTBF}} \quad \text{for repairable parts or equipments}$$

The basic unit of mean time used in reliability calculations is the hour, and the unit of failure rate is therefore the per-unit failures per hour. Because this is a very large unit, failure rate is also expressed as percentage failures in 1,000 h, as per-unit failures in 1,000,000 h, or as parts failing per 1,000,000 parts in 1 h.

EXAMPLE 28.3 One thousand similar equipments which are known to have constant failure rates of 5 per cent per 1,000 h are put into operation at the same time. Calculate the predicted times which will elapse before (*a*) 50 and (*b*) 500 equipments have failed in service.

The reliability is given by eqns. (28.10) and (28.11) as

$$R(t) = e^{-\lambda t} = 1 - \frac{N_F}{N_O}$$

$$\lambda = \frac{5}{100} \times \frac{1}{1{,}000} = 5 \times 10^{-5}$$

(*a*) Let t_1 hours be the time required for 50 failures to occur. Then

$$R(t_1) = \exp(-5 \times 10^{-5} t_1) = 1 - \frac{50}{1{,}000}$$

$$\exp(-5 \times 10^{-5} t_1) = 0{\cdot}95$$

$$5 \times 10^{-5} t_1 = \log_e\left(\frac{1}{0{\cdot}95}\right) \qquad t_1 = 1{,}020\text{ h}$$

Since $R(t_1) \approx 1$, in this case an approximate solution may be obtained by writing $e^{-\lambda t} = 1 - \lambda t$.

(*b*) Let t_2 hours be the time required for 500 failures to occur. Then

$$\exp(-5 \times 10^{-5} t_2) = 1 - \frac{500}{1{,}000}$$

so that

$$5 \times 10^{-5} t_2 = \log_e 2 \qquad \text{and} \qquad t_2 = \underline{\underline{13{,}900\text{ h}}}$$

Note that since 50 out of 1,000 equipments fail during the first 1,000 h, we may say that the probability of an equipment failing in that time is 5 per cent. Also, the probability is 50 per cent that an equipment will fail during the first 13,900 h.

While it is impossible to eliminate failures entirely the chances of a failure may be made very small by designing the equipment to have a very large MTBF compared with its operating time. In the example above, if the operating time of the equipment were only 1 hour then the reliability $R(t) \approx 1 - 5 \times 10^{-5}$.

This means that, on average, there will be only 50 failures in every million equipments, and it follows that most groups of 1,000 equipments will not have any failures during the first hour.

The acceptable value for the reliability of an equipment at a given time depends on many factors, such as human safety, cost of repair and operational usefulness.

There are many examples of the relationship between reliability and human safety, but the most dramatic is that of astronauts. In this case the unreliability of the component parts of the system must be made so low that failure becomes only a remote possibility.

Repeaters which are used in underwater telecommunication systems must operate without failure for a considerable time compared with the duration of a space mission. In this case it is the very high cost of repair and the cost of the *down time* (the period during which the equipment is non-operational) that dictate the need for high reliability.

A third example, that of a digital computer used for scientific or business applications, is also built with a high reliability specification. However, neither safety nor very high repair costs are important in this case, but the cost of down time in large computers is considerable, and thousands of calculations must be made without error; thus operational usefulness is the important characteristic in this case.

28.8 Series Reliability

If a number of parts of a system are operated in such a way that the failure of any one part causes a failure of the system, then those parts are considered functionally to be in series. If a failure of any part is independent of the operation of the other parts then the reliability of the system is given by the product of the reliabilities of the parts. Thus,

$$R(t) = R_1(t) \times R_2(t) \times R_3(t) \ldots R_k(t) \ldots R_n(t) \tag{28.13}$$

where $R(t)$ is the system reliability and $R_k(t)$ is the reliability of the kth part at a time t.

If, in particular, λ is constant then

$$R_k(t) = e^{-\lambda_k t}$$

From eqn. (28.13),

$$\begin{aligned} R(t) &= e^{-\lambda_1 t} \times e^{-\lambda_2 t} \ldots e^{-\lambda_k t} \ldots e^{-\lambda_n t} \\ &= e^{-(\lambda_1+\lambda_2+\cdots+\lambda_k\cdots+\lambda_n)t} \qquad (28.14) \\ &= e^{-\lambda t} \end{aligned}$$

where

$$\lambda = \lambda_1 + \lambda_2 \ldots + \lambda_k \ldots + \lambda_n \qquad (28.15)$$

is the system failure rate.

Note that the condition of independence is not always fulfilled. A secondary failure is an example of one failure occurring because of another.

28.9 Parallel Reliability

An important method of improving the reliability of a system is the use of redundancy, i.e. the inclusion of additional equipment in such a way that a failure of one part does not cause a failure of the whole system. Redundancy may be used in the parts of an equipment or in the complete system. For example, if two components are connected in parallel the equipment can be designed to function with an open-circuit in one of the components. However, in this case, the stresses in the good component may be increased, the equipment performance may be degraded, and the faulty component must be repaired or replaced quickly if the advantages of redundancy are to be restored. As a second example, consider the use of computers in an air traffic control system. For reasons of safety it is necessary that the computing system should function continuously, and however high the reliability of one computer, these requirements can only be met by duplication of the computer.

If the reliability of each part is assumed to be independent of the other parts then the probability of a system failure is

$$Q(t) = Q_1(t) \times Q_2(t) \times Q_3(t) \ldots Q_k(t) \ldots Q_n(t) \qquad (28.16)$$

where $Q(t)$ is the system unreliability and $Q_k(t)$ is the part unreliability. Since $R(t) = 1 - Q(t)$ the system reliability is easily calculated.

EXAMPLE 28.4 After 100 h an equipment has a reliability of 0·75. Calculate the overall reliabilities which would be obtained if the equipment were (*a*) duplicated and (*b*) triplicated.

Since $R_1(100) = 0{\cdot}75$, $Q_1(100) = 0{\cdot}25$. Assume that

$$Q_1(t) = Q_2(t) = Q_3(t)$$

(*a*)
$$Q_D(100) = [Q_1(100)]^2 = 0{\cdot}0625$$

Therefore

$$R_D(100) = 1 - Q_D(100) = \underline{\underline{0{\cdot}9375}}$$

(*b*)
$$Q_T(100) = [Q_1(100)]^3 = 0{\cdot}0156$$

Therefore

$$R_T(100) \doteq \underline{\underline{0{\cdot}9844}}$$

Hence, if an equipment has a reliability of 75 per cent, two such equipments used in such a way that the system fails only if both fail, will have a reliability of 93·8 per cent, and three equipments a reliability of 98·4 per cent.

28.10 Environmental Conditions

Most components or equipments having a high reliability when operated under well-controlled laboratory conditions are found to have a greater failure rate when they are subjected to increased stress of one kind or another. The following environmental conditions for both storage and operation of components and equipments must always be taken into account.

Extremes of *temperature* invariably reduce reliability because of (*a*) the effects of expansion and contraction of materials with change of temperature, (*b*) changes of component values, (*c*) melting, softening or freezing of some component materials, and (*d*) the effect of temperature on chemical action.

High humidity is another cause of increased failure rate, particularly when it is associated with high temperatures. Under these conditions a thin film of water can form on a component, and printed-circuit board surfaces can become ionized and thus form a conducting path. This path, together with the capacitive path provided by the high relative permittivity of water, is likely to cause a component defect. Moisture can also enter equipment by diffusing through a material of which it is made or by entry through a hole in the sealing of the equipment.

Other effects which may be associated with the temperature and humidity effects are *dust*, *air pressure*, *salt spray* and *mould growth*. Dust can reduce surface insulation resistance and spoil the performance of lubricants; low pressure, by causing the release of water vapour or other trapped gases, can change the electrical properties of the component; it can also affect cooling by convection. Salt spray may produce rapid corrosion.

Vibration is a serious cause of unreliability unless the equipment is very carefully designed to cope with it. The important characteristics of the vibration are the amplitude and frequency of its periodic components. The mechanical resonance of parts in the equipment must be designed to be at frequencies outside the range of the expected vibrations and they must be suitably damped.

28.11 Reliability Assessment

From what has already been written it should be clear that no single value of failure rate can be given for a component because the value changes under different stress conditions. Although a simple relationship between failure rates at two stress levels should not be taken for granted, it is found that in many cases the failure rate at one stress condition may be calculated from the failure rate at another by the use of weighting factors. Three important weighting factors are (1) *environment weighting factor* (EWF), (2) *rating weighting factor* (RWF) and (3) *temperature weighting factor* (TWF). Hence

$$\text{Failure rate} = \text{Basic failure rate} \times \text{EWF} \times \text{RWF} \times \text{TWF} \qquad (28.17)$$

Figures for basic failure rates and for various weighting factors have not been included in this book because it is important always to use the latest available information. This should be obtained from recently published papers, and from manufacturers' technical information services.

The temperature at which a component operates will depend on (i) the ambient temperature, (ii) self-heating, (iii) heating from other components in the equipment, and (iv) the method of cooling used. Devices such as resistors and transistors, which are given a specified power rating at 25°C ambient temperature, must be derated as the ambient temperature is increased.

The method of calculating the power dissipated in a device must take into account the permitted tolerances of all the components in the circuit and of the supply. The calculation is usually based on either *worst-case design* or *statistical design*.

In the worst-case method, all the component values are assumed to be at an extreme acceptable value and are so chosen that the circuit is most likely to fail. This idea will be illustrated by a very simple example.

EXAMPLE 28.5 A d.c. power supply which has an output of 24 V $\pm$ 10 per cent is connected to a resistor $R_1 = 100\,\Omega \pm 20$ per cent in series with a resistor $R_2 = 200\,\Omega \pm 10$ per cent. Determine the worst-case dissipation of power in R_2.

If the power dissipated in R_2 is P_2, then

$$P_2 = \frac{V^2 R_2}{(R_1 + R_2)^2}$$

where
$$V = V_0(1 \pm a) = 24(1 \pm 0{\cdot}1)\ \mathrm{V}$$
$$R_1 = R_{10}(1 \pm a_1) = 100(1 \pm 0{\cdot}2)\ \Omega$$
$$R_2 = R_{20}(1 \pm a_2) = 200(1 \pm 0{\cdot}1)\ \Omega$$

Hence

$$P_2 = \frac{V_0^2 R_{20}(1 \pm a)^2(1 \pm a_2)}{(R_{10} + R_{20} \pm \Delta R_1 \pm \Delta R_2)^2}$$

$$= P_{20} \frac{(1 \pm a)^2(1 \pm a_2)}{\left[1 \pm \left(\dfrac{R_{10}}{R_{10} + R_{20}}\right) a_1 \pm \left(\dfrac{R_{20}}{R_{10} + R_{20}}\right) a_2\right]^2}$$

where $P_{20} = V_0^2 R_{20}/(R_{10} + R_{20})^2$.

The power in R_2 will be a maximum when V is greatest and when R_1 is least, but the correct limit for R_2 is not so obvious. However, from the maximum power transfer theorem (Section 2.4) it can be deduced that the power in R_2 is greatest when the value of R_2 is closest to the value of R_1, and in this case the lower limit of R_2 should be chosen. Hence

$$P_2 = \left\{\begin{matrix}\text{Power dissipated by } R_2\\ \text{when all components}\\ \text{have nominal values}\end{matrix}\right\} \times \text{Tolerance factor}$$

In this example,

$$P_{20} = \frac{24^2 \times 200}{300^2} = 1{\cdot}28\ \mathrm{W}$$

and

$$\text{Tolerance factor} = \frac{1{\cdot}1^2 \times 0{\cdot}9}{(1 - 0{\cdot}067 - 0{\cdot}067)^2} = 1{\cdot}45$$

Hence

$$(P_2)_{worst\ case} = 1{\cdot}28 \times 1{\cdot}45 = \underline{\underline{1{\cdot}85\ \mathrm{W}}}$$

Thus the worst-case power dissipation in R_2 is 45 per cent higher than the power dissipated if all the components and the supply had actual values equal to the nominal design values.

The solution given above illustrates a general method. In this particular case a more direct method could have been used.

The disadvantage of using worst-case design is that components with high ratings must be used to take into account the unlikely event that all the tolerances are at their extreme values. Some reduction in the ratings is obtained by using components with closer tolerances, and where this is not possible, as with transistors, by using special techniques which make the circuit operation largely independent of characteristics of the device.

However, the cost of a component and the probability of a degradation failure are both increased by a reduction in the permitted tolerance. It is for these reasons that statistical design should be considered. In this case it is assumed that extreme values will not occur in the same circuit, and the design is based on a calculated probability that certain ratings will not be exceeded. A study of this method of design is beyond the scope of this book.

28.12 Maintainability

The number of times that an equipment becomes faulty is a direct function of its unreliability, but the length of time for which an equipment is not available is a function both of unreliability and of the time it takes to repair it and return it to service. *Maintainability* is a characteristic of an equipment related to the ease with which it can be repaired. There are four parts to the problem of achieving good maintainability:

(i) The method of determining that a defect exists. It may be obvious to a viewer that a defect exists in his television receiver, but it may not be so obvious that the arithmetic unit in a computer is producing some wrong answers.

(ii) The method of quickly identifying a defective component or assembly. The solution to this problem includes the training of servicemen, provision of test apparatus and test procedures, clear identification of assemblies and of parts, provision of test points and also of purpose-built test equipment. It is very important that designers should consider maintainability from the beginning of the project and not add service extras to the equipment as an after-thought.

(iii) Rectifying the fault, which may include replacement of a defective part. All components should be made as accessible as possible, particularly those known to have a high failure rate. Wiring looms should not be positioned so that they prevent the easy removal of an assembly, and if a large number of soldered connections have to be removed then this operation may well produce subsequent faults in the equipment. Preset controls should be sited where adjustments may be made without undue difficulty, but on the other hand they should not be directly accessible because they would then be open to misuse.

(iv) Finally it will be necessary to verify that the system is functioning correctly after the completion of the repair.

A useful definition, of maintainability is:

"Maintainability is the probability that a device will be restored to operational effectiveness within a given period of time when the maintenance action is performed in accordance with prescribed procedures".*

28.13 Reliability Costs

It should be clear that, in general, designing an equipment for high reliability and building a system with a guaranteed reliability will add to the initial cost. However, this higher initial cost will be offset, to a greater or lesser extent, by reduced maintenance costs for any given degree of maintainability. Reduced maintenance costs arise not only because of a reduction in man-hours devoted to maintenance but also due to a reduction in the provision of test equipment and stock of spares required.

The graph of costs against reliability takes the form shown in Fig. 28.2.

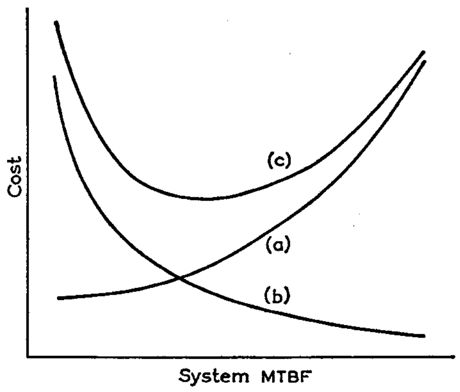


Fig. 28.2 THE COST OF UNRELIABILITY

(*a*) Initial cost
(*b*) Maintenance cost
(*c*) Total cost

* CALABRO, S. R., *Reliability Principles and Practice*, Chap. 9 (McGraw-Hill, 1962).

It is obvious that minimum total cost is not always the most important consideration and that safety must be a first priority.

A part of the cost of achieving high reliability is bound up with the problems of testing components and equipments and analysing these data. If only those components which have been adequately tested and approved are used in a design, then one part of this cost will be kept to a minimum and it will be possible to obtain increased confidence in the reliability of the components.

High reliability is possible only if everyone connected with the equipment plays his part. A reliability statement must be regarded as a part of the specification as important as, say, the function details, and must be designed into the system from the outset. A realization of the relation between the part and its probable total environment is very important. This includes, not only the environmental factors considered above, but also the important part that *human engineering* should play. The position of the equipment in relation to other equipments, the design and position of the controls, the type of indicators and the methods of labelling are all obvious considerations but they are all too often overlooked.* All production personnel should be trained to understand fully the importance of reliability. Components and assemblies may be damaged by poor handling, and bad joints are more easily corrected during manufacture than later, during test.

A programme of *tests* and *quality control* must be carefully considered and carried out by trained technicians of high professional integrity. Also, care must be exercised to see that stresses set up during the tests do not make the components potentially unreliable.

The MTBF of an equipment will be reduced seriously unless sufficient attention is paid to the hazards of *transport* and *installation.* Consideration must also be given to various environmental conditions during transport and storage as well as to the need to protect the equipment from mechanical damage. Rules for installation should be worked out with the known variations in human characteristics kept in mind. *Installation engineers* should not have to be weight lifters, contortionists or giants in order to do their job, and as with other members of the team, they should be made aware of the importance of their part in ensuring the reliability of the system. In particular, they should pay attention to problems of dust, cooling and the protection of interconnecting cables. The *customer* is an important link in the reliability chain because incorrect operation of the

* DUMMER, G. W. A., and WINTON, R. C., *An Elementary Guide to Reliability* (Pergamon Press, 1967).

equipment will usually reduce its reliability. This again is a matter of training, but a good instruction manual and sufficient attention to the ergonomic problems involving the use of the equipment are also of importance. This means that attention should be given to operator fatigue and error due, for example, to the operator having to stand rather than sit, to controls and dials located in awkward positions, to excessive noise, to too little or too much light or heat, to draughts and vibration. While it is not always possible to achieve a perfect operator environment, a considerable improvement can be made by giving the problem careful thought. The term *operability* is used to describe that characteristic of an equipment which is measured by the probability of an operator not making an error in the use of the equipment.

Finally the importance of accurate recording of *fault data*, and the speedy *feedback of information* to the designer, cannot be overstated. It is by these means that the designer can investigate apparent weaknesses in his equipment, and then take remedial action. The report form should be designed with care. It should be as simple as possible, but it should include questions which will give the essential information about the type of defect found, the environmetal conditions, and whether the failure is a primary or secondary one.

PROBLEMS

28.1 A radio receiver has an MTBF of 1,500 h. What is the probability of a failure occurring during the period of the Wimbledon tennis championships?

Assume that the receiver is functioning at the beginning of the period and that it is switched on for a total of 40 h. State any other assumptions which must be made.

Ans. 0·0263.

28.2 A reliability test was made on a batch of 2,000 similar paper-dielectric capacitors at a temperature of 70°C and at the rated voltage. The number of survivors at the end of the 100th hour was 1,960, and at the end of the 200th hour was 1,930. What was the average failure rate over the second 100 h period? Why is the measurement of low failure rate a costly procedure?

Ans. 15·5 per cent per 1,000 h.

28.3 If the failure rate of the capacitors in Problem 28.2 is assumed to be constant, calculate the reliability of the capacitors at the end of the first 500 h period of test. Why is this assumption of constant failure rate probably wrong? Is the true reliability during the first 500 h period likely to be greater or less than that calculated in the first part of the problem?

Ans. 0·92.

28.4 A particular circuit which consists of two transistors and six resistors is operated under conditions for which the MTTF of the transistors is 2×10^6 h and that of the resistors is 10^6 h. Calculate the MTBF of the circuit. What is the reliability of the circuit after 9,000 h?

Ans. 143,000 h; 0·939.

28.5 An equipment uses 20 of the circuits described in Problem 28.4. Calculate the reliability of the equipment after 9,000 h assuming that a defect in any one circuit causes a failure in the equipment.

Ans. 0·284.

Index

Higher Electrical Engineering has come to be accepted as the standard work for the HNC and HND in Electrical and Electronic Engineering and also for much first-year degree-level study. This second edition is in SI units throughout and contains chapters on reliability and logic. As before, there is a large number of worked examples and problems with answers at the end of each chapter.

An introduction to the generalized theory of machines has been given, but conventional treatments, substantially revised, of synchronous and induction machines have been retained. In electronics, the shift of emphasis to semi-conductor devices is reflected in the omission of valve circuits from large sections of the text, in favour of bipolar transistor circuits. The field-effect transistor is also dealt with.

Reviewing the first edition, the *Journal of the Institution of Electrical Engineers* said: 'The selection of material, its arrangement and presentation are excellent and reflect the authors' experience of teaching the subject'.

ISBN 0 273 40063 0

Higher Electrical Engineering

SHEPHERD
MORTON
SPENCE

PITMAN