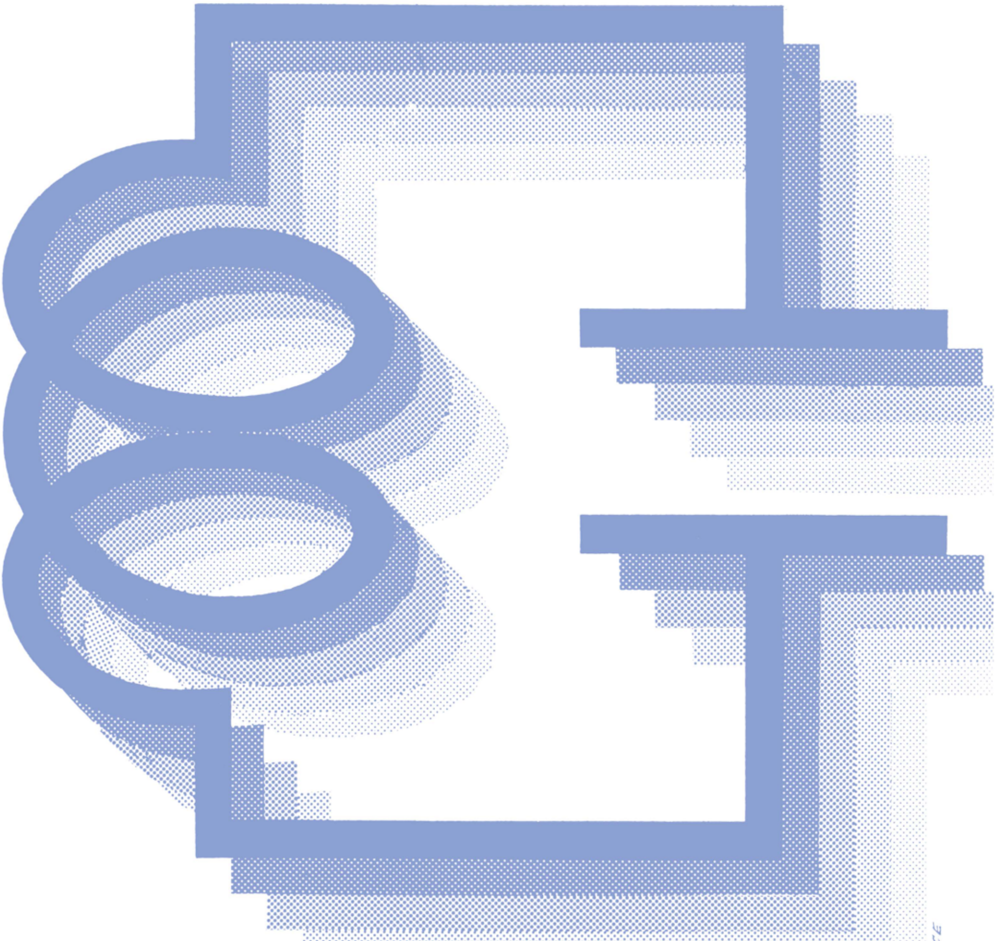
TRANSISTORS IN PULSE CIRCUITS

G. FONTAINE



PHILIPS TECHNICAL LIBRARY

TRANSISTORS IN PULSE CIRCUITS

This book is the fourth in a series by the same author published or to be published in the Philips Technical Library

DIODES AND TRANSISTORS

General theory

AUDIOFREQUENCY TRANSISTORS

RADIOFREQUENCY TRANSISTORS

Amplifiers and oscillators

TRANSISTORS IN PULSE CIRCUITS

PHILIPS TECHNICAL LIBRARY

TRANSISTORS IN PULSE CIRCUITS

G. FONTAINE

Macmillan Education

Softcover reprint of the hardcover 1st edition 1971 978-0-333-02274-0

SBN: 333 022742

ISBN 978-1-349-00168-2 ISBN 978-1-349-00166-8 (eBook)
DOI 10.1007/978-1-349-00166-8

First published in English by
MACMILLAN PRESS LTD
London and Basingstoke
Associated companies in New York, Toronto, Melbourne,
Dublin, Johannesburg and Madras

PHILIPS

FOREWORD

With the wider application of electronic systems in industry, and the ever-increasing use of computers, all types of engineers and technicians must be well-informed on the operating principles and the nature of circuits and components.

Information to be produced, transmitted or transformed is now nearer to a square or triangular signal, rather than the sinusoidal oscillations already met with in audio- and radiofrequency. These new signals have their own characteristics, parameters and behaviour patterns, and it is the purpose of this book to study and explain these. Moreover, semiconductors (diodes, transistors, integrated circuits) can distort the functioning of equipment, and a thorough understanding of the basic physical principles involved is necessary to explain these effects and to find ways of reducing or eliminating them.

This book is divided into three main parts: the first deals with factors relating to the purely electrical behaviour of diodes and transistors, and defines the principle 'switching' parameters, i.e. forward and reverse recovery time of the diodes, rise and fall time of the current, delay and desaturation time, etc. The second part outlines the limiting operating conditions of semiconductors by linking the considerations of maximal power and critical temperature. The third and final part of the book is a thorough study of the principal components of switching circuits: astable, monostable and bistable multivibrators, astable and triggered blocking oscillators.

This knowledge is essential for the design, construction and maintenance of components or equipment for numerical and logical systems. For the sake of simplicity, however, we have omitted complex mathematical treatments, as the design of a multivibrator and the definition of its limiting operating conditions can be worked out by very simple equations.

Furthermore, the systematic examination of density graphs enables useful and practical conclusions to be drawn from an apparently qualitative study; a thorough and detailed explanation of all the terms likely to be met in the literature will be found in this book.

G. FONTAINE

August 1969

CONTENTS

Chapter 2 — TRANSISTORS FOR SWITCHING

Chapter 4 — THERMAL BEHAVIOUR OF SWITCHING TRANSISTORS — LIMITING OPERATING CHARACTERISTICS

Chapter 5 — SWITCHING CIRCUITS AND MULTIVIBRATORS

1. Diodes for Switching

1.1. *PN* junction

Diodes and transistors consist essentially of junctions associated with resistive elements. Their behaviour depends on the signal applied to their terminals. The behaviour of these junctions depends on the behaviour of the signal applied to their terminals and may produce capacitance effects with small signals or inductive effects with large signals. Before explaining the origin of these different effects, knowing the importance of the majority and minority carrier density graphs, a discussion of the physical principles which help to study the static behaviour of a junction would seem appropriate.

Modern diodes and transistors may be divided into two main classes: germanium diodes and transistors on the one hand, and silicon diodes and transistors on the other. For easing the discussion, we will examine the behaviour of a *PN* junction knowing that the *P* region may consist of germanium or *P* type silicon, and the *N* region of germanium or *N* type silicon. Fig. 1 shows a *PN* type junction clearly.

To explain the various currents which may flow through the circuit associated with this junction it is necessary to use majority and minority carrier density graphs as shown in Fig. 2. The *P* region contains 10^{16} mobile holes per cubic centimetre, known as 'majority carriers', and the *N* region 10^{16} free electrons per cubic centimetre; on the other hand these two regions contain respectively 10^{10} free electrons (in the *P* region) and 10^{10} mobile holes (in the *N* region). These two regions are separated by a region containing fewer carriers, known as the 'barrier region'.

Difference of impurity concentration

In practice, modern diodes and transistors have their *N* region more weakly doped than the *P* region. Let us therefore assume an *N* region containing only 10^{14} free electrons; there will then be only 10^{12} mobile holes per cubic centimetre.

The appearance of the *PN* junction is then as shown in Fig. 3. The *P* region, on the left, is very rich in holes (majority carriers), but the *N* region, on the right, is proportionately much less rich in electrons.

The densities of minority carriers (electrons in the *P* region and holes in the *N* region) are represented by horizontal graphs located at exactly opposite levels with respect to those of the corresponding majority carrier density graphs.

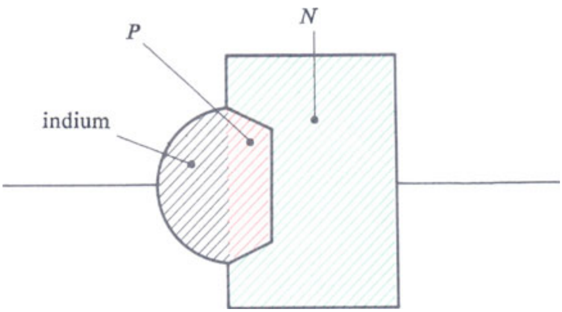


Fig. 1

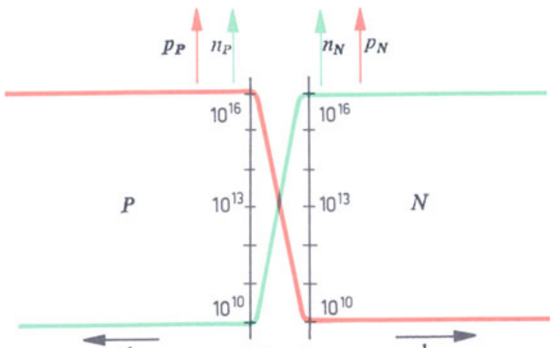


Fig. 2

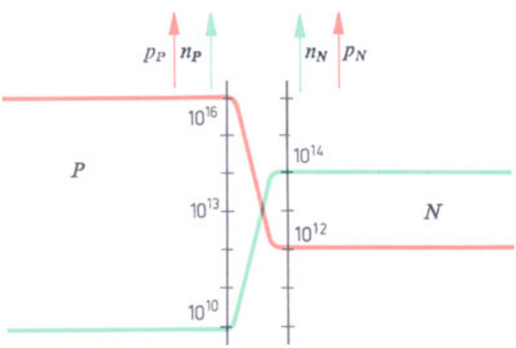


Fig. 3

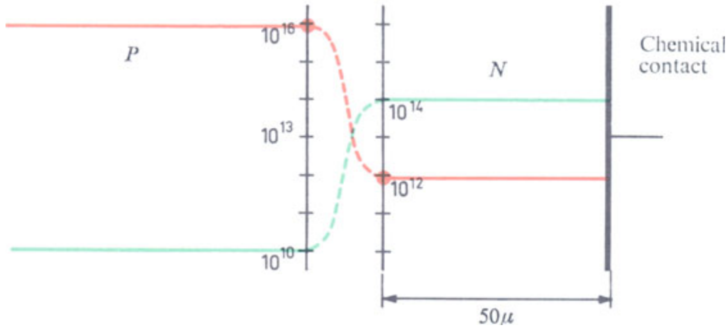


Fig. 4

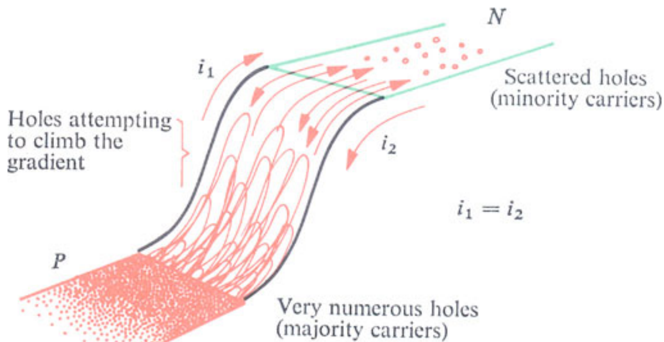


Fig. 5

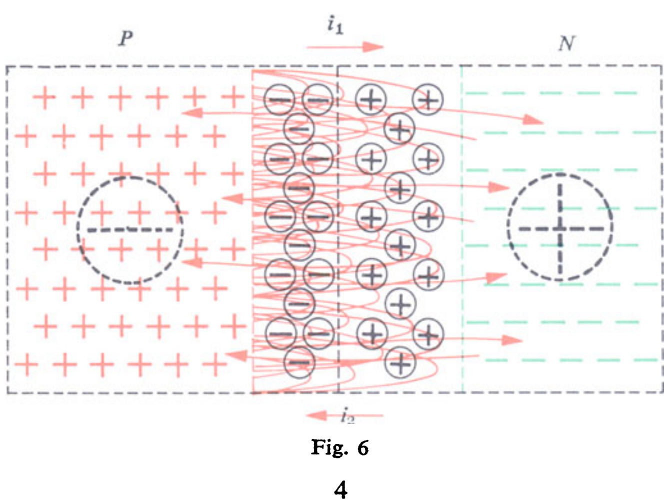


Fig. 6

Narrow N regions

In order to approach the practical case more closely, let us consider a very thin N region (approx. 50 µm). This region (Fig. 4) may at the same time represent the base of an alloyed transistor or the N region of a diode designed primarily for use in switching. It is, however, possible to find diodes with a wider N region: their behaviour in dynamic conditions does not differ from that of diodes with a very thin N region. In this case artificial recombination centres, obtained by the injection of atoms of gold, are often located very near the barrier.

1.2. Unbiased junctions

In the absence of bias a simplified explanation is that no majority carrier can pass through the barrier; the current in the external circuit is consequently zero. However, this reasoning is faulty because it does not explain the presence of a current in a certain state of the junction bias. The analogy shown in Fig. 5 is, therefore, of great interest.

This analogy consists in likening the two regions, P and N, to two planes; one red (P region) containing a very large quantity of balls; the other green (N region) located at a higher level and containing very few balls. The diffusion of the holes from the P region towards the N region is represented by a displacement of balls from the lower to the higher plane. Simultaneously the balls from the higher plane fall back into the lower plane. The state of equilibrium corresponds to a gradient such that the number of balls passing from the red plane to the green plane is exactly equal to the number of balls falling back from the green plane to the red plane. In these conditions the final current is zero.

The diagram in Fig. 6 is a complete physical representation of this analogy. The current i_1 caused by the diffusion of holes is practically equal to the current i_2 of minority holes in the other direction.

The density of the holes moving in the barrier decreases as we pass from left to right, that is from the P to the N region. The same applies to the number of paths of the balls in Fig. 5; this number increases as the level considered approaches the red plane representing the P region.

A rapid examination of Fig. 5 confirms the absence of linearity between the variations of the number of balls in movement in the transition zone and the possible differences of height.

The majority and minority carrier density graphs in Fig. 7 represent this situation exactly. The density of mobile holes, i.e. minority carriers in the *N* region, is equivalent to the equilibrium level 10^{12}. The maintenance of this density at its equilibrium level in the neighbourhood of the barrier depends on whether the departure of the minority holes from the *N* region is compensated by the arrival in this region of majority holes from the *P* region. The currents i_1 and i_2 are equal; the resultant current is, therefore, zero.

The explanation given above to illustrate the movement of holes in either direction through the barrier applies also to the study of the movement of electrons. Using the same principle, the analogy in Fig. 8 represents two planes: one containing a very large number of green balls (*N* region), and the other, at a higher level, fewer balls (*P* region). The state of equilibrium corresponds to a gradient between these two planes such that the number of balls reaching the higher plane (current i_3) is exactly equal to the number of balls falling back to the lower plane (current i_4). In these conditions, the current also is equal to zero.

A more physical explanation of this effect is supplied by Fig. 9. A very large number of electrons leaves the *N* region with paths which become more dense the further the region considered is situated to the right of the barrier. In the state of equilibrium the width of the two space charge zones is considerable and the majority of electrons return to the *N* region (diffusion current i_3 limited to the passage of only two electrons). The minority carriers, i.e., electrons in the *P* region, do not have to overcome the forces represented by these regions and they pass easily through the barrier in the other direction (current i_4).

Since the electron currents i_3 and i_4 are equal, the resulting current is zero. Owing to the intensive use we will be making of the density graphs, it is important to associate with this value of zero current the effective position of the density graph of the minority carriers (electrons) in the interior of the *P* region.

The differences of level between the various densities in the *P* and *N* regions, and especially the surplus of holes in the *N* region with respect to the electrons in the *P* region, draw immediate attention to the dominant influence of the right side of the *PN* junction.

For certain types of diodes and transistors the concentrations of impurities injected into the *P* regions are greater, which again accentuates the differences of the densities between the majority and minority carriers.

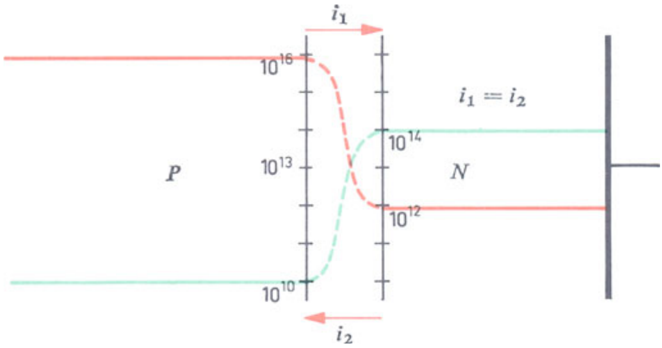


Fig. 7

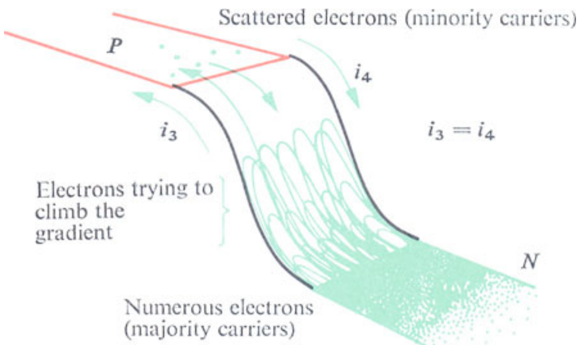


Fig. 8

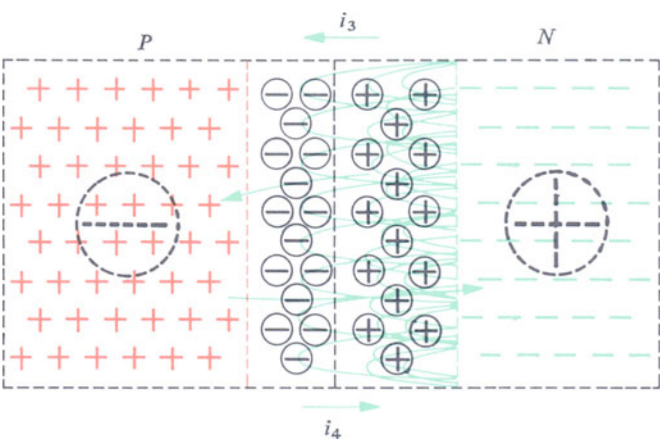


Fig. 9

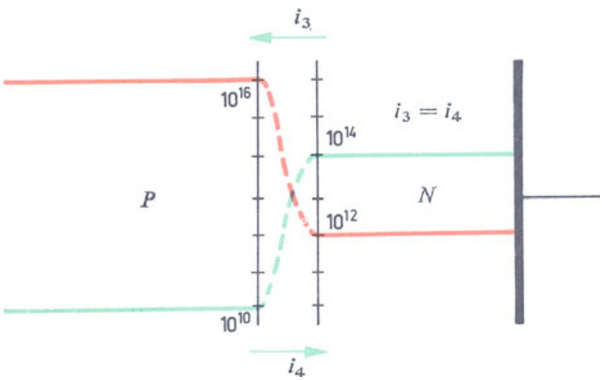


Fig. 10

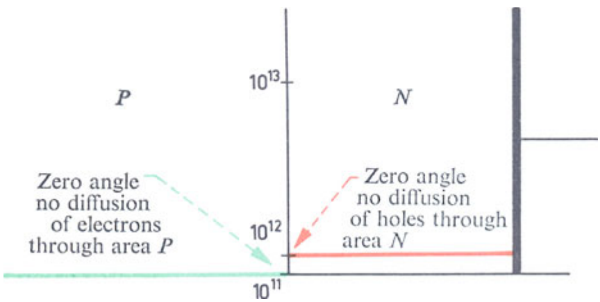


Fig. 11

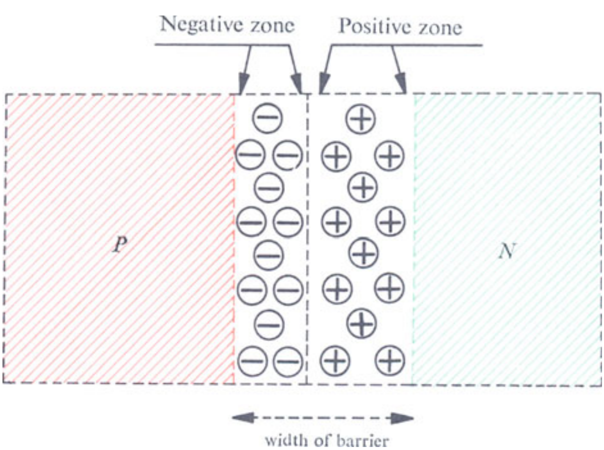


Fig. 12

In the density graphs of the majority and minority carrier shown in Fig. 10, the density of electrons (i.e. minority carriers in the *P* region) is at its equilibrium level in the neighbourhood of the barrier. The electrons which leave the *P* region are replaced by electrons originating in the *N* region. Finally, the equilibrium positions of the minority carrier densities in the neighbourhood of the barrier should be associated with exactly equal currents of majority and minority carriers, hence with a zero current in the external circuit.

Density graphs on a linear scale

At this stage of our discussion the majority carriers are of no interest; moreover, the variations in the width of the barrier are only indirectly dependent on the quantitative expression of the current in the external circuit.

In order to simplify the later stages of this study, it will be enough if we show only the density graphs of the minority carriers in the *P* and *N* regions, with the scale of the ordinates graduated linearly (Fig. 11).

The two graphs are horizontal. This situation corresponds to minority carrier densities exactly equal to the equilibrium densities peculiar to the regions located outside the space charge zones constituting the barrier. The zero value of the current in the external circuit must be associated with these two horizontal graphs. It will, in fact, be found that a more quantitative study will give us equations in which the expression for the current will vary directly with the tangents of the angles between these graphs and the horizontal. In the absence of bias the angles are equal to zero, the tangents are also zero and there will be no current.

Before we pass to the study of the effect of bias on the behaviour of the junction, we must emphasize the importance of the two space charge zones. These are the cause of the limitation of the diffusion currents, and in these conditions they are responsible for the absence of current in the external circuit.

The biasing of a junction consists in the modification of the thickness of this barrier.

A forward bias implies the narrowing of the space charge zones, whilst a reverse bias causes it to widen.

The origin of the space charge zones is explained by the removal of the majority carriers in the neighbourhood of the transition region; any variation in the thickness of these two regions consequently results in the arrival or removal of majority carriers.

1.3. Biased junctions

A junction is biased by means of a battery connected at the ends of the *P* and *N* regions. The biasing can therefore assume two forms; we will first consider forward biasing.

Forward biasing

Figure 13 shows a junction biased in the forward direction. The positive pole of the battery is connected to the left end of the *P* region and the negative pole to the right end of the *N* region. The left section of the *P* region is now at a positive potential, and all the holes move to the right (due to the repulsion of charges of the same sign).

Simultaneously the negative potential of the right end of the *N* region causes the group of negative charges (i.e. free electrons) to move to the left. The holes, i.e., positive charges, enter the negative space charge zone and partly neutralize it; the electrons, i.e., negative charges, reach the positive space charge zone and reduce its thickness. The barrier is now narrower. . . . What influence has this reduction of the barrier on the diffusion currents?

This contraction may be likened to a decrease in the distance between the two red and green planes in Figs 14 and 15.

We will now consider Fig. 14. A smaller change in level between the two planes allows the passage of a large number of red balls (*P* region) on to the green plane (*N* region). The fall of the balls from the green plane to the red plane has not changed; consequently the currents i_1 and i_2 are very different. This effect corresponds to much larger movements of holes from the *P* region to the *N* region than from the *N* region to the *P* region. The above explanation applies also to the study of the movement of electrons from the *N* region to the *P* region (Fig. 15).

Density graphs plotted logarithmically. Forward biasing affects the position of the density minority carrier graphs. The broken lines in Fig. 16 should be interpreted in connection with the absence of bias, and those in full lines with the forward biasing of the junction.

The importance of the diffusion currents with respect to the currents of minority carriers explains this displacement upwards of the two density graphs of the minority carriers.

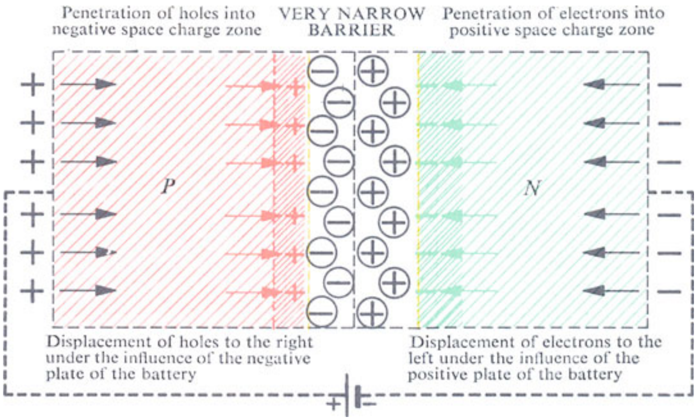


Fig. 13

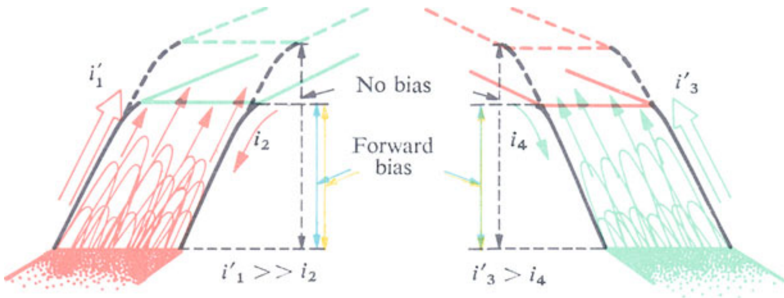


Fig. 14

Fig. 15

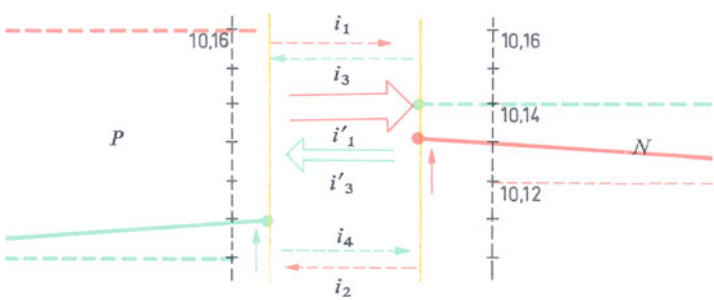


Fig. 16

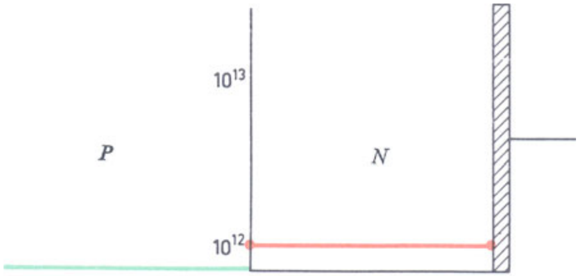


Fig. 17

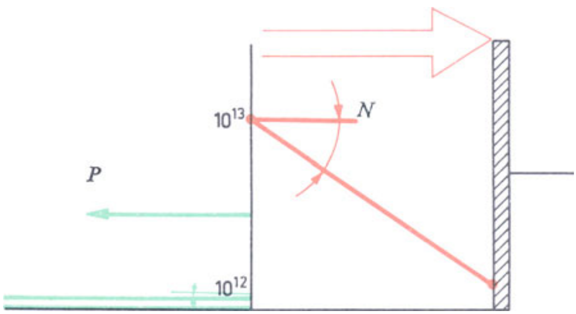


Fig. 18

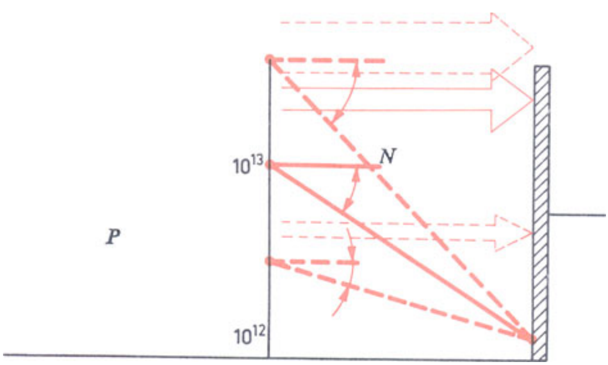


Fig. 19

The changes in the densities of the minority carriers in the neighbourhood of the barrier enable us to calculate the currents of holes and electrons diffusing into the *P* region from the *N* region and vice versa. Because of the highest position of the equilibrium density of the minority carriers in the *N* region with respect to the density of electrons in the *P* region, with the ordinate graduated logarithmically, the same difference in the carrier density is associated with larger hole currents flowing from the *P* to the *N* region than of electrons from the *N* to the *P* region.

The determination of the current in a circuit consisting of a battery and a diode is thus facilitated by the use of density graphs for the majority and minority carriers; in this case the ordinate is on a logarithmic scale. The calculation of this current may be still further simplified by the use of minority carrier density graphs on a linear scale.

Density graphs on a linear scale. When a junction is not biased the minority carrier density graphs will be completely horizontal. Such diagrams are shown in Fig. 17, in which the hole density is at the level 10^{12} throughout the *N* region, and the electron density at the level 10^{10} in the *P* region (the corresponding line is superimposed on the abscissa). The angles between these graphs and the horizontal are zero, and the current associated with the value of the tangents of these angles is also equal to zero.

The forward biasing of the junction (Fig. 18) causes the density graphs to move upwards near to the barrier; these movements correspond to an increase in the minority carrier densities. The currents of holes in the *N* region and of electrons in the *P* region may now be determined by calculating the tangents of the angles of the hole density graphs in the *N* region and of the electron density graphs in the *P* region.

In this type of junction the electron current is practically negligible by comparison with the hole current.

Any increase in the forward bias of the junction causes an increase in the number of injected holes, associated with a larger angle and a higher current. Conversely, a decrease in the forward bias causes a decrease in the angle and, therefore, of the hole current in the interior of the *N* region.

Reverse bias

The second method of using a junction involves biasing it in the reverse direction. The positive pole of a group of batteries, located at the right of the *N* region (Fig. 20) causes an attraction of the negative charges (i.e. free majority electrons in this region); as a result these electrons leave the left end of the *N* region, and the positive space charge zone in the *N* region side becomes wider.

The negative pole, located at the left end of the *P* region, produces a similar attractive effect on the holes. A large number of holes leaves the right end of the *P* region; the negative space charge zone widens. Reverse biasing of a junction is accompanied by a widening of the barrier. How does this widening affect the value of the currents in the barrier?

Reverting to the above analogy, Fig. 21 represents a junction consisting of two planes separated by a much steeper gradient than that shown initially. The widening of the barrier is, therefore, associated with an increase in this gradient.

The balls (mobile holes) which might previously have reached the green (highest) level, cannot now negotiate such a gradient; the current i_1 then becomes zero. Conversely, the balls (minority carriers) of the green level, can always 'fall' on to the red level. The current i_2 has remained at its initial value. The increase in the gradient causes only the minority balls to pass from the higher level to the lower.

This reasoning also enables us to explain the behaviour of the majority carrier electrons in the *N* region. Fig. 22 shows the new situation; the widening of the gap between the two planes now allows only the passage of the balls (free electrons) from the higher to the lower plane. The current i_3 is zero; the electrons making up the current i_4 will then be the only ones appearing in the expression for the external current.

This analogy helps to illustrate the reverse current in the junction. It is, indeed, very difficult to explain the origin of this current by means of the majority and minority carrier density graphs.

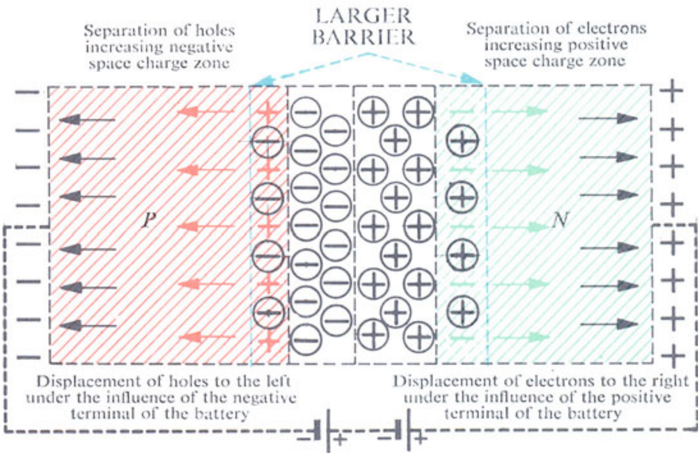
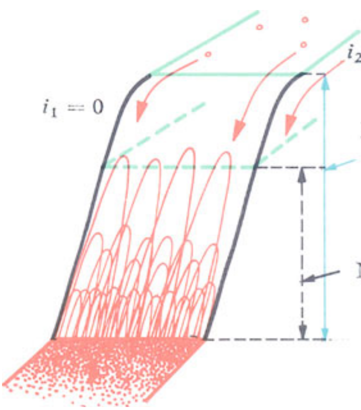


Fig. 20



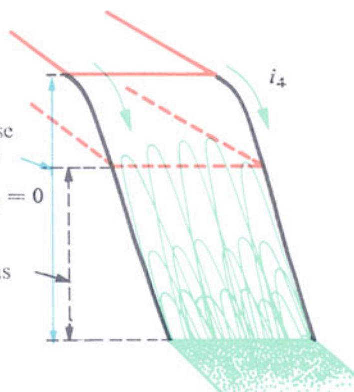


Fig. 21 **Fig. 22**

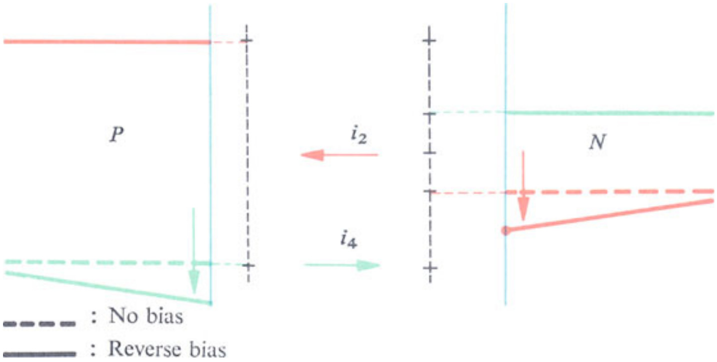


Fig. 23

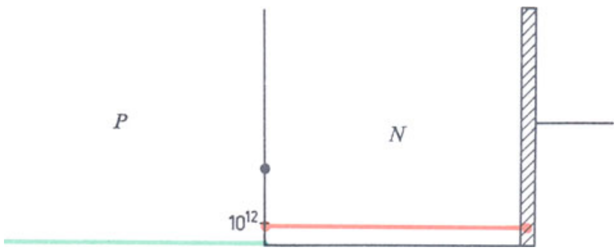


Fig. 24

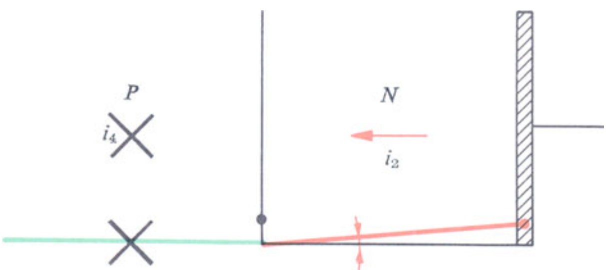


Fig. 25

Use of density graphs on a logarithmic scale. This condition may now be seen in the network of majority and minority carrier density graphs in Fig. 23. The absence of bias is represented by the broken lines. The corresponding graphs are perfectly horizontal.

The reverse bias causes a widening of the barrier, the new ends of which are represented by two vertical blue bands. The movement of the holes from the N region towards the P region produces a decrease in the hole density in the first region; consequently the corresponding graph is at a lower level than its equilibrium position. Similarly the movement of the electrons from the P region towards the N region produces a reduction in the electron density in the P region near the barrier. These two gradients of the minority carriers densities enable us to express the two currents i_2 and i_4 in the circuit.

Use of density graphs on a linear scale. Finally, the absence of bias implies a horizontal position for the two minority carrier density graphs (with the ordinates on a linear scale). A zero value of the angles, and, therefore, of the current, corresponds to this absence of bias (Fig. 24).

Biasing the junction in the reverse direction causes a displacement of the holes from the N region towards the P region with a corresponding decrease in density; as a result the level of the hole density graph will be near to zero (Fig. 25). In practice the hole density is not zero near the barrier, but the 10^7 or 10^8 mobile holes, which correspond to this new bias condition, become negligible by comparison with the 10^{12} holes initially present in this region.

In the same way the flow of the electrons from the P region towards the N region produces a decrease in the electron density near the barrier in this region; but it is now difficult to observe this decrease of density because of the scales used. Since the electron current i_4 is negligible, the external current (which is a function of the sum of the two currents) is exclusively dependent on the number of holes passing from the N towards the P region (current i_2). The current in the circuit is then a function of the tangent of the angle of the hole density graph in the N region.

This angle very rapidly approaches a limiting value with which the saturation current of the diode is associated.

1.4. Diodes in switching

Now that we have completed our study of the static behaviour of diodes, it will be interesting to describe the operation of a circuit including a diode (a junction associated with resistive elements in series and parallel) connected in series with a resistance R and a generator G (Fig. 26). This generator supplies a rectangular voltage. The drive voltage switches the diode from the absence of bias to a certain degree of bias in the forward direction (Fig. 27*a*). In order to describe the shape of the signal received at the output of this circuit, all that is necessary is to represent the variation with time of the voltage at the terminals of the load resistance (R). Since the characteristics of the load used are purely resistive, a simple application of Ohm's law enables us to associate the voltage at the terminals of this resistance directly with the shape and value of the current in the circuit.

A description of the output signal, therefore, requires precise knowledge of the variations of the current with time.

By this means, and by a simple use of networks of static characteristics, it should be possible to associate the variations of this current directly with the variations of the applied voltage (Fig. 27*b*). In fact, a purely static study of the diode is very inadequate for the explanation of the effects peculiar to the transient behaviour.

Switching in the direction of increasing current

In practice, the shape of the current is very different from that of the applied signal. The explanation of this effect will be facilitated by the study of the various bias conditions of the diode. We will first consider the behaviour of the diode from instant t_1 to instant t_3 (Fig. 28).

From instant t_1 to the instant immediately preceding t_2 (Fig. 28) there is no applied signal and the diode is not biased. Conversely, from instant t_2 to instant t_3 the voltage supplied by the generator biases the diode in the forward direction. A study of the shape of the current as a function of time is equivalent to defining the strength of the current on the one hand, from instant t_1 to instant t_2, and on the other hand from instant t_2 to instant t_3.

A reminder of the two static conditions of the operation of the diode will facilitate the determination of the limiting values of the current in the circuit. These values may be calculated either from the forward characteristic of the diode associated with that of the resistance, or by means of the minority carrier density graphs.

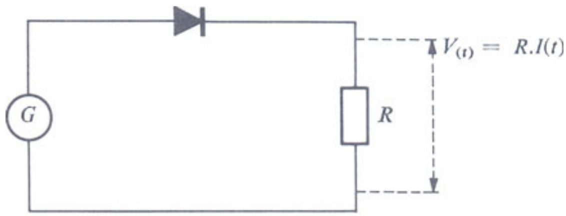


Fig. 26

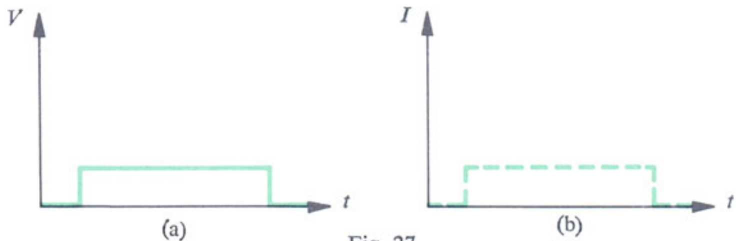


Fig. 27

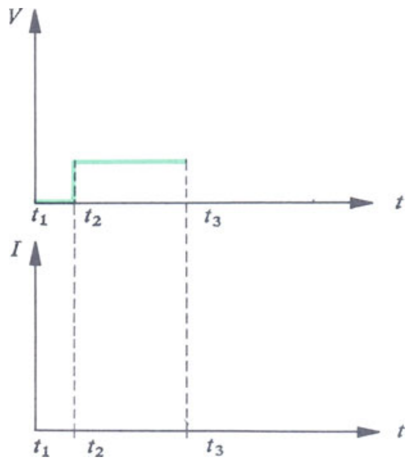


Fig. 28

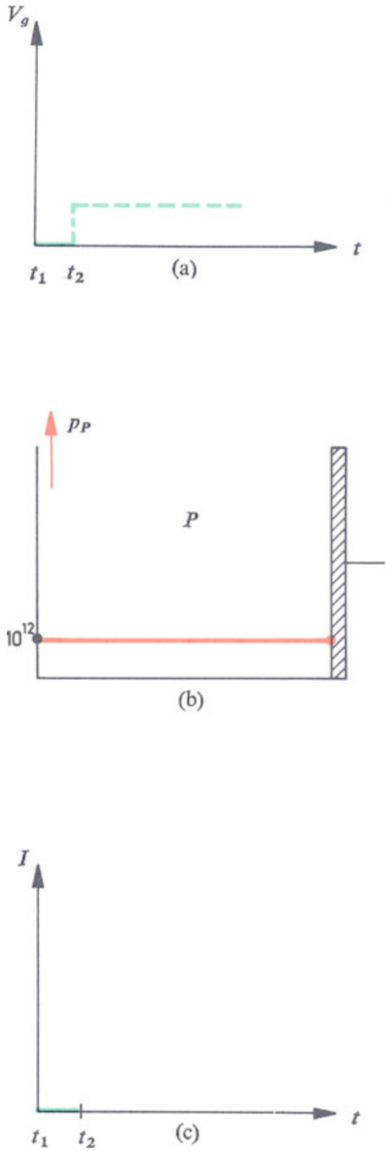


Fig. 29

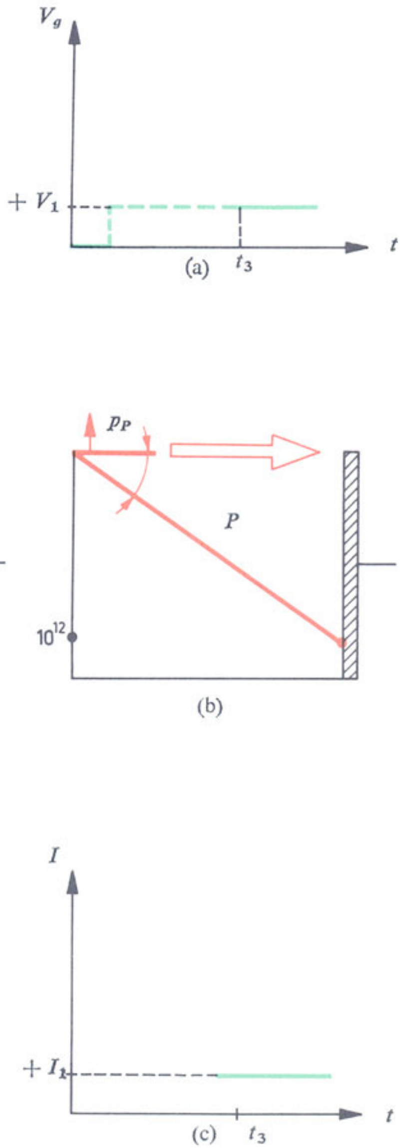


Fig. 30

Absence of bias. From instant t_1 to the instant preceding t_2 no signal is supplied by the generator (Fig. 29a). To this absence of bias corresponds a horizontal position of the density graph of the holes in the N region (Fig. 29b). The pole located to the right of this region is simply an ohmic contact assumed to be absolutely perfect. The hole density in the N region, near this pole, is therefore always equal to its equilibrium level whatever the bias conditions of the junction. The horizontal position of this graph implies a zero value for the angle and, therefore, for the current in the circuit. From instant t_1 to the instant preceding t_2 there is no current in the circuit (Fig. 29c).

Bias in the forward direction. At instant t_2 the signal is suddenly applied to the terminals of the junction. For the moment we will ignore the value of the current in the circuit at this instant.

Let us, therefore, assume an instant t_3 long enough away from instant t_2 to enable the diode to be permanently biased in the forward direction.

At instant t_3 a voltage V_1 is applied to the terminals of the diode and the resistance (Fig. 30a).

This voltage persists for a relatively long period; the generator may therefore be likened to a battery. The barrier being now narrowed, the hole density in the N region increases and the density graph is oblique as shown in Fig. 30b. This position corresponds to a certain value of the angle between the graph and the horizontal; the current of holes in the N region is a function of the tangent of this angle. In Fig. 30c, instant t_3 is associated with the flow of a current in the circuit equal to I_1.

These two static states are associated with a zero current in the circuit from instant t_1 to instant t_2, and with a current equal to I_1 from instant t_3 until the disappearance of the signal. An increase in the applied voltage takes place only in the second state; the absence of voltage is always equivalent to a zero value of current; conversely, an increase in the bias in the forward direction would result in a current higher than the value I_1.

The position of the hole density graph in the N region is above all a function of the value of the voltage applied at the terminals of the diode.

Very precise values of current in the circuit are dependent on the static states, these values being given by the extreme positions of the density graphs. The present problem is to determine the exact shape of the current during the two transient periods:

(a) first transient period, when the voltage rises from zero to its final value;

(b) second period, when the voltage falls from its final value to zero.

These two transient periods are represented on the x-axis by instants t_2 and t_4 (Fig. 31). In the following pages we will be examining the influence of the amplitude of the signal on the shape of the current in the circuit. First we must take account of another element: the resistance R.

Influence of the total resistance of the circuit. The determination of the resistance of a circuit is not limited to the examination of the resistance of the load, equivalent in our case to the load resistance; we must also consider the resistances of the various components making up the circuit as a whole. In Fig. 32, which represents the connection in series of a generator, a diode and a resistance, we see that there are in fact three resistances:

(a) the generator has an internal resistance (R_g); its value may vary greatly;

(b) the diode, in addition to its behaviour typical of a junction, also presents a certain resistance to the flow of the current; there will be an opportunity later to define more precisely the actual value of this resistance;

(c) the load resistance, which may be purely ohmic, but also that represented by the input junction of a transistor.

In order to determine the total resistance of a circuit we must associate these three resistances in the form of a single resistance marked 'R' in Fig. 33: this resistance is equal to the sum of the three others.

The shape of the current is strongly influenced by the value of this resistance.

In order to simplify our discussion, it would be useful to ignore this influence, assuming it to be very little different from zero. Our study of this particular case is purely theoretical, as it is impossible for a circuit containing a generator and a diode to have zero resistance.

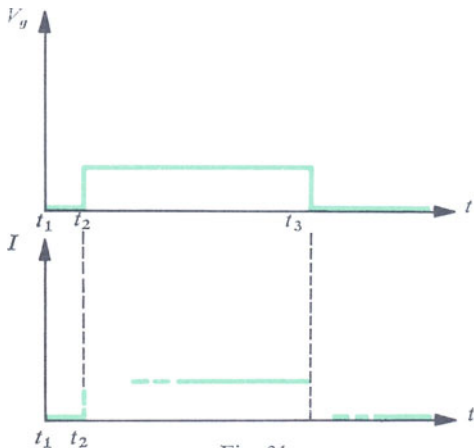


Fig. 31

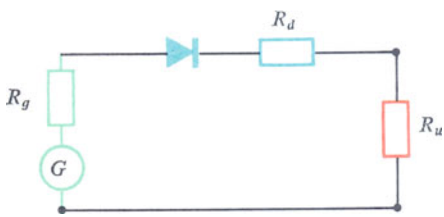


Fig 32

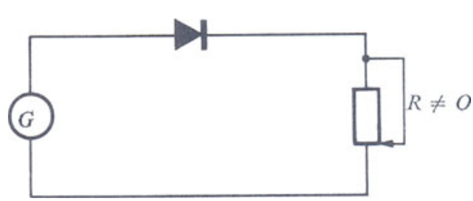
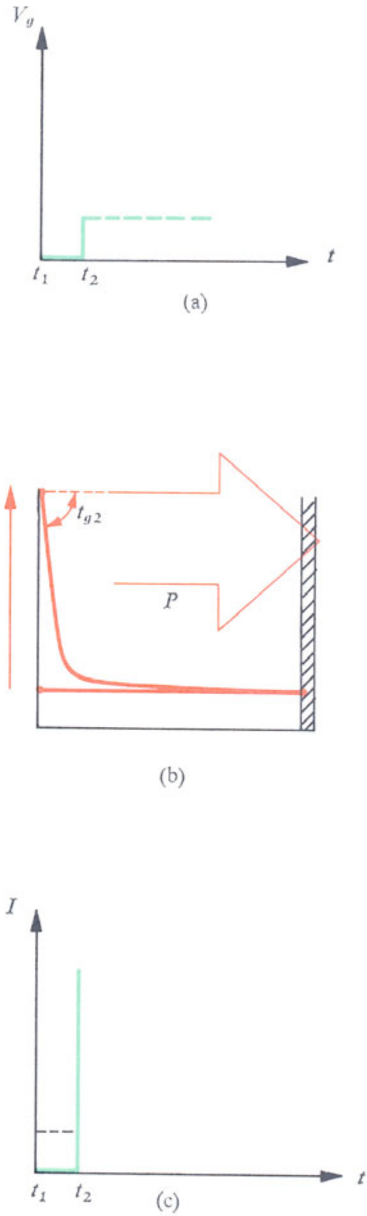
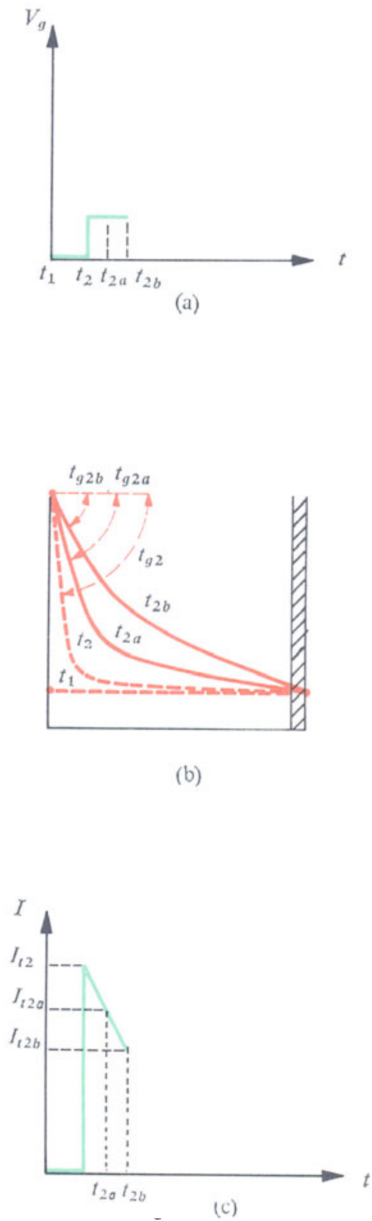


Fig. 33


V_g
t
t_1
t_2
(a)
t_{g2}
P
(b)
I
t
t_1
t_2
(c)


V_g
t
t_1
t_2
t_{2a}
t_{2b}
(a)
t_{g2b}
t_{g2a}
t_{g2}
t_{2b}
t_{2a}
t_2
t_1
(b)
I
I_{t2}
I_{t2a}
I_{t2b}
t
t_{2a}
t_{2b}
(c)

Influence of a sudden variation of voltage. At instant t_2 the voltage supplied by the generator changes abruptly from zero to its value V_1 (Fig. 34a).

The barrier immediately narrows and the holes pass from the P region to the N region. The total absence of resistance in the circuit makes possible a diffusion of holes corresponding to the effective change from the equilibrium density level to its final level. Consequently, the hole density graph in the N region is at once displaced from its equilibrium position (10^{12} level) to the position dependent on this bias condition (level 10^{13}).

There has not been time for the surplus holes in the left side of the N region to diffuse into its interior; it is only the extreme left side which is influenced by this increase in minority carrier density. The hole density graph then becomes oblique as shown in Fig. 34b.

The tangent of the corresponding angle gives the initial value of the hole diffusion current in the N region, or the strength of the current in the circuit. This angle is large (nearly 90°) and its tangent is consequently near infinity: the current corresponding to this value of the tangent is very high (Fig. 34c).

When the hole density graph has reached its final position near the barrier, in the N region, the holes which continue to diffuse can now only modify the density levels in the interior of this region.

From the instant t_2 to the instant t_{2a} (Fig. 35a), with the holes still penetrating into the interior of the N region, the graph is displaced from the extreme left position, represented in Fig. 35b by t_2, towards an intermediate position represented by t_{2a} in the same figure. The angle and its tangent decrease and the current is therefore smaller.

At instant t_{2b} the holes still pass into the N region, and the diagram swings again to the position represented by t_{2b}.

Since the tangent at instant t_2 is larger than that at instant t_{2a}, which is itself larger than the tangent at instant t_{2b}, the corresponding currents decrease proportionally, that is;

$$I_{t2} > I_{t2a} > I_{t2b}$$

From instant t_2 to instant t_{2b} the current varies as shown in Fig. 35c.

In fact a strict application of the theoretical clause concerning a zero value of resistance in the circuit would imply a current of infinite value at instant t_2. The choice of a value I_{t2} is necessary to simplify interpretation by the reader.

After this note concerning the restrictions brought to the representation of the instantaneous values of the current, we come to the definition of the last transient stages governing the increase of the current from its value I_{t2b} to its final value I_1.

Fig. 36 shows the variations of the voltage supplied by the generator at instants t_2, t_{2a}, t_{2b}, t_{2c}, t_{2d} and t_2'; the signal will be seen to maintain its initial value.

Instant t_{2b} represents a very exact position of the density graph in Fig. 37; this position gives us the value of the angle between the graph and the horizontal in the neighbourhood of the barrier and thence the tangent of this angle. It is then easy to calculate the value I_{t2b} of the current.

With the signal still applied, the diode remains in the same state; the holes continue to penetrate and diffuse into the N region. As the hole density in the neighbourhood of the barrier is stabilized at the level 10^{13} imposed by the applied voltage, the line becomes curved as shown by the broken lines in Fig. 37. At instant t_{2c} the diagram is near to its final position. The current has decreased again and is therefore near to I_1.

At instant t_{2d}, which is not shown in Fig. 37, the density graph is very near its final position and the current is still nearer to the value I_1.

At instant t_2', the graph is inclined as shown by the full line in Fig. 37 and the current is equal to I_1.

Fig. 38 shows the transient shape of the current when the bias of the junction changes suddenly from zero to a given forward value.

From instant t_1 to the instant preceding t_2, the current is zero in the circuit; at instant t_2, the current assumes a value near to infinity (I_{t2}); it then decreases until it reaches its final value (I_1) at instant t_2'.

The current is then maintained at this value during the whole period that the voltage is applied at the terminals of the circuit.

This strong forward current pulse associated with the first stage of the transient behaviour of a junction is of special importance in the use of diodes in switching circuits.

It will, however, be strongly influenced by:

(a) the value of the total resistance of the circuit;

(b) the amplitude of the applied signal.

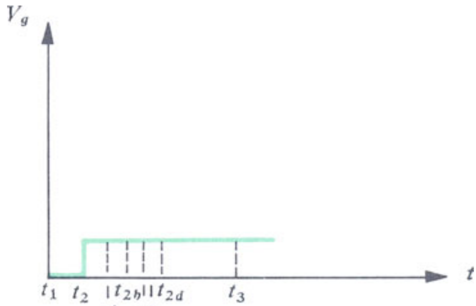


Fig 36

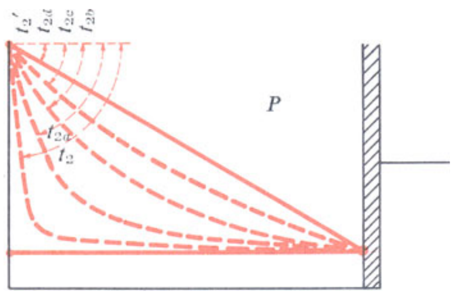


Fig. 37

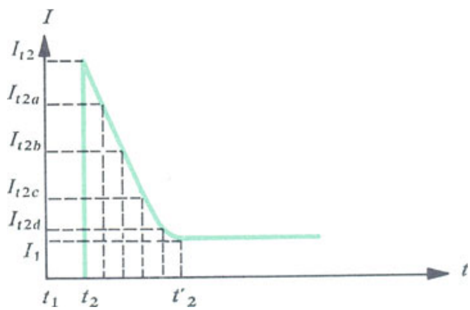


Fig. 38

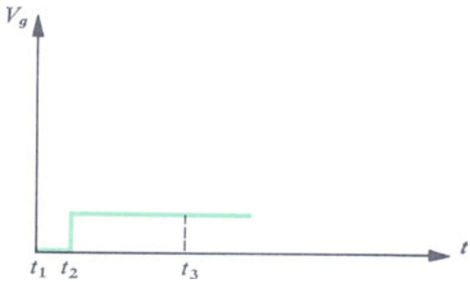


Fig. 39

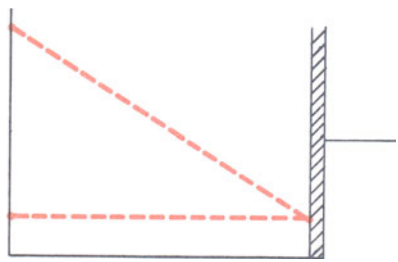


Fig. 40

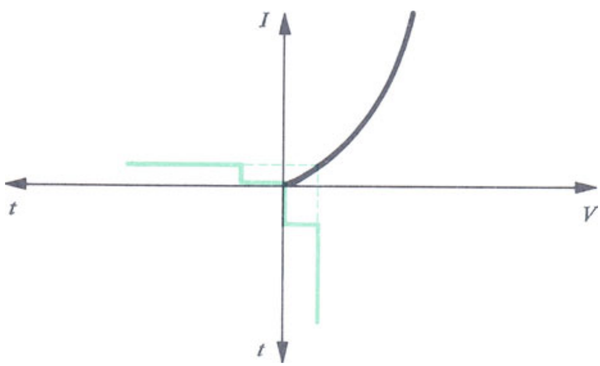


Fig. 41

Note on the use of the networks with static characteristics. When a rectangular voltage is applied to the terminals of a junction, the current in the circuit is very different from the applied signal. Fig. 39 shows the voltage variations with time. These variations give the diode the following stages:

(a) in the absence of bias from instant t_1 to the instant preceding t_2;

(b) with a given forward bias, from instant t_2 to instant t_3.

From instant t_1 to the instant preceding t_2, the graph is horizontal corresponding to the equilibrium hole density in the N region. A zero value of current should be associated with this position of the graph.

Conversely, at instant t_3, the transient effects are assumed to have 'disappeared', and the diagram is inclined (Fig. 40); this position determines the final value of the current in the circuit.

The study of the static behaviour of circuits containing semiconductor devices (diodes or transistors) is facilitated by the use of their networks of characteristic curves. At audio frequencies these characteristics are in general adequate for the determination not only of the effective values of the currents or voltages appearing at the different points of the circuit, but also of the precise shape of the signals.

A similar reasoning for rectangular signals would give us, for the type of circuit used, the network of characteristic curves in Fig. 41. The top-right quadrant represents the forward characteristics of the diode; the bottom-right quadrant represents the variations in the applied voltage, and the top-left quadrant the variations with time of the current in the circuit.

In this figure the absence of bias results in a zero value for the current, and the corresponding steps are superimposed on the two abscissae.

On the other hand, biasing in the forward direction (or a certain positive voltage in the bottom-right quadrant) causes a current to flow in the circuit, the value of which is obtained by successive projections of the voltage on the forward characteristic and on the ordinate.

This simplified reasoning helps to explain the absence of a perfectly rectangular current, and it justifies us in ignoring the transient effects.

Definition of the forward recovery time of the diode. The density graphs have enabled us to point out the striking differences in shape between the signal applied at the terminals of the circuit and the corresponding current. A perfectly rectangular generator voltage is accompanied during the first transient stage by a very strong current pulse in the circuit.

Without the aid of the graphs it might have been thought that if the current were not perfectly rectangular, it would require a certain time to reach its final value. The angle between the graph and the horizontal in the neighbourhood of the ohmic contact is a function of the time required for the production of the holes essential for the establishment of the current.

In practice the very high values of the angle between the hole density graph in the N region and the horizontal in the neighbourhood of the barrier, produce very intense diffusions, the continuity of which is ensured by a displacement of a similarly large number of electrons in the external circuit.

This time during which the diode 'no longer reacts' to the changes in the applied voltage is defined in data sheets by the term: 'forward recovery time of the diode' (Fig. 43).

This recovery time is very important, not only because of its excess of current, but also because during this time the variations in the signal supplied by the generator in either direction no longer have any effect on the value of the current.

When using a diode of this type, either in rapid switching circuits for example or in signals with a relatively high repetition rate, the recovery time should be less than the time separating the two 'leading edges' of the pulses which constitute the information. The current variations shown in Fig. 43 apply to a purely theoretical case. There will be an opportunity later to define the less 'sharp' pulses accompanied by longer forward recovery times (the influence of the resistance of the circuit).

Simplified diagram of the diode. In addition to the characteristics and the recovery times supplied by the manufacturer, there is another very important item of information: the electrical equivalent circuit of the diode.

The various static studies enabled us earlier to draw attention to the exclusively resistive elements. The appearance of such a current pulse in the circuit is the consequence of the capacitive behaviour of the diode (Fig. 44).

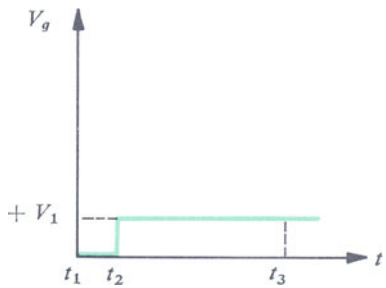


Fig. 42

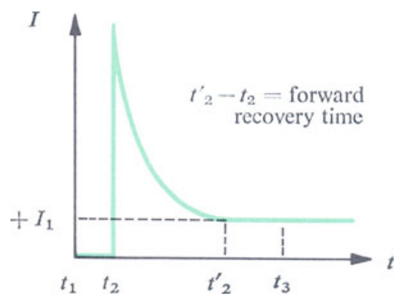


Fig. 43

Fig. 44

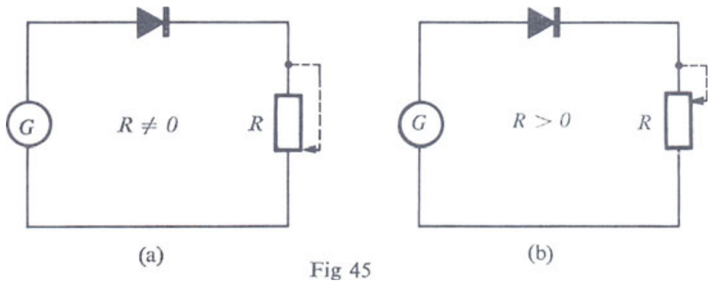


Fig 45

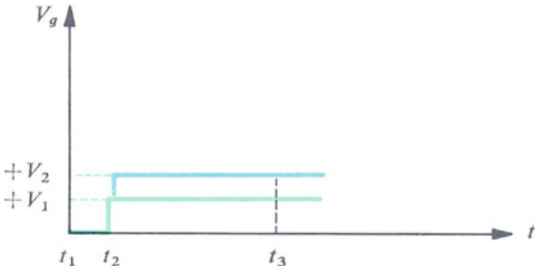


Fig. 46

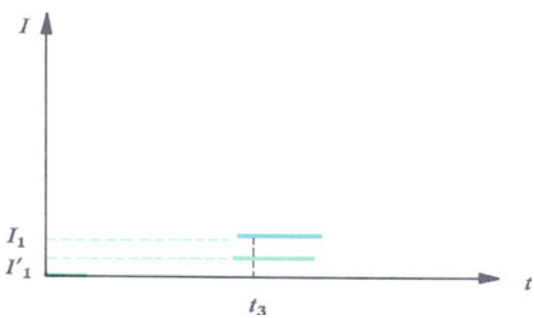


Fig. 47

Influence of a resistance in the circuit. The circuit in Fig. 45a, namely a generator associated in series with a diode and a short-circuited potentiometer, illustrates the limiting theoretical conditions peculiar to the study discussed above.

In practice a resistance near to zero is unthinkable as each of the components of this circuit offers a certain resistance to the flow of the current, namely:

(a) the resistance of the generator;

(b) the resistance of the diode;

(c) the load resistance.

It is therefore necessary, if we wish to gain a more positive idea of the real behaviour of the diode when a rectangular signal is applied to its terminals, to know the influence of this combination of resistances.

The circuit shown in Fig. 45b differs from the preceding one by the displacement upwards of the contact of the potentiometer. We must ask ourselves how this resistance affects the shape of the current.

Before passing to a detailed examination of the effects peculiar to the behaviour of the diode in the transient state it is advisable, in order not to make a quantitative distinction between the above study and present-day developments, to define again the static operating conditions of the circuit.

Figure 46 shows the two first stages of the rectangular signal supplied by the generator. From instant t_1 to the instant preceding t_2, the voltage is zero. On the other hand, from instant t_2 to instant t_3 the voltage remains constant at a value V_2 (shown in blue in the figure). The horizontal green line, which corresponds to the voltage V_1, applies to a circuit with a theoretical zero resistance.

The static study therefore consists of the determination of the current during the two stable states shown in this figure, namely:

(a) from instant t_1 to instant t_2;

(b) at instant t_3.

From instant t_1 to the instant preceding t_2, there is no applied voltage and the current in the circuit is therefore zero (Fig. 47).

At instant t_3 a voltage V_2 is applied to the terminals of the circuit (Fig. 46). A current of value I_1 corresponds to this voltage.

If we had maintained the voltage at the value V_1, the current, due to the resistance connected in series in the circuit, would have been lower and equal to I_1'.

We must ask ourselves why for a diode and resistance in series, we chose an applied voltage V_2 higher than V_1?

Since there is no resistance in the circuit we must use the forward characteristic of the diode (curve shown in green in Fig. 48).

A voltage V_1 supplied by the generator is associated with a certain value of the current, obtained successively by the setting up of the perpendicular to the abscissa at the point V_1 and by the projection on the ordinate of the point of intersection of this perpendicular with the characteristic.

The connection of a resistance in series with the diode causes the characteristic to be displaced to the right (blue curve in the figure). This new characteristic is obtained by the point by point summation of the resistance characteristic and the diode characteristic.

A forward voltage bias of V_1 only produces in the circuit a current I_1' smaller than I_1; if it is desired to maintain the current at its initial value I_1, all we have to do is to increase the applied voltage and to locate it at the position V_2 on the abscissa of the network of characteristics. Any increase in the value of the resistance connected in series in the circuit would cause a displacement of the characteristic to the right and would require a higher applied voltage to maintain the final current at a constant value.

This graphical explanation may easily be confirmed by an examination of the circuit, in which we see at once that any drop in voltage caused by the flow of a current in the resistance R can be observed by a decrease in the voltage applied at the terminals of the diode.

Study of the static behaviour. Fig. 49 shows the new variations of the voltage as a function of time. From instant t_1 to the instant preceding t_2, this voltage is zero; the hole density graph in the N region is horizontal (Fig. 50). The angle is zero and the current is also zero (Fig. 51).

At instant t_3 the voltage has been applied for a relatively long time, and this enables us to liken the generator to a battery supplying a voltage V_2.

The hole density graph in the N region has taken up its final position (Fig. 50); the tangent of the angle between this graph and the horizontal enables us to calculate the current in the circuit (Fig. 51).

The static behaviour of a diode is not necessarily affected by the value of the resistance placed in series with it, and in order to maintain the same current we need only modify the voltage supplied by the generator.

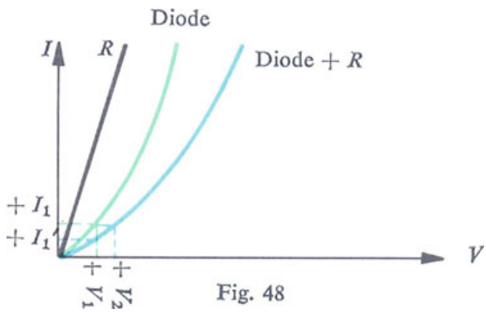


Fig. 48

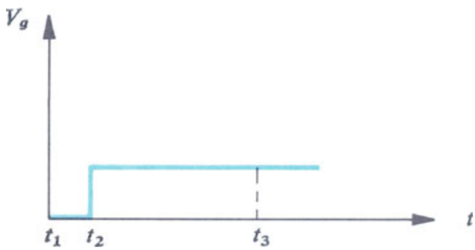


Fig. 49

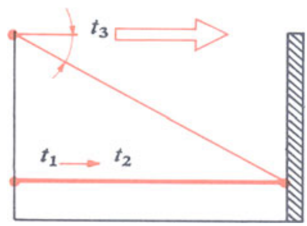


Fig. 50

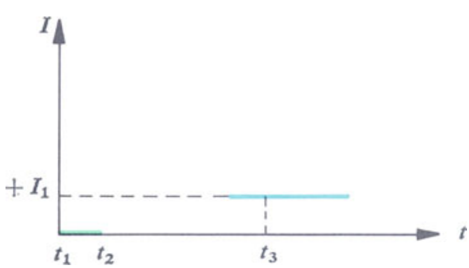
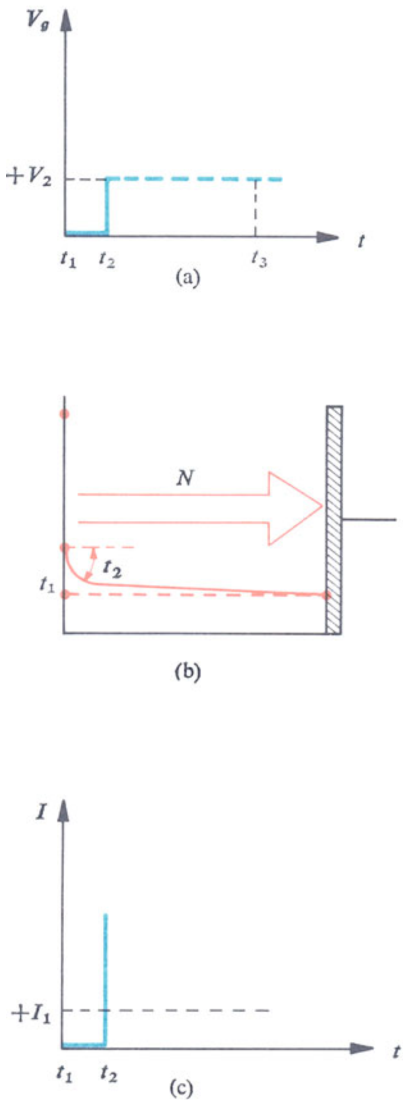
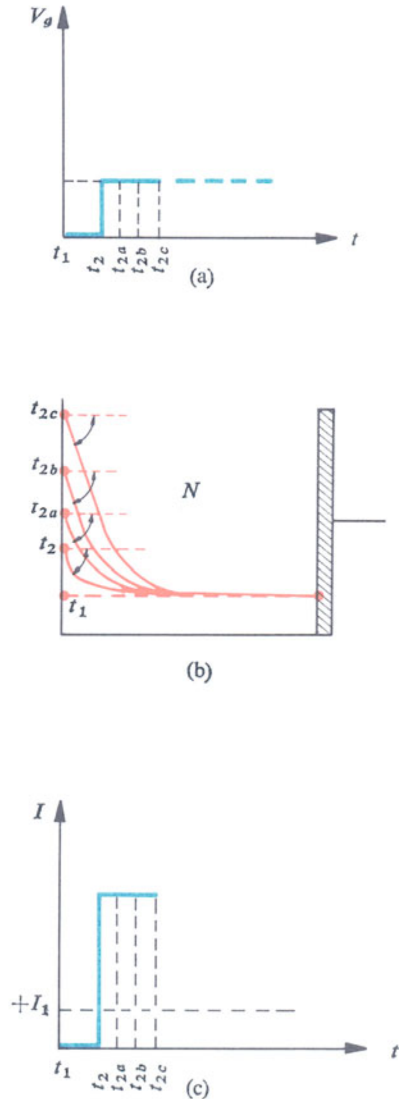


Fig. 51



V_g
$+V_2$
t
t_1
t_2
t_3
(a)
N
t_2
t_1
(b)
I
$+I_1$
t
t_1
t_2
(c)

Fig. 52



V_g
t
t_1
t_2
t_{2a}
t_{2b}
t_{2c}
(a)
t_{2c}
t_{2b}
t_{2a}
t_2
t_1
N
(b)
I
$+I_1$
t
t_1
t_2
t_{2a}
t_{2b}
t_{2c}
(c)

Fig. 53

Study of the dynamic behaviour. What happens at instant t_2 when the voltage supplied by the generator increases suddenly from zero to V_2?

The barrier narrows immediately and the holes diffuse from the *P* region towards the *N* region. In the absence of resistance all the holes needed for the change of the density near the barrier from its equilibrium level to its final value can pass through the barrier. The insertion of a resistance in the circuit entails a limitation in the number of holes.

Consequently at instant t_2 the density graph is only in an intermediate position. This position, shown in Fig. 52b, is associated with a new angle and a different value of current in the circuit. The reduction of this angle involves a decrease in the tangent; the current pulse is then limited to a value lower than in the former case (Fig. 52c).

The applied voltage remains at the value V_2 until instant t_4. The holes continue to diffuse from the *P* region towards the *N* region and will, whilst penetrating into this region, displace the density level near the barrier until it reaches its final value corresponding to the bias chosen.

If we examine the different stages necessary for the displacement of the density graph from its position at instant t_2 to that occupied at instant t_{2a}, we notice that it occupies successively the positions associated with instants t_{2a} and t_{2b} shown on the abscissa in Fig. 53a.

The increase of the density level in the neighbourhood of the barrier is accompanied by a continual penetration of holes into the interior of the *N* region; consequently the density levels increase simultaneously near the barrier and in the left hand part of the *N* region. We may therefore assume that at instant t_{2a} the angle between the new position of the graph and the horizontal is practically equal to the previous one (Fig. 53b).

If at instant t_{2a} the value of the angle of the hole density graph in the *N* region has not changed the current at this instant has maintained its initial value.

As the angles associated with instants t_{2b} and t_{2c} are also equal to the angles associated with instants t_{2a} and t_2, the current remains at this value from instant t_2 to t_{2c} (Fig. 53c).

The insertion of a resistance causes two important modifications in the shape of the current in the first part of the transient stage;

(a) a decrease in the peak of the forward current;

(b) an increase in the period during which this current maintains its maximum value.

At instant t_{2c} the density level reaches its final value in the neighbourhood of the barrier. However, the signal is still applied at the terminals of the circuit (Fig. 54). What happens at instants t_{2d}, t_{2e} and t_2''?

Starting from instant t_{2c}, since the hole density level in the N region cannot increase near the barrier, the holes continue to diffuse into this region, and will gradually swing the diagram to its final position.

At instant t_{2d} (Fig. 55), the hole density graph has moved upwards; the holes coming from the P region have now reached the zones more to the right in the interior of the N region. The angle between this new graph and the horizontal is now smaller, the tangent has consequently decreased and the current assumes a value lower than in the previous case.

At instant t_{2e}, since the holes continue to penetrate into the interior of the N region, the angle between the graph and the horizontal is still smaller, and the current associated with this angle is therefore lower than the value at instant t_{2d}.

Finally, at instant t_2'', tha graph assumes its final position; the angle then reaches its final value and the current reaches I_1, as shown in Fig. 56.

When a rectangular signal is applied to the terminals of a circuit consisting of a diode associated with a resistance, the increase of the current from zero to its final value I_1 is divided into three successive stages:

(a) At instant t_2 a strong forward current pulse, much higher than the value I_1, appears in the circuit;

(b) from instant t_2 to instant t_{2c} the current maintains this value; the corresponding time being required for the displacement of the density diagram from position t_2 to position t_{2c} in Fig. 55;

(c) from instant t_{2c} to instant t_2'', a period dependent on the approach of the density diagram to its final position, the angle and the current decrease with time.

Fuller explanations could be made of the quantitative conditions; the equations for the calculation of the forward recovery time associated with any given value of the resistance R depend in principle on the three previous stages.

In addition to the physical parameter the circuit resistance affects:

(a) the value of the current peak, inversely;

(b) the duration of the recovery time, directly.

Finally, the study of the behaviour of a diode in terms of a rectangular

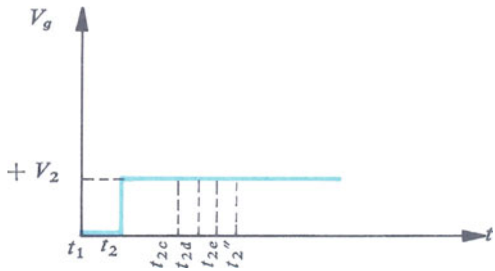


Fig. 54

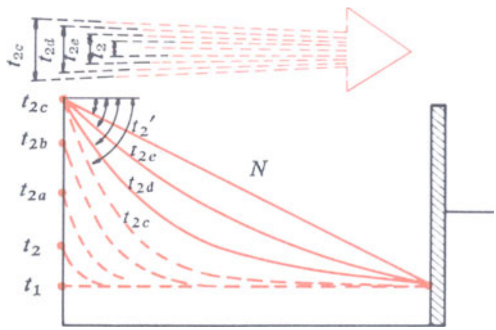


Fig. 55

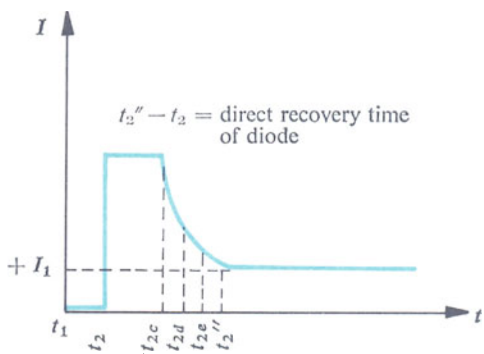
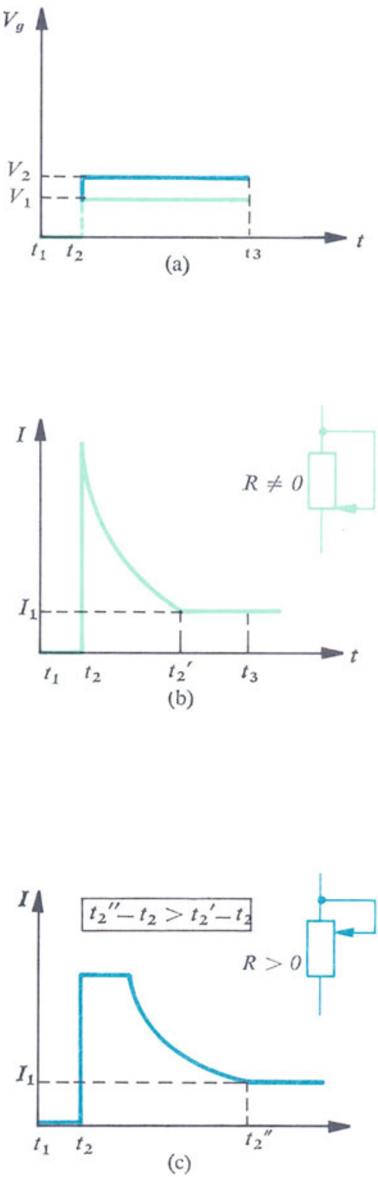
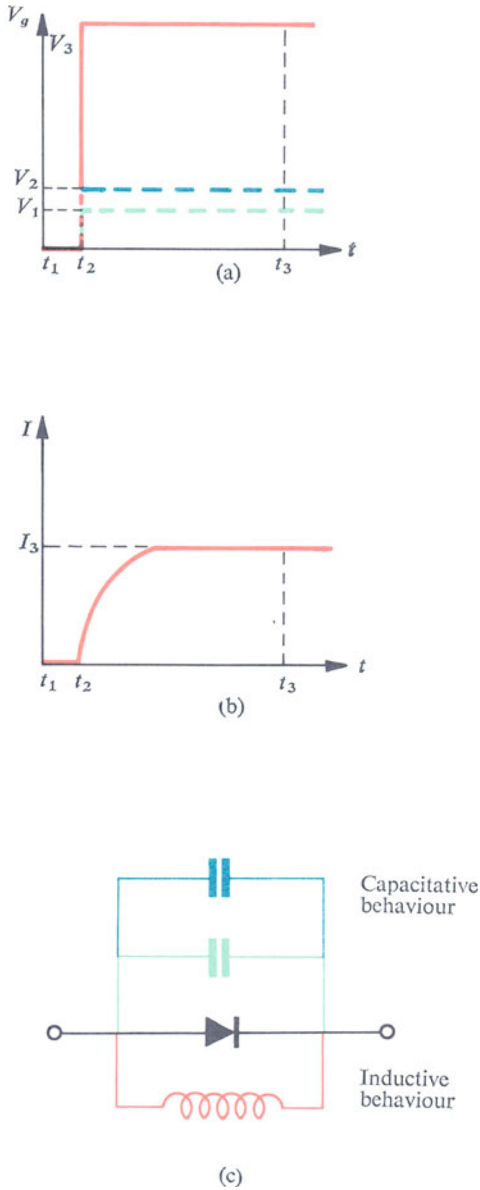


Fig. 56



V_g
V_2
V_1
t_1 t_2 t_3 t
(a)
I
$R \neq 0$
I_1
t_1 t_2 t_2' t_3 t
(b)
I
$t_2'' - t_2 > t_2' - t_2$
$R > 0$
I_1
t_1 t_2 t_2''
(c)

Fig. 57



V_g
V_3
V_2
V_1
t_1 t_2 t_3 t
(a)
I
I_3
t_1 t_2 t_3 t
(b)
Capacitative behaviour
Inductive behaviour
(c)

Fig. 58

signal applied at its terminals is divided into three sections:

(a) in the absence of a resistance in the circuit (purely theoretical case) Fig. 57b;

(b) with a resistance in the circuit (practical case) Fig. 57c.

When there is no resistance, a very strong current pulse appears at instant t_2; the current then falls rapidly until it reaches the value I_1. The period $t_2'-t_2$ shown in Fig. 57b, determines the minimum value of the forward recovery time of the diode. The value I_1 of the current is quantitatively dependent on the application of the voltage V_1 at the terminals of the diode (Fig. 57a).

When there is a resistance connected in series in the circuit the voltage supplied by the generator, for the final current to be the same, should be equal to V_2 (Fig. 57a). The current pulse appearing at instant t_2 is smaller (Fig. 57c); this current is maintained, on the other hand, at this value for a certain period and then falls much more slowly than in the previous example. The time required for the current to pass from zero to the value I_1 is given for a resistive load by $t_2''-t_2$ in Fig. 57c; this time is longer than before:

$$t_2''-t_2 > t_2'-t_2$$

Any increase in the resistance connected in series in the circuit causes an increase in the forward recovery time of the diode.

Influence of the amplitude of the signal

The signals supplied by the generator are of small amplitude; we will examine the effect of increasing the applied voltage.

In Fig. 58a, the voltage passes suddenly at the instant t_2 from zero to the value V_3 which is much larger than V_1 or V_2

The 'running away' effect noticed before no longer occurs; on the contrary the current rises exponentially as a function of time (Fig. 58b).

The diode no longer behaves in the small signal condition; this change in operation has an indirect effect on the electrical equivalent diagram of the diode.

The capacitance, the main item in the dynamic equivalent circuit, associated with the concept of forward recovery, is replaced by an inductance electrically dependent on the rise time of the current in the circuit.

Before explaining the reasons behind the origin of this inductive behaviour, it would be useful to remind ourselves of some physical properties of the P and N regions of which the diode is composed.

The diode consists essentially of a junction, defined as the association of a monocrystal with two types of conductivity: the one of the *P* type on the left of Fig. 59, and the other of the *N* type. The two conductive regions are separated by a barrier in which electric fields accelerate the passage of the carriers.

Electrical neutrality in the P region. Is it possible to define the state of the electrical charge in the two regions, *P* and *N*?

The *P* region on the left of the junction contains a greater number of holes (majority carriers) than electrons, the minority carriers. The surplus of holes is explained by the injection of impurity atoms, indium atoms which have accepted a fourth electron and have become negatively ionized (Fig. 60). Can we assume, given the surplus of holes (positive charges) over electrons (negative charges), that the *P* region is positive?

This is definitely not so, since the number of holes is directly dependent on the number of ionized atoms.

Indeed, whatever part of the *P* region is considered, the number and charges of the majority carrier holes in this region are exactly compensated by the negatively ionized impurity atoms; this region of the monocrystal is therefore electrically neutral.

Electrical neutrality of the N region. We will now consider the *N* region. In Fig. 61 the electrons, the majority carriers, are a hundred times more numerous than the holes, the minority carriers.

The surplus of negative charges should therefore be associated with the concept of a negative *N* region. In practice, the larger number of electrons is explained by the injection of pentavalent impurity atoms (for example arsenic atoms). These atoms have lost their fifth valency electron and have become positively ionized. The number of free electrons is therefore a function of the number of positively ionized atoms; the *N* region is also electrically neutral.

Because of the importance of the study of the behaviour of the *N* region as a function of the signals applied at the terminals of the junction, it will always be necessary to bear in mind its electrical neutrality.

This is independent of:

(a) the bias;

(b) the position of the majority and minority carrier density graphs.

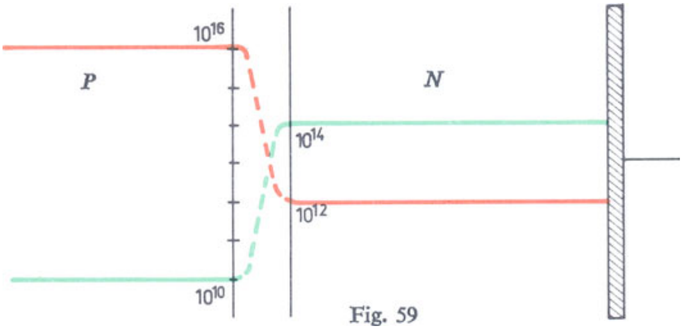


Fig. 59

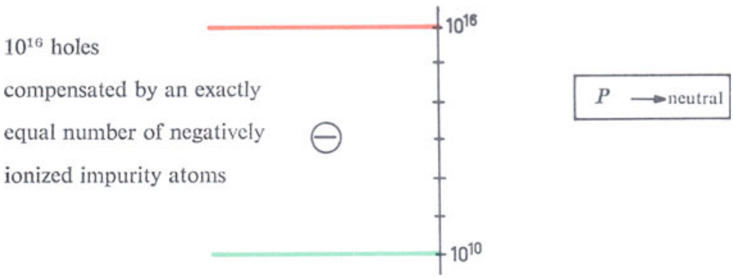


Fig. 60

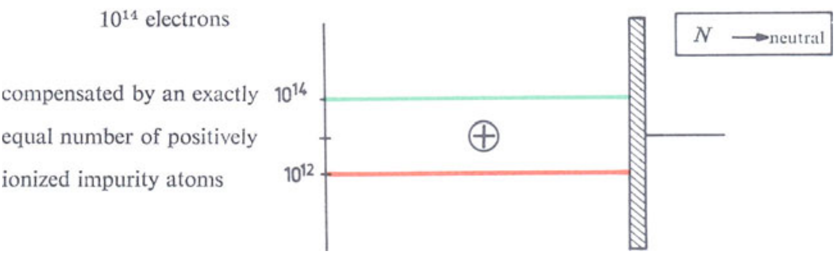


Fig. 61

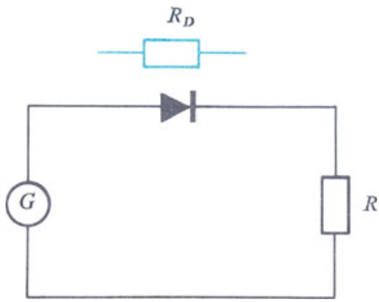


Fig. 62

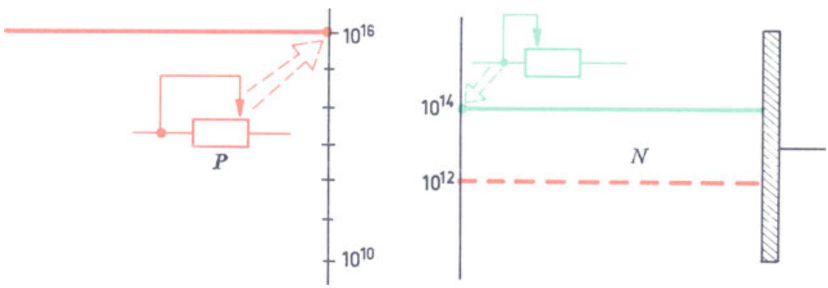


Fig. 63

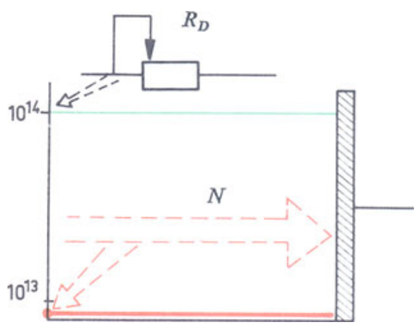


Fig. 64

Resistance of a diode

In the course of the above discussions we referred to the presence of a resistance intrinsic to the diode. This is represented by R_D in Fig. 62.

Can we explain the 'justification' of this resistance and especially of its origin in the interior of the *PN* junction?

The resistance of a substance depends on the number of mobile carriers it contains per cubic centimetre. If we examine a block of *P* type germanium or silicon and a block of *N* type germanium or silicon, we can express the value of their resistance in terms of their mobile carrier densities. In this case it is only the majority carriers that are of interest, because of their large number compared with the minority carriers.

If we look at Fig. 63 we see that the *P* region contains 10^{16} holes per cubic centimetre. The instrinsic resistance of this substance, represented in red in the figure, is therefore directly proportional to the number of majority carriers belonging to this region. Since this number is relatively high, the resistance is low. We will represent it by a short-circuited potentiometer.

In the *N* region the number of majority carriers is much lower. Modern diodes and *PNP* transistors are characterized by an *N* region more weakly doped than the *P* region. If the *N* region contains only 10^{14} free electrons the resistance offered to the current is higher. It is represented by a potentiometer, marked in green in Fig. 63, and its slider is at the left.

Any resistance effect due to the barrier can be ignored: the very strong electric fields present in its interior accelerate the passage of carriers into it.

The determination of the resistance of a diode consists in fact, in the present case, in the calculation of the resistance of the *N* region. The resistance R_D is therefore necessarily dependent on the value of the resistance of the *N* region (Fig. 64).

This resistance is directly proportional to the position of the density graph of the majority carriers in the interior of the *N* region. The corresponding densities are located at the 10^{14} level for majority carriers and at 10^{12} for the minority carriers. Any variation in the majority carrier density therefore influences the value of this resistance.

In certain types of diodes, the percentages of impurity injected into the *P* and *N* regions are different from those so far chosen. Since the difference between the two regions remains practically the same, the reasoning required for the explanation of the operation does not change.

Influence of forward bias on the neutrality of the N region

In the absence of bias (Fig. 65), the two majority and minority carriers' density graphs are horizontal. Fig. 66 shows the two densities on a graph with the ordinates graduated linearly. (As a reminder, the positive ionization of impurity atoms explains the neutrality of this region).

What happens when the applied signal changes suddenly from zero to V_1?

At the instant when the voltage is applied to the terminals of the junction (Fig. 67), and in the case of a very low resistance connected in series in the circuit, the hole density graph in the N region changes almost at once from the position shown by a broken line to that with a full line in Fig. 68.

This new position of the density graph of the minority carriers in the N region is associated with a surplus of holes in the interior of this region near the barrier. The first result of this increase in the number of holes is the diffusion of carriers from the point of highest density to that of the lowest, namely from the left to the right. The diffusion current is a function of the tangent of the angle between this graph and the horizontal.

However, the surplus of holes, positive charges, is no longer compensated by a corresponding number of negative charges in the N region; it is the cause of a positively charged zone at the left hand end. (Fig. 69).

The electrons are now attracted towards the left and there is no opposition to their concentration in a position exactly parallel to that of the holes. A current of electrons, caused by this electric field, appears flowing from right to left in the interior of the N region.

The surplus of electrons in the left hand section of the N region explains the new position of the density graph in Fig. 70. This decreasing density of free electrons from left to right is accompanied by a diffusion of electrons from the point of highest density towards that of lowest density.

Three currents appear simultaneously in the N region:

(a) a diffusion current of holes, flowing from left to right;

(b) an electron field current flowing from right to left;

(c) a diffusion current of electrons, flowing from left to right.

The values of the two electron currents (field and diffusion) are very similar. Consequently it is only the hole current that enters into the expression of the external current.

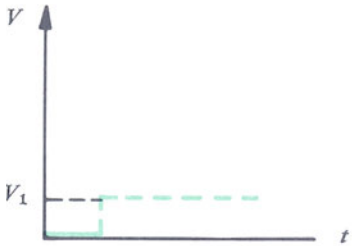


Fig. 65

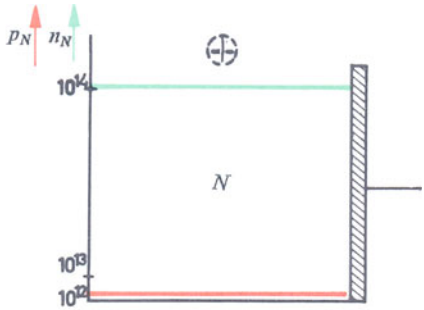


Fig. 66

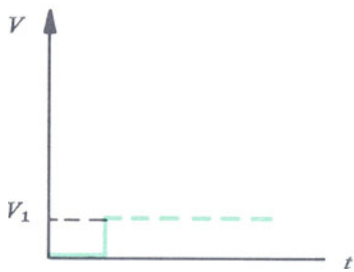


Fig. 67

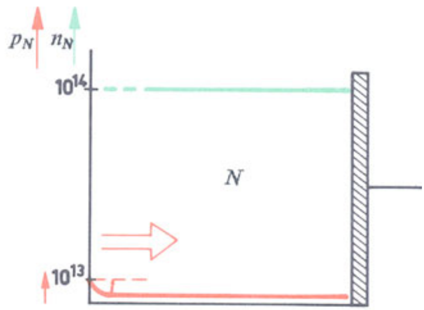


Fig. 68

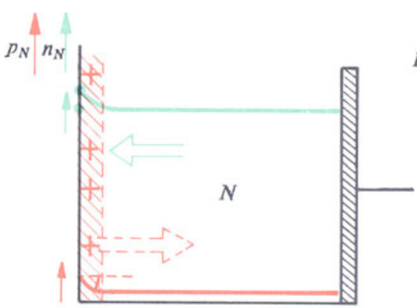


Fig. 69

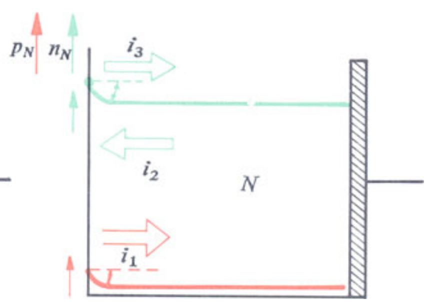


Fig. 70

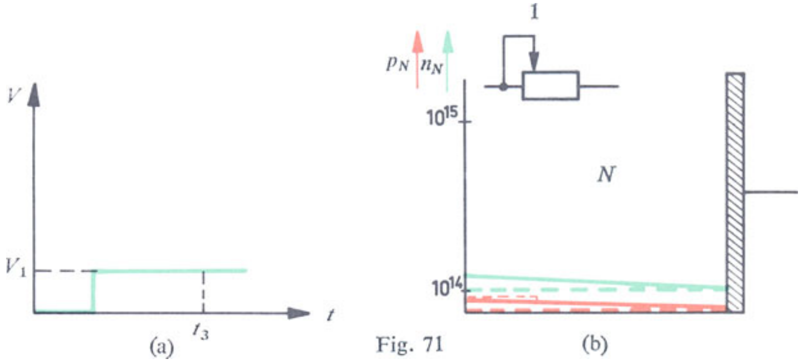


Fig. 71

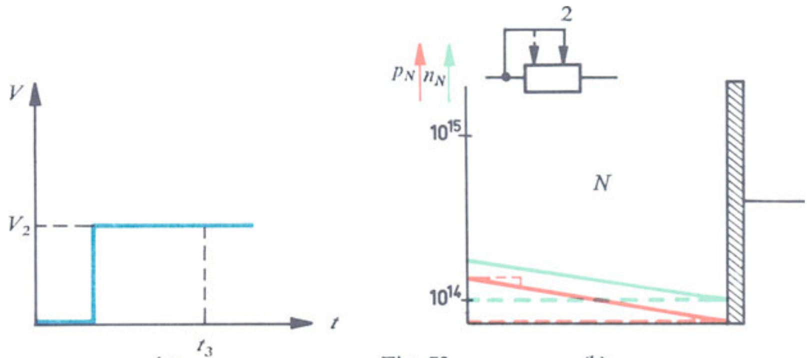


Fig. 72

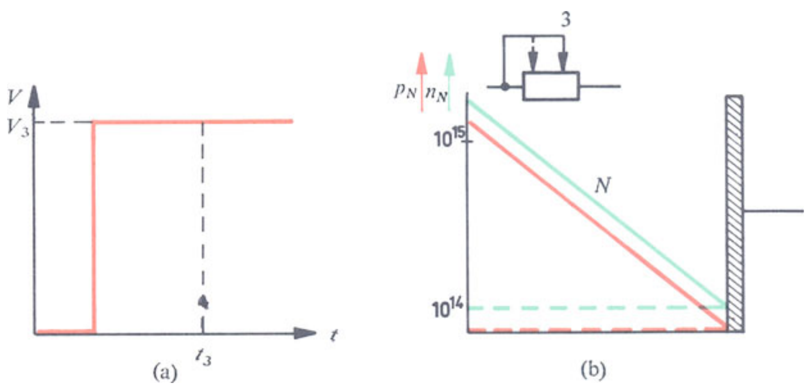


Fig. 73

Variations of the resistance of a diode as a function of the applied signal

Any displacement of the density graph of the holes in the N region results in an identical displacement of the electron density graph in the same region.

With small signals the voltage supplied by the generator (Fig. 71a) causes only a very small displacement of the hole density graph with respect to that of the equilibrium position (Fig. 71b). The current in the external circuit is accounted for by the variations in the tangent of the angle between this graph and the horizontal. The electron density graph, when plotted with the ordinate graduated linearly, is exactly parallel to the hole density graph.

The N type region is, therefore, always neutral in spite of the forward bias of the junction. The displacement upwards of the two graphs has practically no influence on the number of carriers in this region. Since the resistance of the N type material has not varied, the slider of the potentiometer is always located at the left hand end (Fig. 71b).

Let us now assume that the signal has a larger amplitude, namely voltage V_2 in Fig. 72a; the positions of the density graphs of the majority and minority carriers in the N region are at a higher level (Fig. 72b). A larger forward bias accentuates the narrowing of the barrier and speeds up the movement of the holes from the P to the N region. The increase in the number of majority and minority carriers (with the two graphs still remaining parallel) causes a displacement of the slider of the potentiometer to the right. The resistance of the N type material decreases.

Let us now assume a voltage V_3 much higher than V_2 (Fig. 73a). The final position of the hole density graph is now much higher than that of the electron density graph in its equilibrium position. In order to prevent the appearance of any field in the interior of the N region, the electron density graph assumes a parallel position (Fig. 73b).

In this region the number of carriers per cubic centimetre has considerably increased; the increase in the number of electrons and holes in these ratios causes a considerable decrease in the resistance of the N type material and consequently in the resistance of the diode.

This decrease in resistance is associated with a very large movement to the right of the slider of the potentiometer (Fig. 73b).

Rise time of the current in the circuit

We must now examine the reasons why the behaviour of a diode driven by a signal of a very large amplitude changes until it can be electrically likened to an inductance.

In the absence of a signal (Fig. 74a) the width of the barrier is at its equilibrium value; the number of holes diffusing from the *P* region to the *N* region is exactly equal to the number of holes 'falling back' from the *P* to the other region.

The two majority and minority density graphs are in their equilibrium positions. These two graphs are horizontal (Fig. 74b) and the small number of majority carriers in each cubic centimetre of the *N* region causes it to present a high resistance to the passage of the current (left hand position of the slider of the potentiometer).

Since the graphs are horizontal the corresponding angle is equal to zero, and the current in the circuit is therefore zero. From instant t_1 to the instant preceding t_2 (Fig. 74c) no current flows in the circuit.

At instant t_3 the voltage V_3 is still applied at the terminals of the circuit. This voltage is very high and it causes a considerable narrowing of the barrier. A large number of holes penetrates into the *P* region and the level of the density diagram is then as shown in the full line in Fig. 75b.

The hole density is now much higher than the electron equilibrium density. As will be seen from the above explanations the electron density will be exactly parallel to and greater than the hole density. The very large increase in the electron and hole densities in the *N* region causes a definite improvement in the conductivity of this material.

The resistance of the *N* type material (in fact the real resistance of the diode) may now be represented by the same potentiometer with the slider far to the right. A current equal to I_3 is associated with this voltage (Fig. 75c).

A new increase in the applied voltage will be manifested by an upward displacement of the density graphs of the majority and minority carriers in the *N* region, and will therefore cause another rise in the current.

This current change is to be associated exclusively with the highest positions of the density graphs of the electrons and holes in the *N* region.

The improvement in the conductivity of the *N* type material becomes the most important consideration when we are concerned with large signals, provided the diode is associated with components of relatively low resistance.

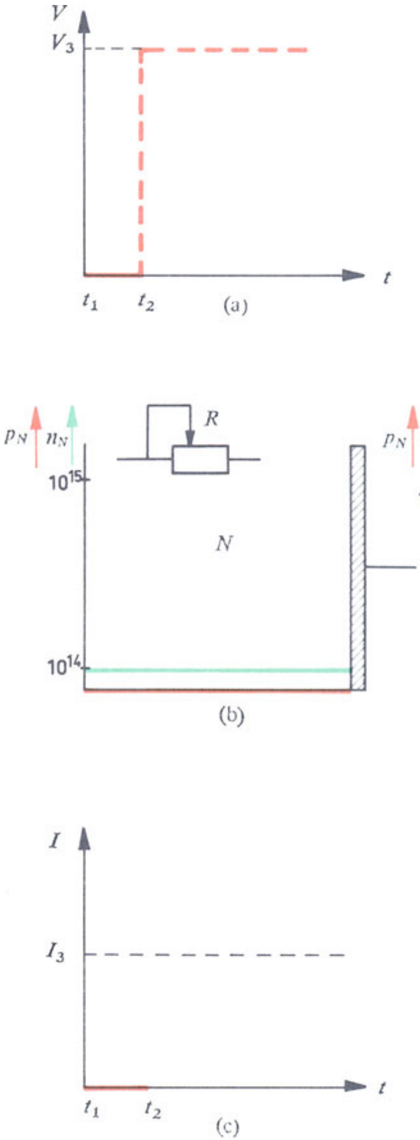

V
V3
t
t1
t2
(a)
pN
nN
R
10^15
N
10^14
(b)
I
I3
t
t1
t2
(c)

Fig. 74

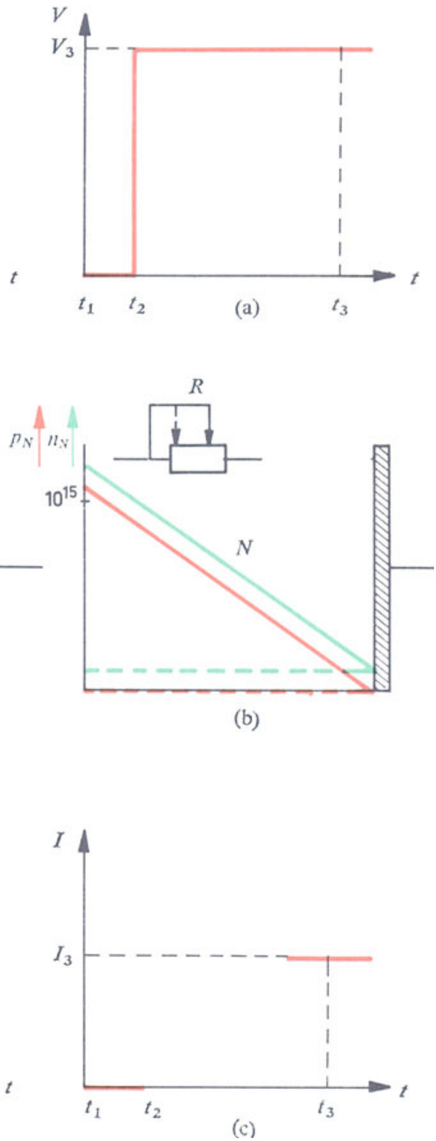

V
V3
t
t1
t2
(a)
t3
R
pN
nN
10^15
N
(b)
I
I3
t
t1
t2
t3
(c)

Fig. 75

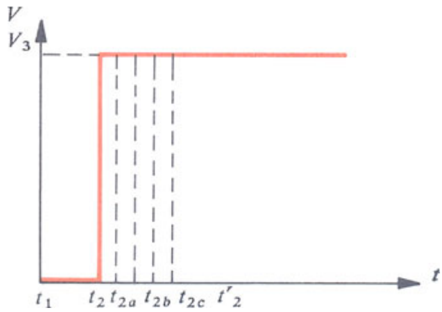


Fig. 76

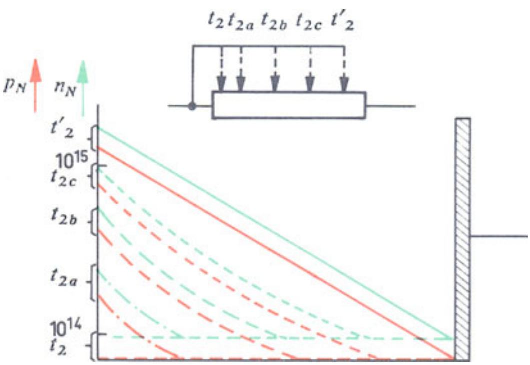


Fig. 77

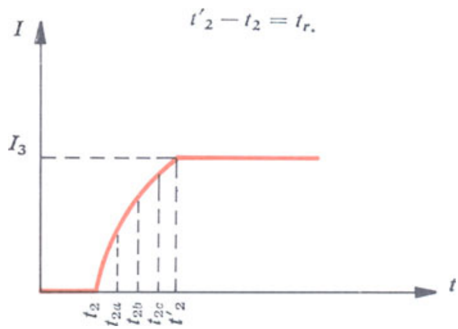


Fig 78

A very strong voltage pulse is applied to the diode at instant t_2; the diode is now in a circuit consisting of a generator and resistance in series. For numerous reasons, and particularly those associated with thermal problems, the load resistance associated with a semiconductor diode is chosen with a low value. The intrinsic resistance of the diode has consequently an important influence on the shape of the current in the circuit.

Instant t_2 on the abscissa of Fig. 76 is followed by instants t_{2a}, t_{2b}, t_{2c} and t_2'.

The signal is suddenly applied at instant t_2, and the hole and electron density graphs are displaced slightly from their initial positions. The number of majority and minority carriers hardly changes and the slider of the potentiometer remains in its initial position.

At instant t_{2a} the voltage V_3 has already been applied for some time. The holes have started to penetrate into the interior of the N region and have displaced the minority carrier density graph into the position shown by the time t_{2a} in Fig. 77.

Another position of the electron density graph is associated with this new position of the hole density graph. The result of the displacement of these two graphs is an improvement in the conductivity of the N material; this is shown by a movement of the slider of the potentiometer to the right (position t_{2a}).

If we express the value of the current by means of a simple application of Ohm's law ($I = V/R$), any change in the resistance for a given voltage causes an increase in the current in the circuit (t_{2a} in Fig. 78).

At instant t_{2b} following t_{2a}, the holes coming from the P region have contributed to the displacement upwards of the two density graphs, whilst the neutrality of the N region is maintained.

The slider of the potentiometer is again moved to the right and the current has increased (t_{2b} in Fig. 78).

The method is the same at instants t_{2a} and t_2' and the current continues to rise until it reaches its final value I_3 (instant t_2' in Fig. 78).

The time required for the rise of the current from zero to its final value (I_3) is called the 'rise time of the current in the circuit', that is

$$t_2' - t_2 = t_r$$

This current shape can be represented by a quantitative expression, but on account of its complexity and especially of the reader's unfamiliarity with the parameters, we will not consider it.

Influence of the amplitude of the applied signal

The level of the signal has an important influence on the shape of the current in the circuit. We will consider this shape by taking an intermediate position between small and large signals.

Small signals (Fig. 79). A voltage V_1 (Fig. 79a) is suddenly applied at instant t_2 to the terminals of a circuit consisting of a diode and a resistance. It is only the displacements of the hole density graph in the N region that explain the current variations as functions of time in the circuit (Fig. 79c).

Small to medium signals. Let us assume a somewhat larger signal, but still at a comparatively low voltage, e.g. V_2 at instant t_2 in Fig. 80a. The angle between the density graph of the holes in the N region and the horizontal plays an important part in the expression for the current. However, for higher positions of the density graph, the decrease in the resistance of the N type material causes a slight increase in the current in the circuit. Figure 80c illustrated this clearly. The capacitive effect is still most important but an inductive effect is beginning to appear.

Medium to large signals. For a still larger signal (Fig. 81a) the variations in the angle of the density graph of the holes in the N region becomes less and less important in relation to the decrease in the resistance of this material (Fig. 81b).

At instant t_2 a small current surge appears but is very soon masked by the tendency of the current to increase with time; the inductive effect has now become important.

Large signals. For a new value of applied voltage, namely V_4 at instant t_2 in Fig. 82a, the angle of the hole density graph plays no further part in the expression of the current in the circuit. It is only the changes in position of the electron and hole density graphs that affect the resistance of the N type material (Fig. 82b). The resistance decreases with time and consequently causes a fall in the current (Fig. 82c).

We can in fact liken the behaviour of the diode:

(a) to a capacitance effect in the first two cases;

(b) an inductive effect in the last two cases.

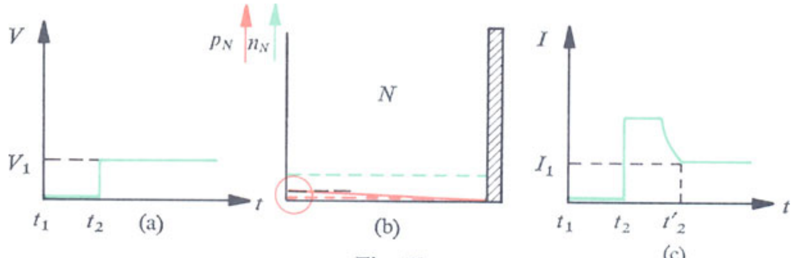


Fig. 79

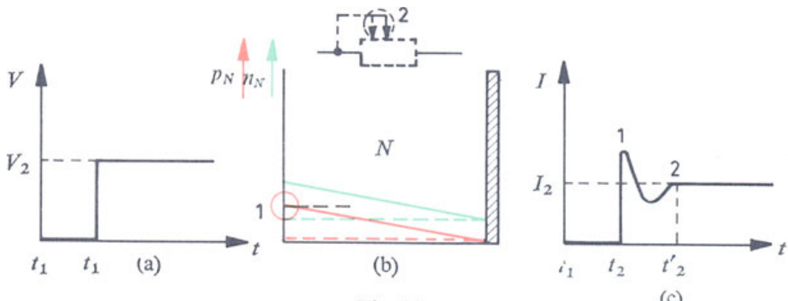


Fig. 80

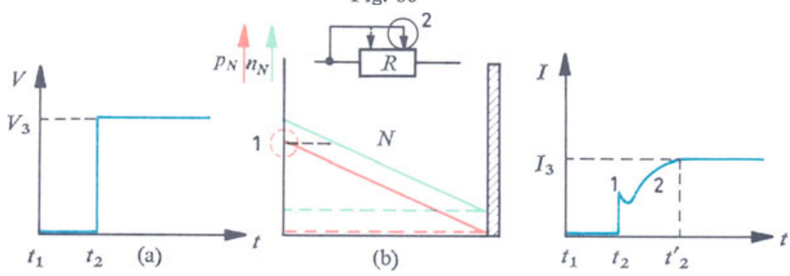


Fig. 81

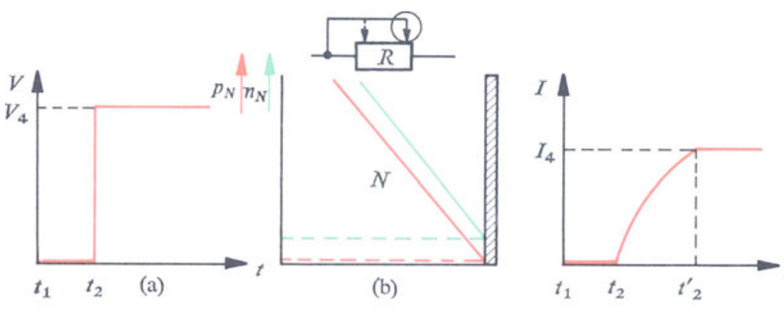


Fig. 82

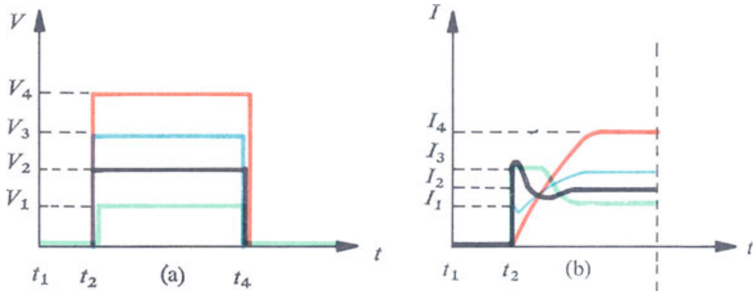


Fig. 83

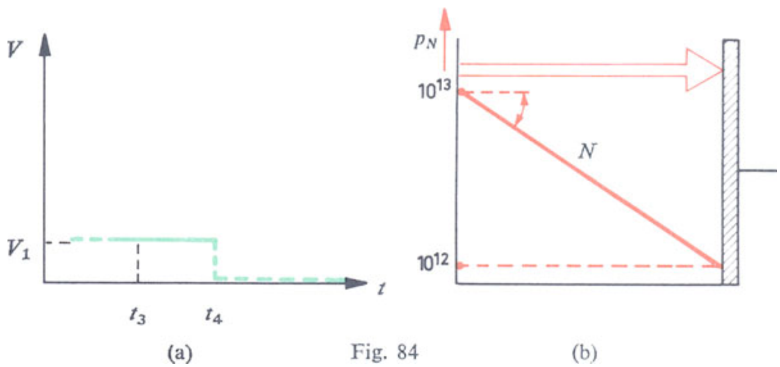


Fig. 84

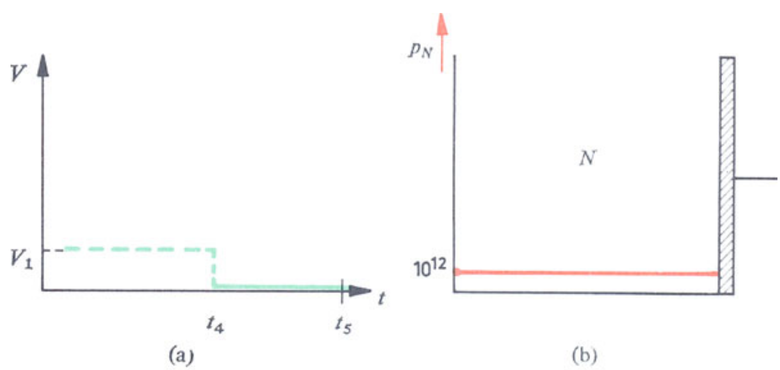


Fig. 85

Figure 83 shows the different results obtained in the above studies. The various levels of applied voltage are shown in Fig. 83a, whilst Fig. 83b represents the corresponding variations of the current with time.

The case of very small signals is shown in green. The voltage changes abruptly from zero at instant t_2 to V_1; the current changes are then defined by the concept of the 'forward recovery time of the diode': t_{fr}.

Shown in black in the same figures, a higher voltage V_2 represents the case of small to medium signals. There are initial inductive effects but they are negligible compared with the capacitive behaviour of the diode.

The voltage V_3 applied at instant t_2, shown in blue, applies to the case of medium to large signals. The capacitive effect has now become negligible by comparison with the inductive effect.

Shown in red in these two figures, the voltage V_4 shows the limit for very large signals. It is only the inductive effect that occurs and the current grows exponentially with time.

Switching from forward bias to zero bias

We have up to now defined the behaviour of a diode switched from zero to a forward bias. We must now enquire what happens at instant t_4 when the signal disappears from the terminals of the circuit.

Study of static behaviour. By contrast with the case of the rise of the current in the circuit, it is no longer necessary to differentiate here between the levels of the applied signal. In order to simplify matters we will limit our discussion to the behaviour of a system operating with small signals.

From instant t_3 to that preceding t_4, a voltage V_1 is applied to the terminals of the circuit (Fig. 84a). A certain position of the density graph of the holes in the N region is associated with this signal (Fig. 84b). The tangent of the angle between the graph and the horizontal determines the value of the hole diffusion current in the N region. This value has to be associated with the current I_1 in the circuit.

After instant t_4 (i.e. instant t_5 in Fig. 85a) the generator has stopped to deliver a signal. Instant t_5 is very far from instant t_4; the density graph of the holes in the N region is again horizontal (Fig. 85b). A zero value of current is associated with this position of the graph.

Transient effects. When the voltage disappears at instant t_4 (Fig. 86a) the bias is immediately removed from the diode. From instant t_3 to the instant preceding t_4 the current is maintained at a constant value equal to I_1. Conversely this current has totally disappeared at instant t_5.

Just as for the forward recovery time, the various resistances placed in series in the circuit affect the shape of the current. These resistance are due:

(a) to the diode (resistance R_d);

(b) to the generator (resistance R_g);

(c) to the load (resistance R_L), (Fig. 87a).

They may be regrouped in a single resistance (R) associated with a generator (without internal resistance) and a diode (Fig. 87b).

In order to simplify our discussion we will again approximate to the theoretical case by assuming that the total resistance of the circuit is near to zero. We are dealing at this stage with small signals, but this assumption about the resistance is difficult to reconcile because of the value of the resistance of the N type material. It will therefore be necessary to define, below, the effect of this resistance on the shape of the current.

Absence of resistance in the circuit. From instant t_3 to the instant preceding t_4 the applied voltage is equal to V_1 (Fig. 88a). The hole density graph is oblique, as shown in Fig. 88b.

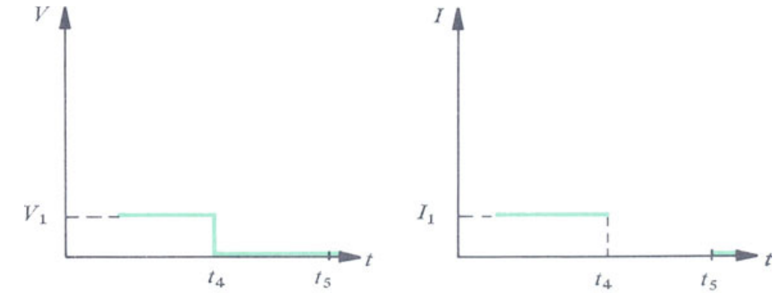
It is possible to obtain from this graph the value of the hole diffusion current (marked with arrow No. 1). During the first transient period, that of the rise of the current in the circuit, it was also possible for a hole current to appear from the angle '3' in this figure.

In fact, during the rise time two hole currents appear simultaneously in the N region: one which causes the density graph to swing to its final position (2) and the other to account for the current in the circuit (1). The diffusion current (1) is a function of the tangent of the angle (3) and of the 'storage' current of the holes (2) required for the displacement of the graph from the horizontal to an oblique position (4).

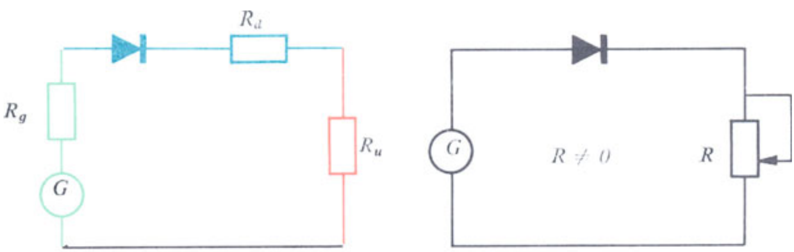
Whether the state is static or dynamic, the current outside the diode is always a function of these two currents:

(a) diffusion current on the one hand;

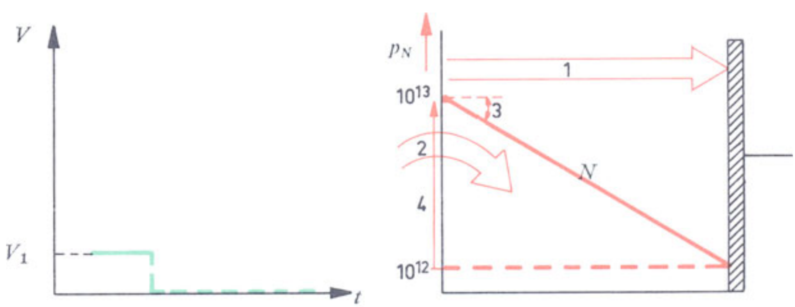
(b) carrier 'storage' current on the other hand.



(a) Fig. 86 (b)



(a) Fig. 87 (b)



(a) Fig. 88 (b)

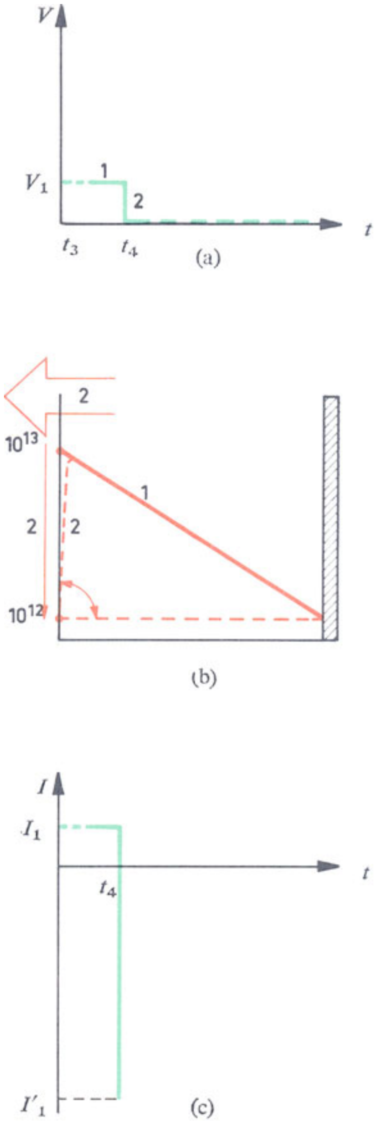

V
V_1
1
2
t
t_3
t_4
(a)
2
10^{13}
1
2
2
10^{12}
(b)
I
I_1
t
t_4
I'_1
(c)

Fig. 89

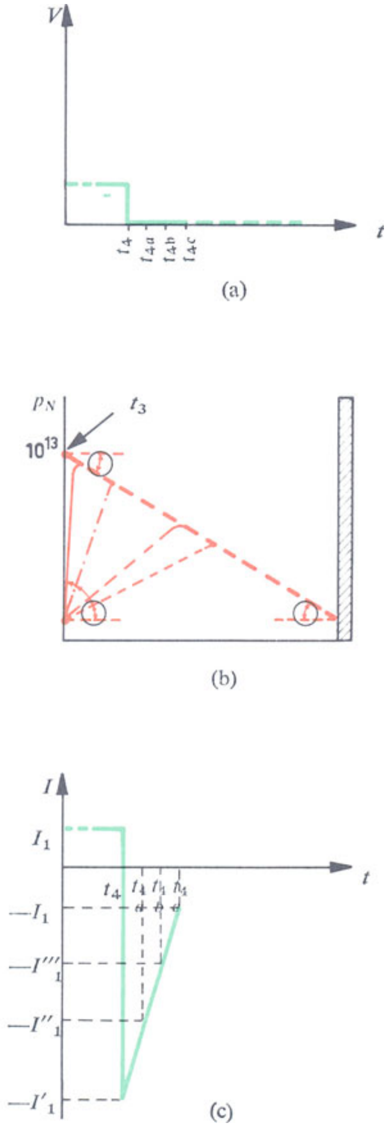

V
t
t_4
t_{4a}
t_{4b}
t_{4c}
(a)
p_N
t_3
10^{13}
(b)
I
I_1
t
t_4
$-I_1$
$-I'''_1$
$-I''_1$
$-I'_1$
(c)

Fig. 90

At instant t_4 the voltage supplied by the generator disappears suddenly (Fig. 89a). The barrier widens immediately and reverts to its equilibrium position; the holes in the P region can no longer penetrate into the N region; however, the surplus holes in the N region can easily pass once more into the interior of the P region (the reader is reminded of the analogy of the two planes at the beginning of this chapter).

If the circuit presents no resistance to the current, all the surplus holes in the N region return immediately to the interior of the P region. Consequently, at instant t_4 the hole density in the neighbourhood of the barrier in the N region, changes abruptly from the value 10^{13} to its equilibrium value 10^{12}.

Only the extreme left hand section on this region is affected by the departure of the holes; in these conditions the graph makes a very large angle with the horizontal. A large gradient towards the left appears (position 2 in the density diagram, Fig. 89b) and a very large hole current flows in the reverse direction. This large reverse current pulse (I_1') is shown in Fig. 89c.

From instant t_4 to instant t_{4c} (Fig. 90a) no signal is applied to the terminals of the junction.

At instant t_{4a} when the holes have left the left hand section of the N region, the graph is in an oblique position intermediate between that corresponding to instant t_4 and the horizontal. The angle decreases and the current assumes the value ($-I_1''$) shown in Fig. 90c.

At the instant t_{4b} the holes continue to diffuse towards the P region and the graph takes up a new oblique position below the previous one. This decrease in the angle is accompanied by a reduction in the value of the tangent; the current becomes still lower (namely I_1''' in Fig. 90c).

At instant t_{4b}' the density graph is in a special position; the left hand angle is equal to the right hand angle. The current, corresponding to instant t_{4c}, is exactly equal (in the reverse direction) to the forward current associated with instant t_3.

Finally, when the forward bias of a junction changes abruptly to zero, a very strong reverse current appears instantaneously in the circuit; the current then decreases until it reaches an absolute value equal to the forward current I_1.

The generator supplies no signal after instant t_4. On the abscissa of Fig. 91 the time is divided into equal intervals from instant t_4 to instant t_4'.

At instant t_3 the diode is forward biased by the voltage V_1. The density graph of the holes in the N region is as shown by the dotted lines in Fig. 92. The current I_1 is directly proportional to the tangent of the angle between the graph and the horizontal near the barrier.

As there is no resistance in the circuit, at instant t_4 the hole density graph passes instantaneously from the level 10^{13} to 10^{12} near the barrier. The corresponding angle is at the origin of a large reverse current (Fig. 93).

From instant t_4 to instant t_{4a} the holes continue to diffuse into the P region and the graph gradually approaches a horizontal position. The angle and the current in the circuit both decrease (instant t_{4c} in Fig. 93).

At this instant the reverse current is identical with that flowing previously in the forward direction. This equality was verified on the previous page by the comparison of the angles made by the graph.

At instant t_{4d} the holes are still passing into the P region: the corresponding density decreases in the interior of the N region. The value of the angle between this graph and the horizontal is still smaller as shown in Fig. 93 by a new decrease of current in the circuit.

A very small angle of the hole density graph (Fig. 92) and a low value of current (Fig. 93) are associated with instant t_{4e}.

At instant t_4' the graph is once more horizontal; there is no current in the circuit and the tangent of the angle between this graph and the horizontal is practically equal to zero, (Fig. 92). At this instant all the surplus holes in the N region have been removed.

When the bias of a diode is switched from the forward direction to a value near zero, a very strong reverse pulse appears in the circuit, and the value of this pulse increases as the total resistance of the circuit approaches zero.

$t_4'-t_4$ defines an inactive period of the diode. No information transmitted to the diode from instant t_4 to instant t_4' can affect the value of the current in the circuit.

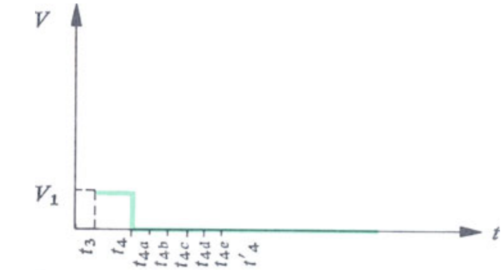


Fig. 91

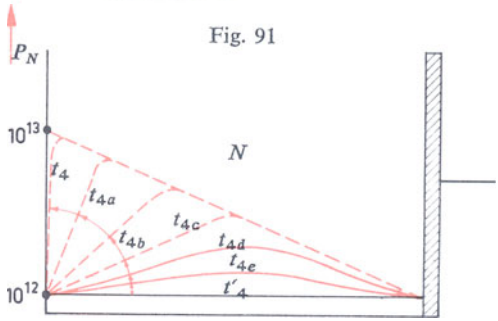


Fig. 92

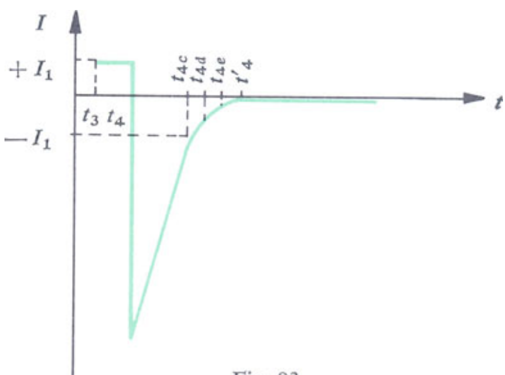


Fig. 93

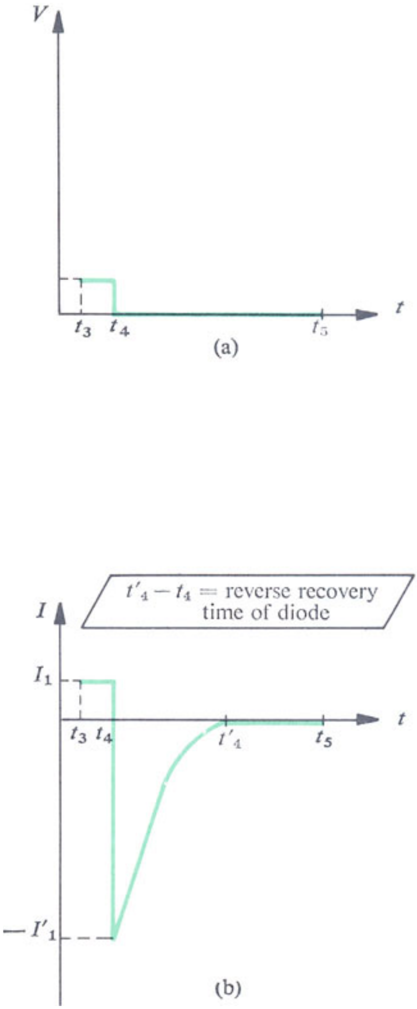

V
t
t3
t4
t5
(a)
t'4 − t4 = reverse recovery time of diode
I
I1
t
t3
t4
t'4
t5
− I'1
(b)

Fig. 94

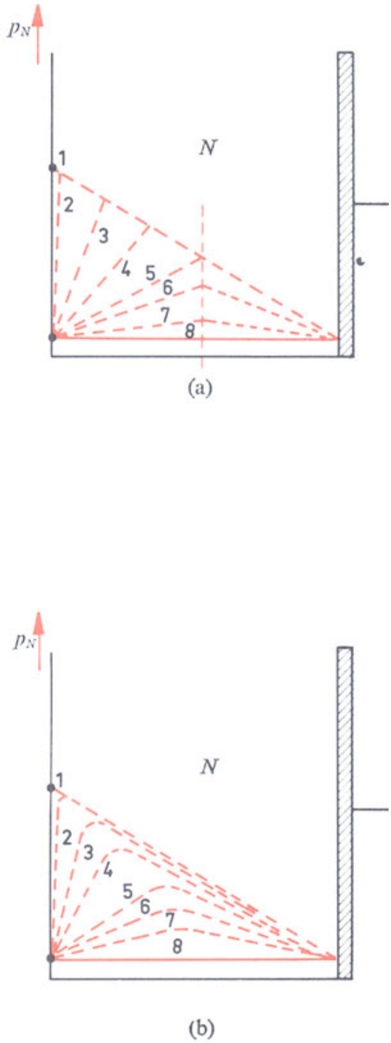

pN
N
1
2
3
4
5
6
7
8
(a)
pN
N
1
2
3
4
5
6
7
8
(b)

Fig. 95

Definition of the reverse recovery time of the diode. Finally, from instant t_3 to the instant preceding t_4 the diode is biased in the forward direction (Fig. 94a). During this time the density graph is oblique; the holes diffuse into the N region and a current equal to I_1 flows in the circuit (Fig. 94b).

A very strong current pulse appears at instant t_4. This pulse is accounted for by the disappearance of the voltage supplied by the generator. With a theoretical zero value of the resistance in the circuit, this pulse would normally be near to infinity. However, the resistance has not been taken as equal to zero, but only with a value not far from zero; the value of the current is therefore limited to I_1' in Fig. 94b.

The absence of bias, from instant t_4 to instant t_5 in Fig. 94a, causes a removal of the holes stored in the N region and the displacement of the density graph. In these conditions, the reverse current decreases and approaches zero, the value reached at instant t_4' in Fig. 94b. The interval required for the current to reach zero is known in data sheets as 'the reverse recovery time of the diode': t_{rr}.

This reverse recovery time of the diode is very important for the following reasons:

(a) it is accompanied by a very strong reverse current pulse, account of which must always be taken in circuit calculations;

(b) during this interval the diode cannot 'react' to any information fed into it.

In purely theoretical studies it is often felt to be sufficient to represent the different positions of the density graph by straight lines, as shown in Fig. 95a. The positions 1, 2, 3, 4, 5, 6, 7 and 8 correspond to different time intervals between t_4 and t_4'.

In practice, however, all the zones of the N region are affected by the removal of the carriers and the group of graphs moves from left to right. The effective positions of the density graph for the above defined are shown in Fig. 95b. The simplified version in Fig. 95a is usually satisfactory for purely quantitative studies.

Influence of the resistance of the circuit on the reverse recovery time of the diode. For purely academic reasons the total resistance inserted in the circuit has been assumed to be near to zero. We must now ask what effect an increase in this resistance has on the amplitude and the shape of the current.

Figure 96 shows a circuit consisting of a diode in series with a generator and a resistance R of 'relatively high' value. Before defining the transient behaviour when the voltage is interrupted it is well to remind ourselves of the reasons for our choice of a voltage V_1'' higher than V_1 (Fig. 97). The reason for this increase in the applied voltage is to maintain the forward current at the value I_1' (Fig. 98) during a period of static operation.

From instant t_3 to the instant preceding t_4 a voltage V_1' is applied to the terminals of the circuit (Fig. 98a). The diode is forward biased and the density graph of the holes in the N region is oblique (Fig. 98b). The tangent of the angle between this graph and the horizontal determines the value of the current of holes in the N region. It is now easy to deduce from this the current strength in the circuit (I_1 in Fig. 98c).

At an instant t_5, far removed from t_4, the generator still does not supply a signal (Fig. 99a). Since the time interval between the instant of the interruption of the voltage and instant t_5 is very long, it is possible to assume that all the surplus holes have in fact left the N region, and the density graph will again be horizontal (Fig. 99b). A zero value of the angle, and therefore of the current in the circuit (at instant t_5 in Fig. 99c), is associated with this position of the diagram of the hole density in the N region.

It is usual before considering the transient effects to define very precisely the two static states existing before and after the effect under study. This rule should always be observed as it enables the limiting values of the currents in the circuit to be determined.

It also has the advantage that it locates the extreme positions of the hole density graphs in the N region, and that it enables the exact value of the maximum hole density near the barrier to be determined.

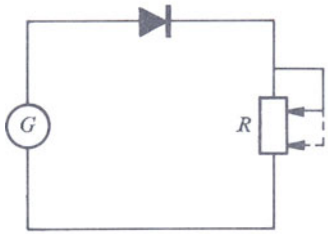


Fig. 96

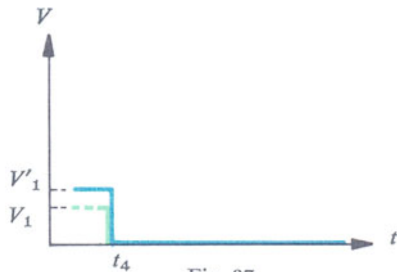


Fig. 97

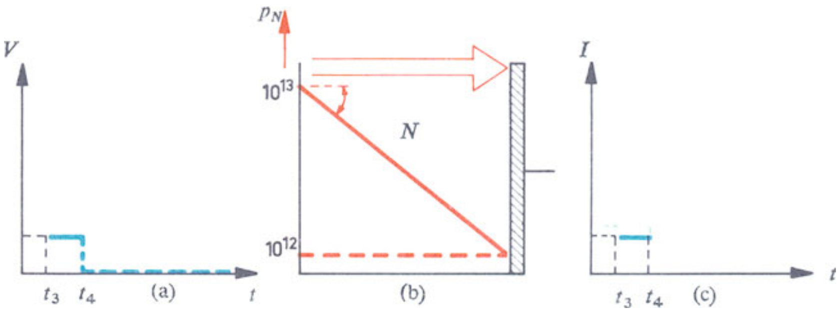


Fig. 98

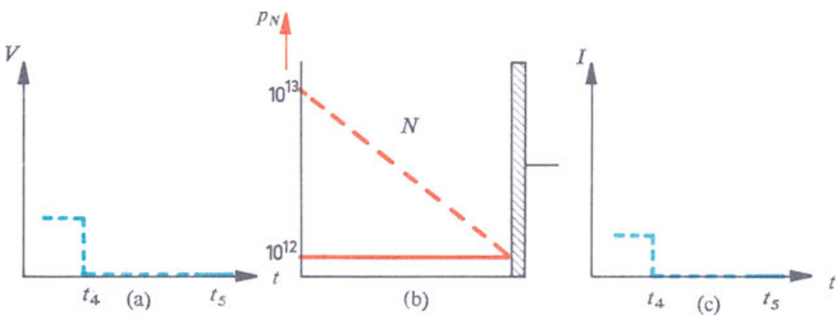


Fig. 99

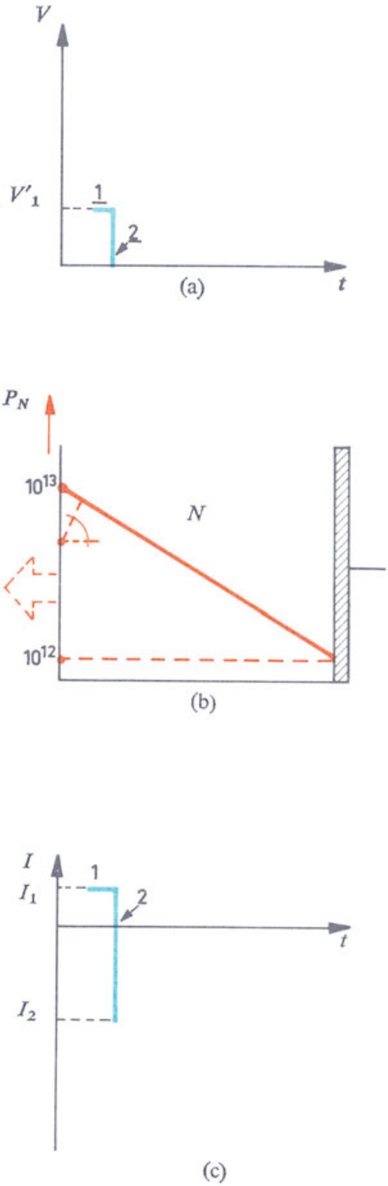


Fig. 100

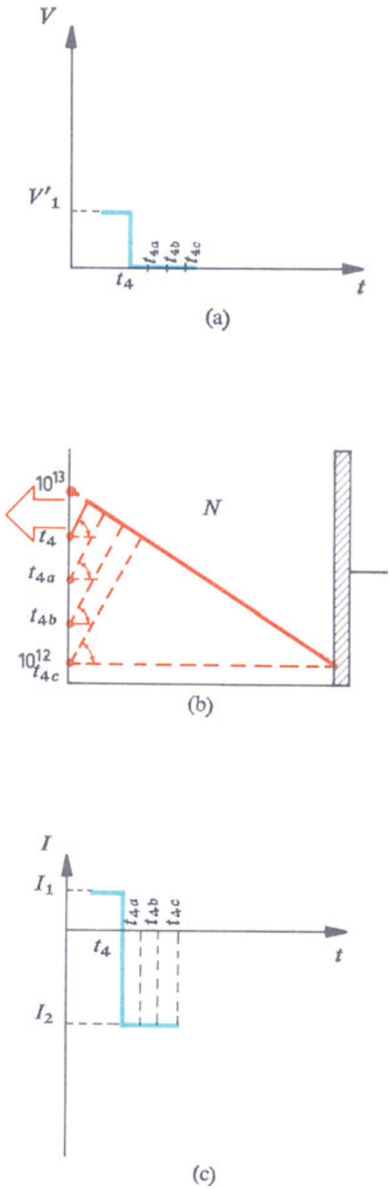


Fig. 101

Transient behaviour of the diode (with a resistance). The voltage supplied by the generator before instant t_4 was equal to V_1' (stage 1 in Fig. 100a). The density graph was oblique (Fig. 100b); the diffusion current of holes in the N region was a function of the tangent of the angle between this graph and the horizontal. This value of the current is shown as 1 in Fig. 100c.

At instant t_4 the voltage V_1' is suddenly interrupted, and the barrier immediately widens. The diffusion of the holes from the P to the N region stops although they may well travel in the reverse direction. As there is a resistance in the circuit all the holes do not return instantaneously into the interior of the P region. This resistance slows up the flow of the carriers through the barrier, and the holes pass through it in only limited numbers compared with the previous example.

Consequently the hole density near the barrier in the N region does not at once reach its equilibrium level, but it assumes a higher level, namely position 2 in Fig. 100b. The angle between this graph and the horizontal occurs in the expression of the hole current flowing towards the P region; this current is equivalent to the flow of a very large reverse current. At instant t_4 a large reverse current pulse appears in the circuit (I_2 in Fig. 100c).

No signal is applied to the terminals of the diode after instant t_4. However the N region contains a certain number of surplus holes compared with the equilibrium density; at instant t_4 the graph is practically in the position marked t_4 in Fig. 101b.

As the holes continue to diffuse towards the P region at instant t_{4d}, the density near the barrier decreases and approaches its equilibrium level; simultaneously the central part of the N region (position t_{4d}) is more and more influenced by the departure of these carriers. In these conditions the angle of the density graph of the holes in the N region remains unchanged (position t_{4d}). If this angle maintains its value the hole current does not change; this is shown in Fig. 101c. The current at instant t_{4d} is still equal to I_2 defined at instant t_4.

At instant t_{4b} the holes are still passing into the P region; the density level near the barrier continues to decrease as well as the hole density in the most central regions of the block of N type material. As before, the angle remains constant and so the current at this instant is still equal to I_2.

At instant t_{4c} the holes are still passing into the P region and the corresponding density graph reaches its equilibrium position near the barrier. The angle between this graph and the horizontal has still not changed; the current is still equal to I_2 (Fig. 101c).

From instant t_{4c} the hole density in the N region reaches the level 10^{12} near the barrier. In these conditions, it can no longer vary and the graph will then tilt and approach the horizontal position.

Let us take the time interval t_{4d} (Fig. 102a) by way of example. No voltage is applied to the diode; the holes diffuse from the N into the P region, the density decreases and the graph turns downwards (t_{4d} in Fig. 102b).

A smaller angle is associated with this position of the hole density graph, consequently the current has decreased (Fig. 102c).

At instant t_{4e} the graphs are in the form of an isosceles triangle on a horizontal base. The two angles are equal and the reverse current equal to the forward current present earlier (t_{4e} and I_1 in Fig. 102c).

The absence of signal is maintained from instant t_{4e} to instant t_4' in Fig. 103a.

The hole density graph continues to approach the horizontal position whilst maintaining the different intermediate stages shown by instants t_{4f}, t_{4g} and t_{4h} in Fig. 103b. A very precise angle of the density graph of the holes in the interior of the N region is associated with each of these positions. This makes it possible to determine the corresponding values of current (Fig. 103c).

From instant t_{4e} (reverse current equal to the forward current) to instant t_4' (total disappearance of the current in the circuit) the current decreases gradually, taking up the different values determined by instants t_{4f}, t_{4g} and t_{4h}.

The insertion of a resistance in the circuit results in a big decrease in the reverse current pulse (although larger than the value of the forward current) as well as a very large increase in the reverse recovery time of the diode.

With slow switching this recovery time is not too inconvenient, because of the time interval between two successive pulses constituting the information transmitted to the circuit.

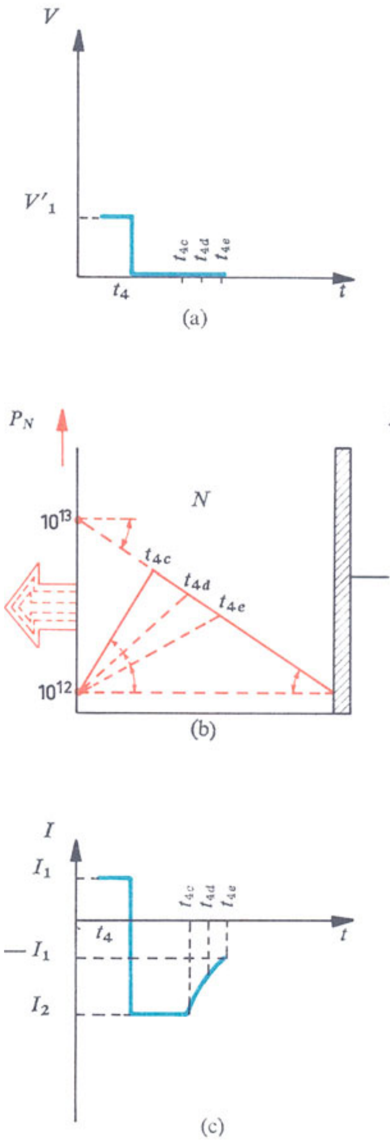


Fig. 102

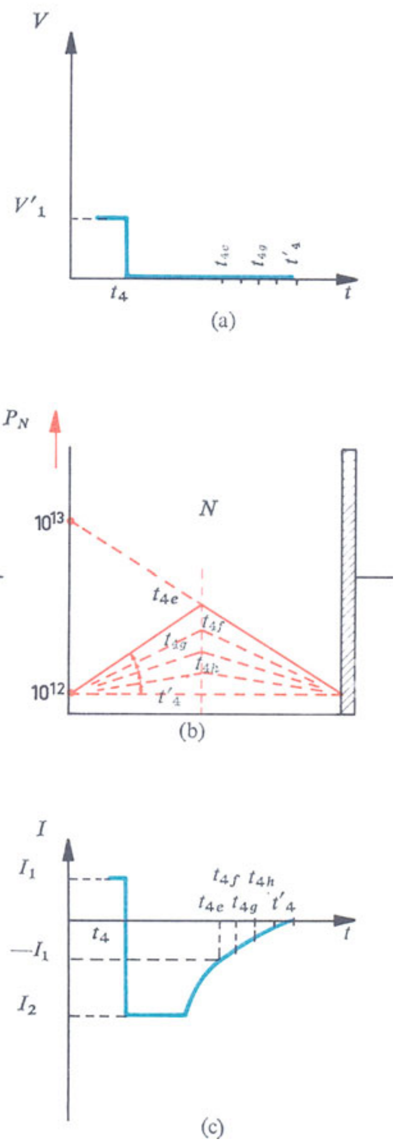


Fig. 103

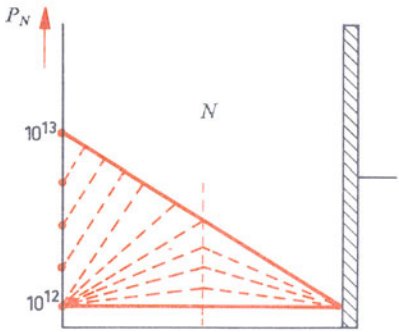


Fig. 104

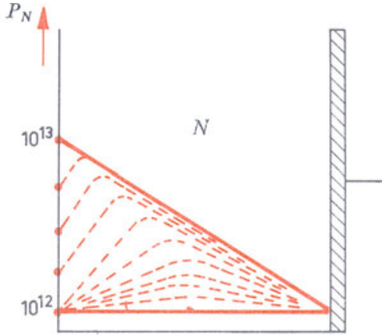


Fig. 105

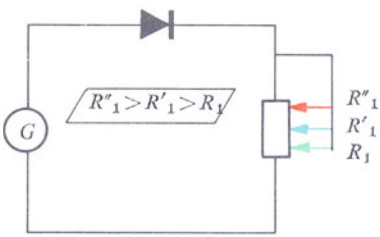


Fig. 106

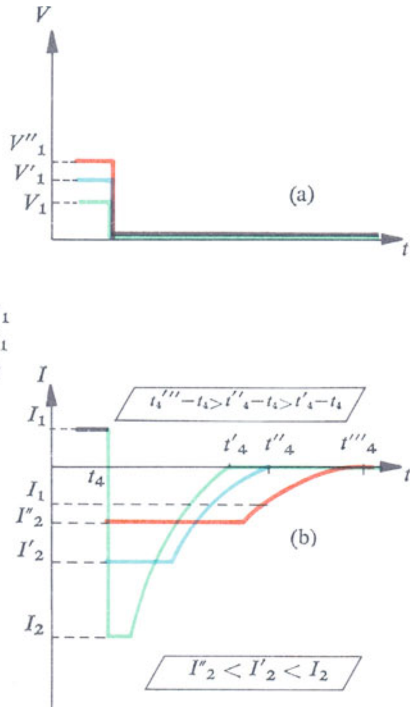


Fig. 107

On the other hand, for signals with a high recurrence rate, this reverse recovery time becomes very troublesome as there is a danger that it may block the diode whilst information is being transmitted.

As in the study of the reverse recovery time of the diode, it is possible in the absence of resistance:

(a) either to use the simplified representation in Fig. 104; the different positions of the density graphs of the holes in the N region are represented by straight lines changing direction with respect to the equilibrium points of this density near the barrier in the neighbourhood of the ohmic contact;

(b) or to respect the exact forms of the various graphs which are represented by the curves in Fig. 105.

In practice the simplified version in Fig. 104 is quite adequate.

Finally, when the bias of a diode changes from the forward direction to zero, a strong reverse current pulse appears in the circuit; the duration of this pulse determines the reverse recovery associated with the concept of the 'reverse recovery time of the diode' as defined in data sheets.

Influence of the resistance on the reverse recovery time. In order to determine the influence of the resistance placed in the circuit, we will examine three possible values of R: namely R_1 smaller than R_1', which itself is less than R_1'' (Fig. 106).

These three resistances are associated with three possible positions of the slider of the potentiometer.

(a) R_1 low value of resistance (green position);

(b) R_1' mean value of resistance (blue position);

(c) R_1'' high value of resistance (red position).

For a low value of resistance the interruption of the voltage gives rise to a strong reverse current pulse (I_2 Fig. 107b). The current remains at this value for a very short time and then decreases until it disappears at instant t_4'. $t_4'-t_4$ then gives the reverse recovery time of the diode.

For a medium value of resistance R_1' a reverse current pulse I_2' such that $I_2' < I_2$ appears at instant t_4. The current remains at this value for a longer time than before (Fig. 107b), and it finally reaches zero at instant t_4''.

If we take a still larger resistance (R_1'' marked in red in Figs 106 and 107) the voltage variation from V_1'' to 0 at instant t_4 is accompanied by

(a) a current pulse I_2'' smaller than I_2';

(b) a reverse recovery time $t_4'''-t_4$ longer than $t_4''-t_4$.

The increase in resistance causes an increase in the reverse recovery time of the diode, as well as a decrease in the current pulse.

$$R_1 < R_1' < R_1''$$
$$I_2 > I_2' < I_2''$$
$$t_4' - t_4 < t_4'' - t_4 < t_4''' - t_4$$

The concept of the reverse recovery time tied to the presence of a very strong current pulse in the circuit does not in fact involve the reverse bias of the diode. The current changes were determined only by the change from forward bias (V_1 in Fig. 108a) to zero bias.

In this case, and in order to recapitulate the steps explaining the different values of current, the reader is referred to Fig. 108b.

Phases 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 represent the various positions of the diagram of the hole density graph in the interior of the N region. These positions are associated with very precise values of the reverse current in the circuit (Fig. 108c). At instant t_4' the diode is not biased because no signal is supplied by the generator and the current in the circuit is practically zero.

Change from forward to zero bias

What happens if at instant t_4 instead of suppressing the forward bias of the junction we apply simultaneously a strong reverse voltage pulse ($-V_2$ in Fig. 109a)?

The density graph is first shown in full lines (before instant t_4) in the upper oblique position in Fig. 109b; by means of this figure we can calculate the corresponding values of current flowing through the circuit.

A certain time after instant t_4 the diode is biased in the forward direction and the density graph is consequently in its lowest position. To this new position of the graph corresponds a current of value equal to $-I_2$ in the circuit. (Fig. 109c).

How does the graph shift from the higher to the lower position?

The method developed earlier can be used again, giving the different positions shown by dotted lines in Fig. 109b. We can then calculate the corresponding values of current and transfer them to Fig. 109c.

The fact of having changed the bias of a junction (i.e. from forward to reverse) has very little effect on the shape of the current in the circuit; on the other hand, the reverse recovery time of the diode shortens. The strong reverse bias causes a widening of the space charge regions of which the barrier consists, that is a substantial reduction in the N region with, as a consequence, a marked decrease in the number of holes to be removed.

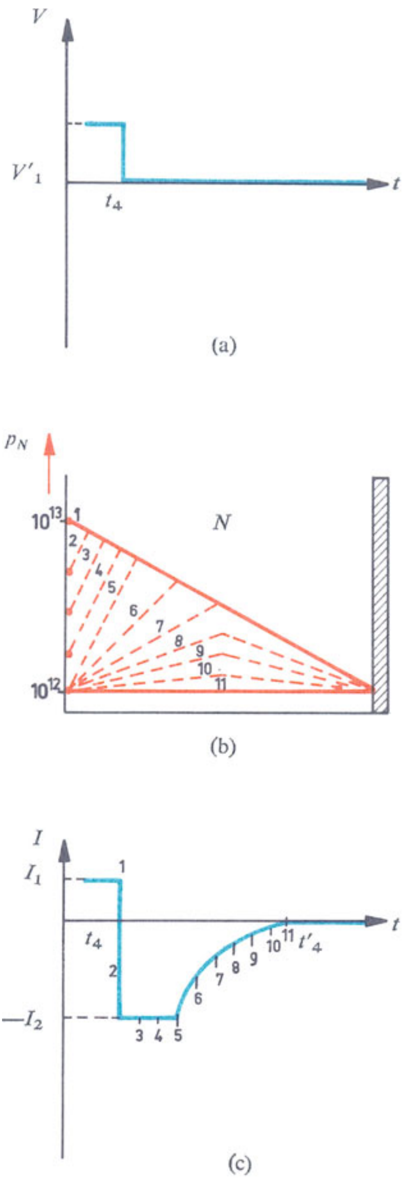


Fig. 108

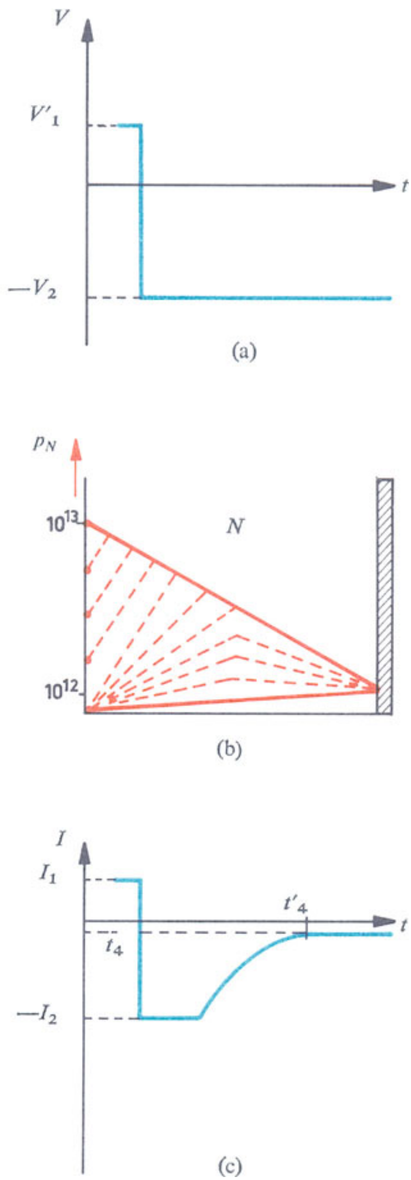


Fig. 109

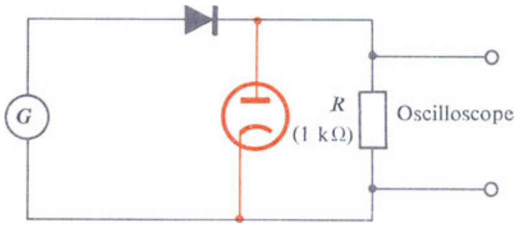


Fig. 110

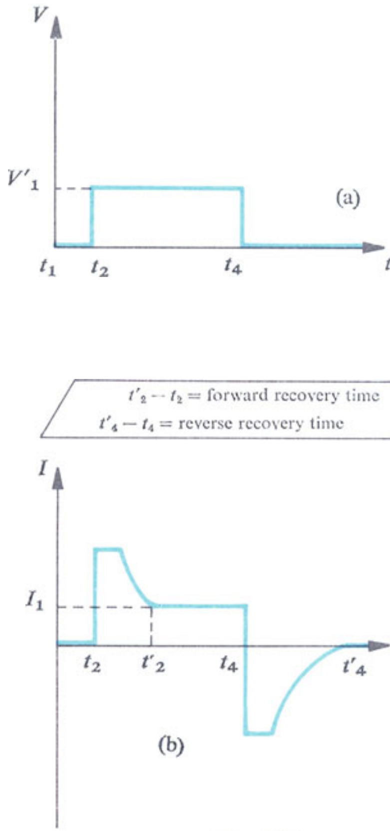


Fig. 111

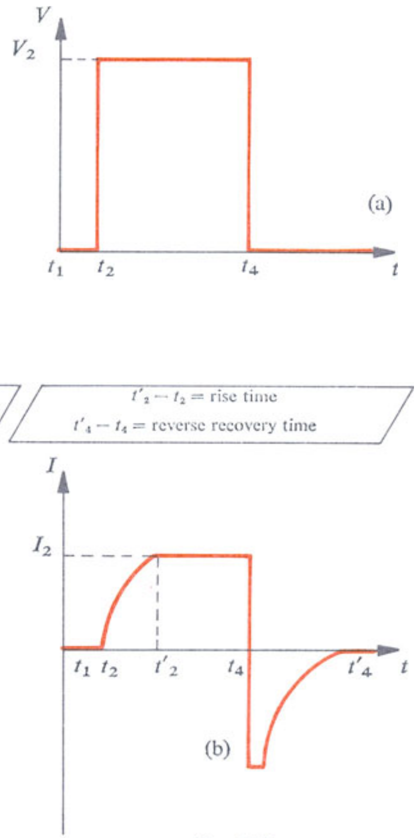


Fig. 112

Measurement of the reverse recovery time of the diode. Manufacturers of diodes and transistors very often include in their data sheets the forward and reverse recovery times. A measurement of these parameters, especially the reverse time, may be of interest.

Figure 110 represents one of the very simple circuits by means of which this time can be measured.

The circuit consists of a generator with a resistance in parallel with an oscilloscope; the diode is connected in series between these two units. Further, a thermionic diode (shown in red in the figure) is connected in parallel with the resistance.

The role of this diode is to short-circuit the resistance during the conductive period of the semiconductor diode. When the diode under test is non-conducting, the termionic diode is also in the reverse state, and therefore prevents the flow of any current. All changes in the current in the circuit pass through the resistance R, and the voltage, which is a function of these changes, appears at the terminals of the resistance; it is consequently applied directly to the input terminals of the oscilloscope.

Overall forms of the signals. Finally, there are two important possibilities for the use of a diode in switching circuits.

Small-signal state. The operation of a diode in the small-signal state during its conductive state is illustrated in Fig. 111a. From instant t_1 to the instant preceding t_2, there is no signal applied and the current in the circuit is zero, Fig. 111b. At instant t_2 and until instant t_2', a forward current pulse appears in the circuit (a pulse associated with the concept of the forward recovery time of the diode). From instant t_2' to instant t_4 the current does not vary. From instant t_4 to instant t_4' a strong current pulse appears in the reverse direction, and this pulse is determined by the reverse recovery time of the diode.

Large-signal state. With large signals (Fig. 112a), from instant t_1 to the instant preceding t_2, no current flows in the circuit (Fig. 112b). From instant t_2 to instant t_2' the current increases steadily as a function of time. $t_2' - t_2$ determines the rise time of the current in the circuit (t_r). From instant t_2' to instant t_4 the current remains at a constant value. From instant t_4 to instant t_4' a very strong reverse current pulse appears in the circuit.

2. Transistors for Switching

To study the operation of a transistor when used for switching means to define the different shapes of the currents and voltages which may appear in the components of the circuits with which they are associated.

Before we examine in detail the transient effects which are liable to affect more particularly the output signal, we must remind ourselves of the physical principles associated with the static study of transistors.

2.1. Basic physical study

We will start with a *PNP* transistor (Fig. 113); it consists of two *P* regions separated by a very thin *N* region at the emitter and collector ends.

The quantitative study of a transistor, and indeed the determination of the values of the current flowing through the emitter, base, and collector electrodes of this transistor is considerably simplified by the use of density graphs.

Absence of bias

In order to be able to represent simultaneously the density graphs of majority and minority carriers, we will use an ordinate graduated logarithmically (Fig. 114). The emitter is represented by the left side, the *P* region; it contains 10^{16} holes per cubic centimetre and only 10^{10} electrons. The centre section, the *N* region, represents the base; it contains only 10^{14} electrons and 10^{12} holes. Finally, the right hand section represents the collector (the *P* region); it contains 10^{16} holes per cubic centimetre and 10^{10} electrons.

These three regions are separated by barriers consisting of space charge regions. The barriers are shown graphically by decreasing or increasing densities, depending on the direction of the observation. The thickness of the centre region, or the base of the transistor, is very low (of the order of 50 μm).

The density graphs of the majority and minority carriers are always very difficult to use simultaneously, especially for the accurate determination of the values of the currents flowing through the three electrodes of the transistor.

We will, therefore, limit our study to the use of simplified graphs representing only the minority carrier densities on ordinates graduated linearly.

In Fig. 115 we again show the three regions of the transistor: emitter, base and collector, and the hole density graph in the base is represented by a horizontal line located at the level 10^{12}. On the other hand the two regions, emitter and collector, are represented by two green straight lines

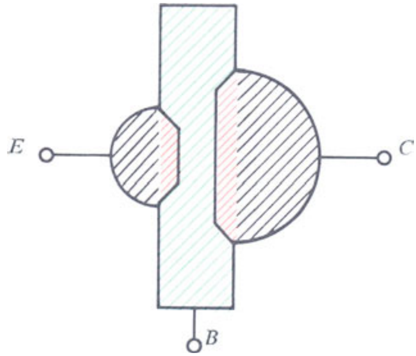


Fig. 113

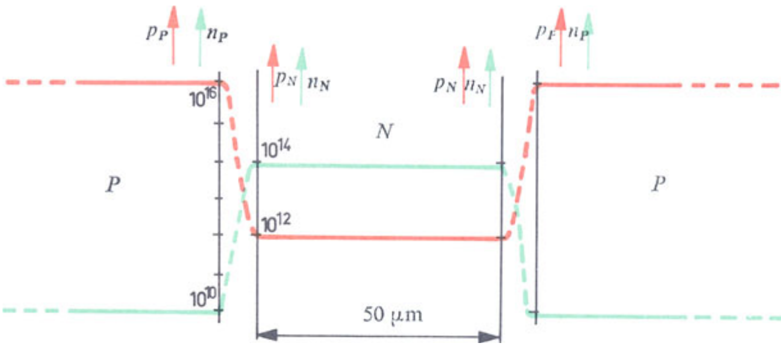


Fig. 114

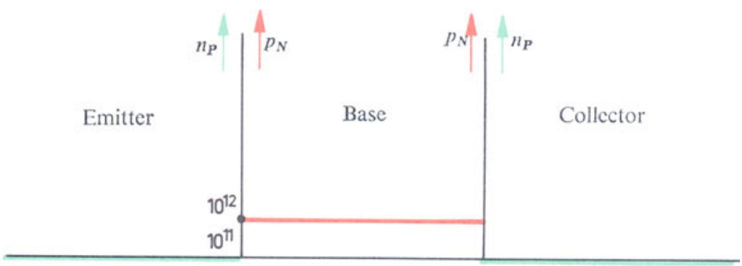


Fig. 115

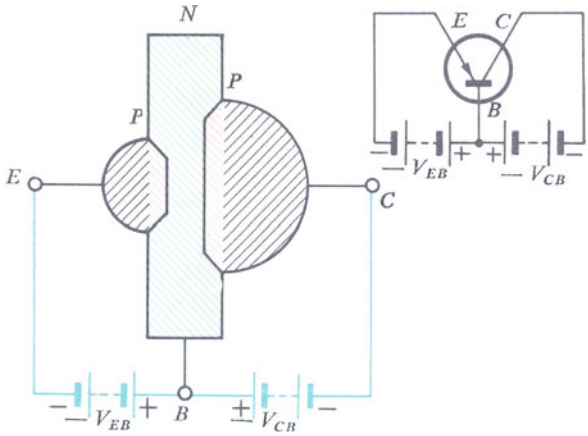


Fig. 116

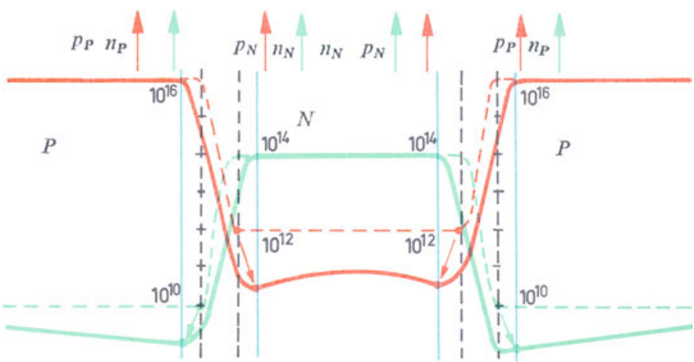


Fig. 117

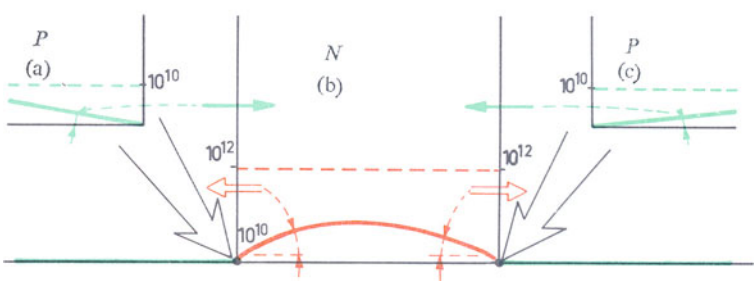


Fig. 118

superimposed on the abscissa; the level 10^{10} is in fact superimposed on this axis. The two barriers on a linear scale will be shown only for special operating states of the transistor.

2.2 Operating conditions of a transistor

A transistor consists of two junctions. There are two possibilities to bias a junction; it is, therefore, possible to specify four operation states of the transistor. The table below summarises these four states.

emitter–base junction	collector–base junction	operation states
reverse direction	reverse direction	cut-off
forward direction	reverse direction	normal
reverse direction	forward direction	reverse
forward direction	forward direction	saturation

2.3. Cut-off state

In the first possibility the two junctions are biased simultaneously in the reverse direction (Fig. 116). In this case the emitter and collector are strongly negative with respect to the base.

The two barriers become wider and only the minority carriers can cross it. The minority carrier density graphs move downwards (Fig. 117).

The new positions of the density graphs are shown by full lines on the linearly graduated axes (Fig. 118). In order to illustrate clearly the angles of the density graphs of the electrons in the *P* regions, an enlargement of the region near the origin is shown in Figs. 118a and c.

The currents are directly connected with the value of the angles made by the various diagrams, namely:

(a) for the emitter current by the density graph of the electrons in the *P* region on the left hand side (Fig. 118a) and by the density graph of the holes in the *N* region on the left of the *N* region (Fig. 118b);

(b) for the collector current by the density graph of the electrons in the *P* region on the right (Fig. 118c) and by the density graph of the holes on the right of the *N* region (Fig. 118b). These currents are very small and the transistor is then in the cut-off state.

The base current is equal to the sum of these two currents.

Common-emitter configuration

The common-emitter configuration is one of the most commonly used in switching circuits. The cut-off conditions for a circuit of this type are shown in Fig. 119.

This condition is determined by the reverse bias of the two junctions. In order to obtain a result of this sort all that is needed is to make the collector and emitter negative with respect to the base; the battery V_{BB} must, therefore, be connected as follows: negative pole, emitter side; positive pole, base side; similarly the collector is negative when the group of batteries represented in the figure by $-V_{CC}$ has the negative pole connected to the collector through the load resistance.

Cut-off region on the network of output characteristic curves. It is interesting to be able to locate in the network of characteristics shown in Fig. 120 the region corresponding to the use of the transistor in its cut-off state. This state by definition involves the reverse bias of the two junctions; how is it possible in this network to represent the reverse bias of the base–emitter junction of the transistor?

All the characteristics situated above the characteristic for $I_B = 0$ are associated with negative values of the base current when the base is negative with respect to the emitter. The reverse bias of this emitter then corresponds to a positive value of the base current, that is to the region below the characteristic for $I_B = 0$. The reverse bias of the collector–base junction is produced by all the values of collector–emitter voltage greater than $-V_{CEK}$.

The cut-off region is therefore located in the blue shaded section in Fig. 120. If we take account of the load line, the cut-off region imposes to the working point a position between the limits A and B shown in this figure.

Equivalent static circuits. The data supplied by the manufacturer often includes dynamic equivalent circuits of the transistor. Before we embark on the study of these circuits it would be useful to define the simplified equivalent circuits known as 'static equivalent circuits' of the transistor. We have two cases to consider:

(a) the transistor is voltage driven (Fig. 121);

(b) the transistor is current driven and the minority carriers reaching the base can no longer make use of the external circuit connected between the base and emitter; in these conditions the collector current 'benefits' from the transistor effect. This effect involves the connecting in parallel of a current generator with the base–collector junction (Fig. 122).

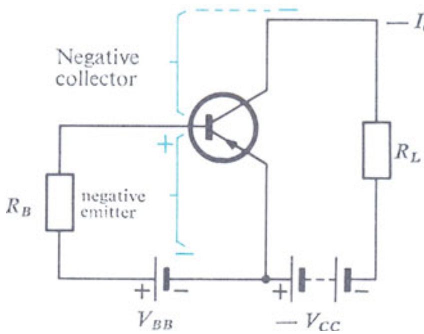


Fig. 119

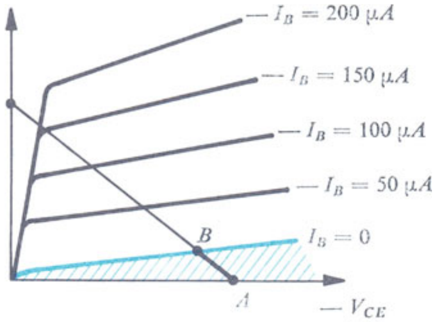


Fig. 120

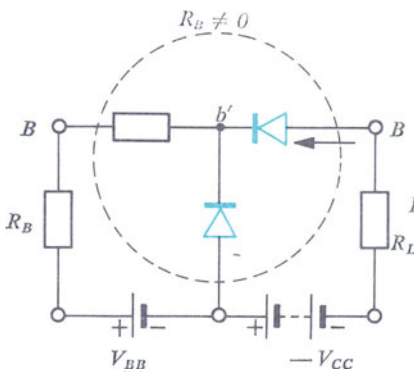


Fig. 121

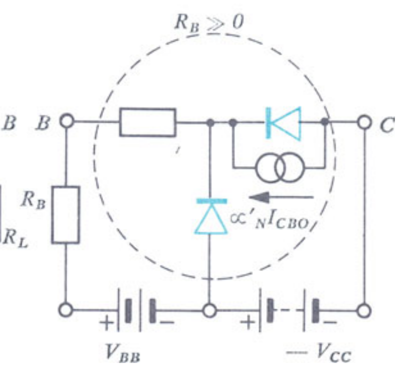


Fig. 122

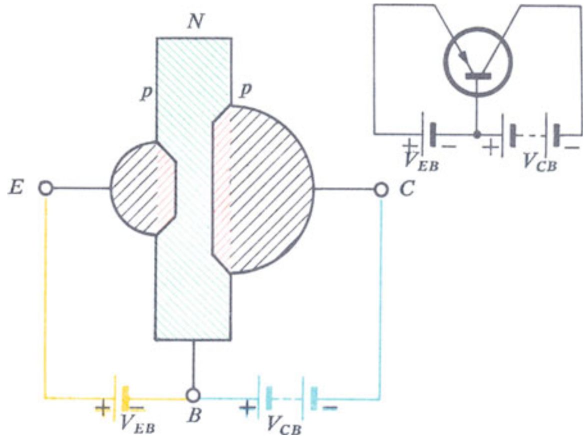


Fig. 123

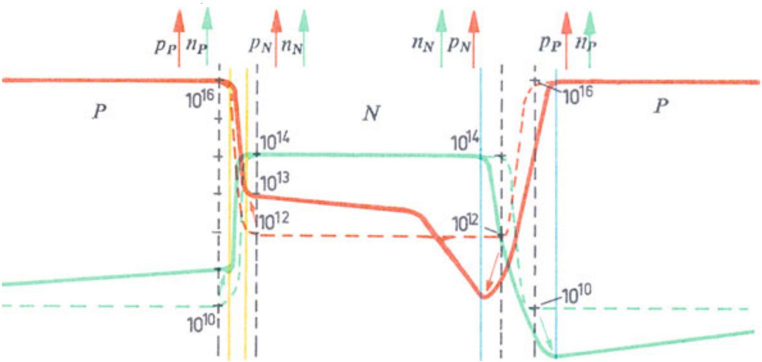


Fig. 124

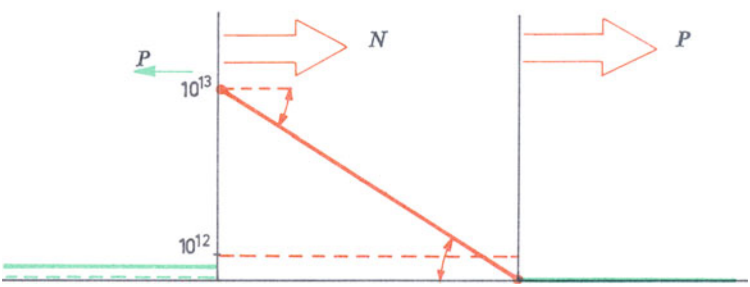


Fig. 125

2.4. Normal state

A second method of operation, very common in audio and radio-frequency practice, consists in using the transistor in its normal state. By definition this condition involves the forward bias of the emitter–base junction and the reverse bias of the collector–base junction. A result of this type can be obtained simply by connecting the positive pole of a group of batteries V_{EB} to the *P* region side (emitter, Fig. 123).

Forward bias of the emitter–base junction causes a narrowing of the corresponding barrier. The majority carriers, holes and electrons, can now pass through the barrier, in the direction *P* region to *N* region for the holes, and *N* region to the *P* region for the electrons (Fig. 124). The densities of the minority carriers increase and they are located above their equilibrium position.

Conversely, the reverse bias of the base–collector junction causes a widening of the corresponding barrier. Only the minority carriers pass through the new barrier. In these conditions the densities of the minority carriers decrease.

It is very difficult to make directly an interpretation of the emitter, base and collector currents from the complete presentation of the density graphs for the majority and minority carriers in Fig. 124. It is better in this case to work exclusively with the diagrams of the minority carrier densities with the ordinates graduated linearly. Fig. 125 illustrates an arrangement of this kind. In the *P* region, which constitutes the emitter of the transistor, the diffusion current of electrons is a function of the tangent of the angle between this graph and the horizontal; this angle is small and so the current is very low. In the *N* region, in which the base of the transistor is located, the hole current from the emitter is determined by the tangent of the angle between the hole density graph emitter side and the horizontal; the angle is relatively large and so the current also is large. The hole current flowing from the base to the collector is a function of the tangent of the diagram of the hole density in the *N* region near the barrier on the collector side (on the right in Fig. 125). This current is very nearly equal to the hole current entering the *P* region on the other side; the two angles are substantially equal as 'alternate-interior angles'.

Representation of the normal state on the network of output characteristic curves

With the common emitter configuration the normal state imposes forward biasing to the base–emitter junction. This requirement is respected in Fig. 126 when the negative pole of the battery is connected to the base and the positive pole to the emitter.

Conversely, the collector–base junction must be biased in the reverse direction, and this can be done simply by having a larger negative voltage connected through the resistance R_L to the collector of the transistor.

On the network of output characteristics the normal operating state is first of all linked to the forward bias of the input junction. The characteristics associated with a negative value of the base current are the only valid ones; they are therefore located above the characteristic for $I_B = 0$ (yellow shaded area, Fig. 127). On the load line the normal state corresponds to the yellow section in Fig. 127.

Static equivalent circuits

As for the cut-off state it is interesting to define the corresponding static equivalent circuit.

With voltage drive let R_B be very nearly zero (Fig. 128); the static equivalent circuit of the transistor includes, in addition to the resistance r_{bb}', an emitter–base junction forward biased, a base–collector junction reverse biased, and a current generator connected to the terminals of the latter.

With current drive, that is with R_B much greater than zero (Fig. 129), and with the same components as above, we must add a second current generator in parallel with the reverse biased junction. The role of this generator is to represent quantitatively the influence of the residual current on the final value of the current.

These two static equivalent circuits of the transistor with the common-emitter configuration are very important; they are very often used for the study of the audio-frequency and radio-frequency applications of the transistor.

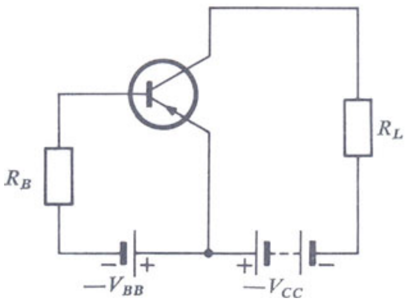


Fig. 126

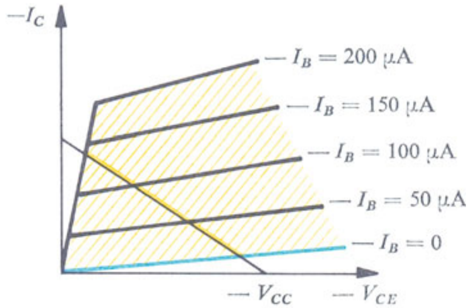


Fig. 127

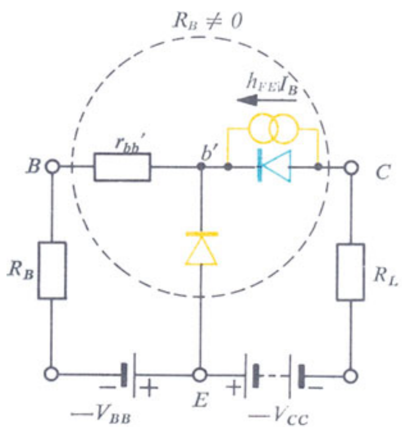


Fig. 128

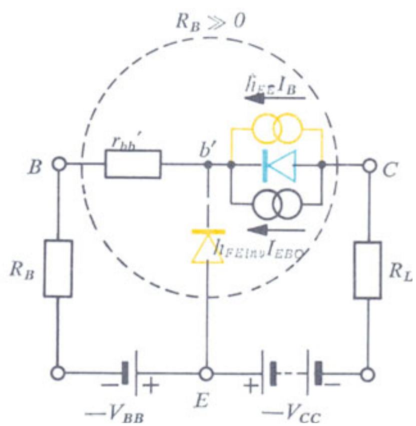


Fig. 129

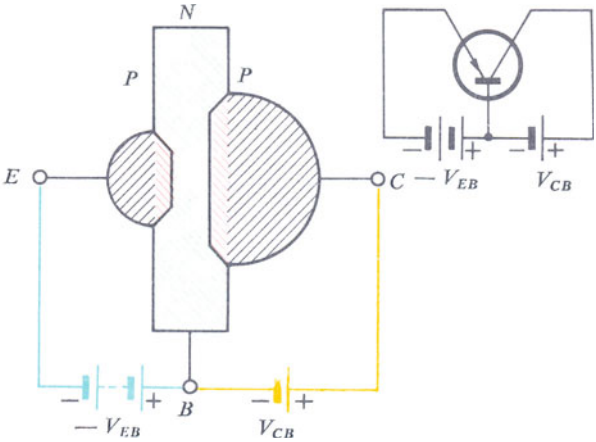


Fig. 130

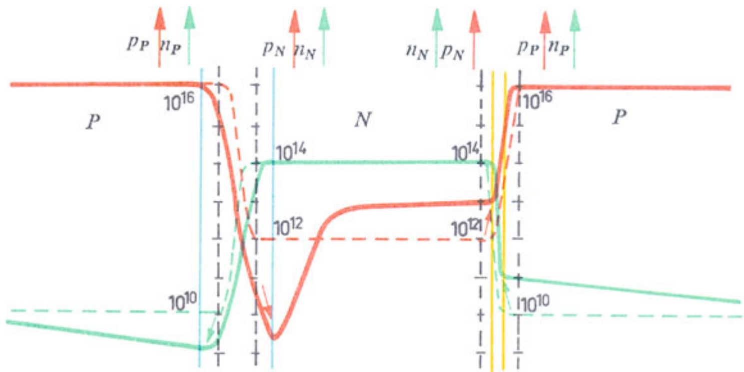


Fig. 131

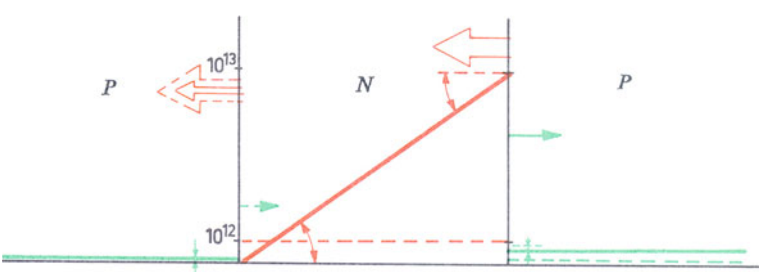


Fig. 132

2.5. Reverse operating state

The third type of operating state of the transistor consists in biasing the emitter–base junction in the reverse direction and the collector–base junction in the forward direction; this is the reverse operating state of the transistor.

All that is needed for this is to connect the negative pole of a group of batteries $-V_{EB}$ to the emitter side (Fig. 130), and at the same time to connect the positive pole of a battery V_{CB} to the collector side.

Biasing the collector–base junction in the forward direction causes a narrowing of the barrier; the holes then diffuse from the P region towards the N region, and the electrons from the N region towards the P region. The densities of the minority carriers increase and are located at a higher level (Fig. 131). The reverse bias of the emitter–base junction causes a widening of the corresponding barrier; only the minority carriers can now pass through, that is the holes from the N region towards the P region and the electrons from the P region towards the N region. In these conditions the minority carrier densities decrease in the neighbourhood of the new barrier.

Just as for the normal operating state, it is very difficult to define the currents quantitatively from such a group of graphs in which the densities of the majority and minority carriers overlap. It is therefore preferable to use a simplified density graph (Fig. 132) in which the ordinates are graduated linearly and in which only the minority carrier densities are shown.

The collector current is defined as the sum of the current of electrons from the N towards the P region on the right, the recombination currents in the N region and the current of holes in the P region (on the right) towards the N region. These different currents are functions:

(a) for the diffusion current of electrons, of the tangent of the angle of the density graph of the electron in the P region on the right;

(b) for the current of holes, of the tangent of the density graph of the holes in the N region on the collector side.

The emitter current is first of all linked to the number of holes which pass from the base into the emitter. This diffusion current depends on the tangent of the angle of the graph of the hole density in the N region (Fig. 132). As the two angles are equal, the hole currents are identical; the large disymmetry of the two junctions causes more numerous recombinations. In these conditions the emitter current is much lower than the collector current in the normal operating state.

Network of characteristics

With the common-emitter configuration, reverse operating state involves a reverse bias of the base–emitter junction and a forward bias of the collector–base junction. All that is needed to obtain this result is to connect the positive pole of a group of batteries to the base (with the negative pole connected to the emitter) and the positive pole of another group of batteries to the collector through the load resistance (with the negative pole connected to the emitter (Fig. 133)).

Representation of the output characteristics of a transistor operating in these conditions calls for a definition of a new network. In practice the reverse operating state corresponds only to the use of a transistor in the transient periods; the network of static characteristics in these conditions is only of slight interest.

Equivalent circuits

On the other hand, it is very useful to be able to represent the static equivalent circuit of a transistor. For this purpose we must differentiate between the two possible systems of driving, namely:

(a) voltage drive corresponding to a low value of the driving resistance (R_B not very different from zero or R_B smaller than r_e);

(b) current drive dependent on a high value of the driving resistance (R_B much higher than zero or R_B much higher than r_e). In these two expressions r_e represents the input resistance of the transistor.

With voltage drive, the equivalent circuit includes, in addition to the two junctions, (one of which is biased in the forward direction and the other in the reverse direction) and the resistance $r_{bb'}$, a current generator connected to the terminals of the base–emitter junction (Fig. 134).

With current drive, in addition to the components used previously, a second current generator must be connected to the terminals of the emitter–base junction in order to represent the residual current and the influence of the transistor effect on this current.

In these two equivalent circuits the 'principal' current generators (yellow in the figures) are quantitatively dependent on the concept of the reverse current amplification factor of the transistor.

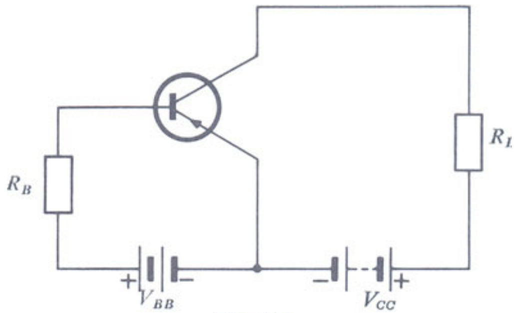


Fig. 133

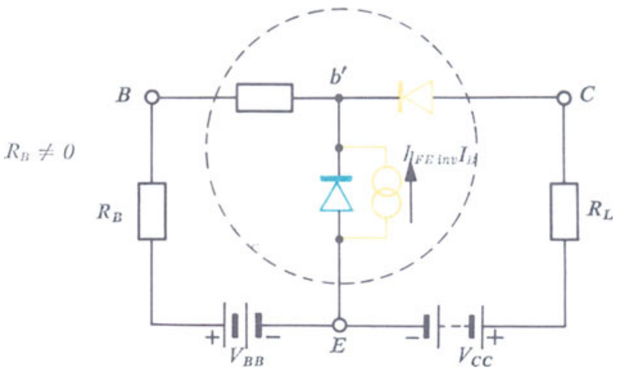


Fig. 134

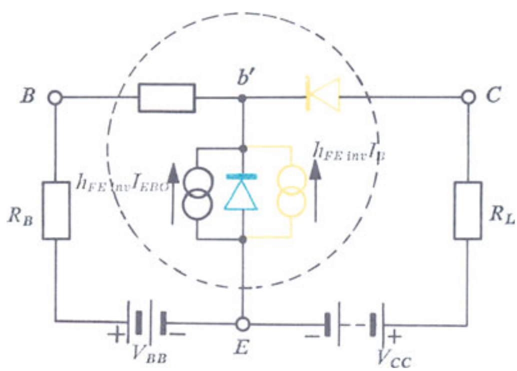


Fig. 135

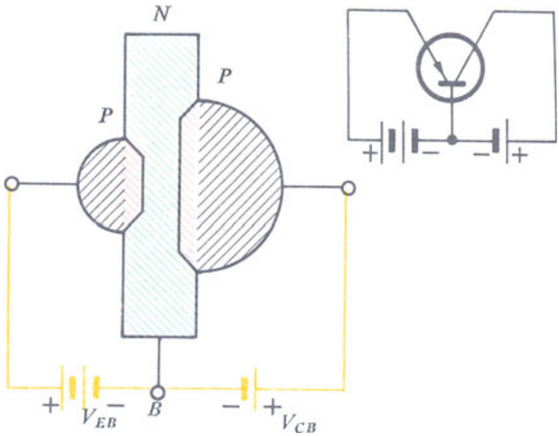


Fig. 136

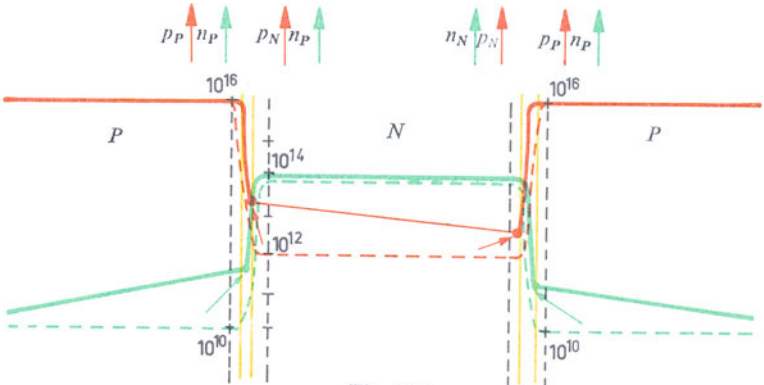


Fig. 137

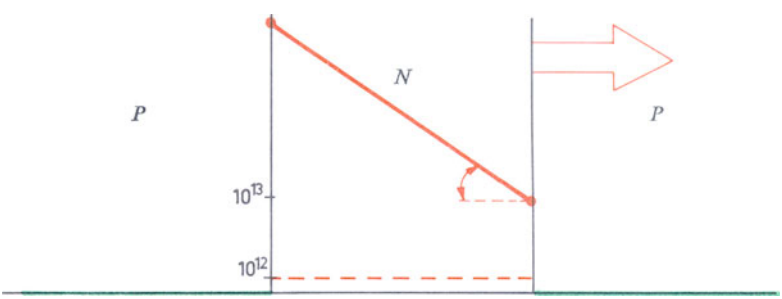


Fig. 138

2.6. Saturation state

The fourth operating state consists in having the transistor in saturation state. This requires the simultaneous forward biasing of the two junctions. This may be obtained by means of the following connections:

(a) the positive pole of a group of batteries (V_{EB}) emitter side, with the negative pole connected to the base;

(b) the positive pole of another battery (V_{CB}) collector side, with the negative pole connected to the base (Fig. 136).

Biasing the emitter–base junction in the forward direction causes a narrowing of the barrier. The holes, the majority carriers in the *P* region, diffuse into the *N* region; the same applies to the electrons from the *N* regions which diffuse towards the *P* region; the two minority carrier densities increase in the neighbourhood of the barrier, and become located at a higher level than their equilibrium value (Fig. 137).

The forward biasing of the right hand junction produces the same effects; the holes from the *P* region diffuse towards the *N* region and the electrons from the *N* region towards the *P* region. If the forward bias of the collector–base junction has been made lower than the bias of the other junction, the number of carriers passing through this barrier is lower, and the density gradients are less steep.

As for the other methods of operation, it is difficult to determine the currents passing through the three transistor electrodes by means of a graph of this kind. A considerable simplification can be made by using only the density graphs of the minority carriers, with the ordinates graduated linearly.

The various currents can be calculated with the aid of these graphs (Fig. 138). The forward bias of the two junctions causes an upward shift of the hole density graph. The angle between this graph and the horizontal enables us to calculate the diffusion current of holes from the *N* region towards the *P* region on the right. The electron currents on the emitter and collector sides are negligible. Any increase in the forward bias of the base–collector junction involves an upward shift of the right hand end of the density graph of holes in the base and a decrease in the current in the collector circuit of the transistor.

Common-emitter configuration

When a transistor is operated in the common-emitter configuration, the maintenance of the saturation state is ensured by biasing the base–emitter and collector–base junctions in the forward direction. This can be done by connecting the negative pole of a battery to the base and the positive pole to the emitter, whilst the negative pole of another group of batteries is connected to the collector through the resistance R_L (Fig. 139).

Networks of characteristics. By means of the network of output characteristics of the transistor (Fig. 140), it is useful to be able to determine the saturation region. The forward bias of the base–emitter junction corresponds to the characteristics located at a level higher than the characteristics for $I_B = 0$. The forward bias of the collector–base junction implies a less negative collector than the base (that is, a collector positive with respect to the base). On the load line any increase in the collector current (upward shift of the working point) corresponds to a decrease in the negative voltage applied to the collector with respect to the emitter. When the working point approaches the upper position the collector–emitter voltage decreases until the instant when it becomes practically equal to the base–emitter voltage. In these conditions any further decrease in the collector–emitter voltage causes a forward bias of the collector–base junction; the transistor is then in its saturation state. This saturation region of the transistor is located in the section of the Fig. 140 which is shaded in red.

Equivalent circuits. As with the earlier operating states, it is useful to define the static equivalent circuit of the transistor. This circuit includes the resistance r_{bb}' (given special values in this case) and the two forward-biased junctions. To each of these junctions corresponds a current generator, namely:

(a) the current generator $h_{FE} \cdot I_B$ connected to the terminals of the base–collector junction;

(b) the current generator $h_{FE\,inv} \cdot I_B$ connected to the terminals of the emitter–base junction.

The static equivalent circuit of the transistor does in fact contain all the information needed to establish the other circuits mentioned earlier.

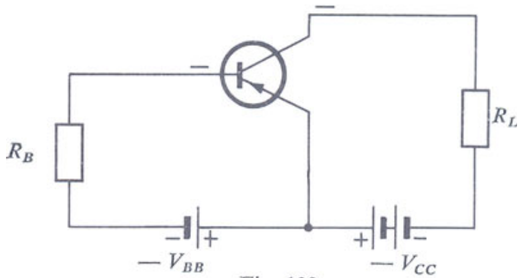


Fig. 139

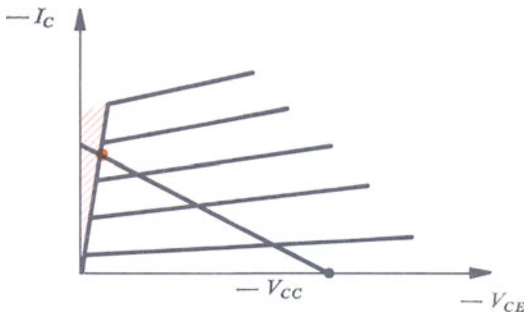


Fig. 140

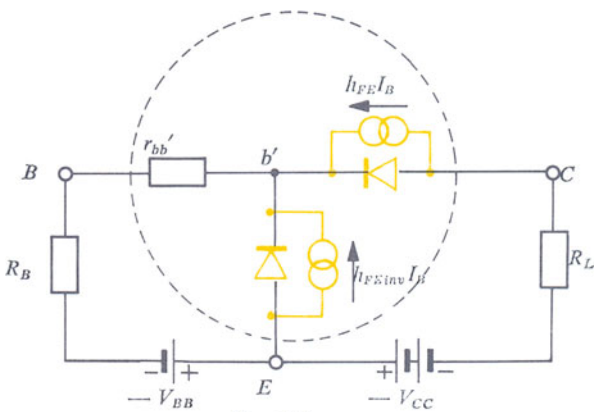


Fig. 141

'Amplification' function

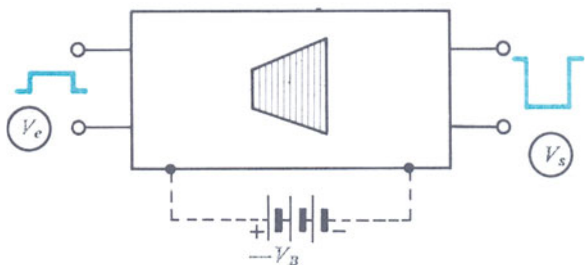
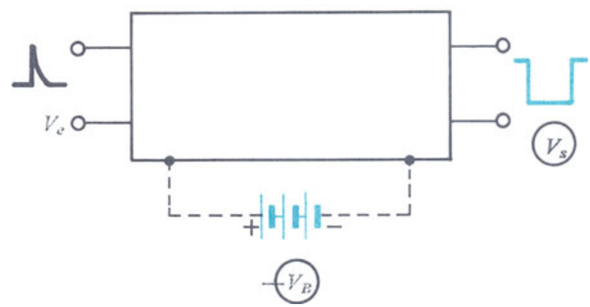


Fig. 142

'Switching' function



2.7. Operating principles of transistors with pulse drive

The study of the behaviour of transistors with pulse drive depends on the various conditions of using these devices in telecommunication, logic, and industrial electronic circuits.

Two main functions can be associated with this type of application:

(a) the 'amplification' function;

(b) the 'switching' function.

The amplification function is of outstanding importance, especially in television receivers. It consists in the supply, starting from a very small signal, of a signal of larger amplitude more suitable for the output stage (for example, the video-frequency signal for the picture tube should be of much larger amplitude than the voltage available at the output of the detector stage).

We can define the amplification function by comparing the signal available at the output of a stage or group of stages with the signal fed in at the input.

The power gain is determined by the quotient of the output power by the input power.

In practice the amount by which the output power exceeds the power supplied to the input is accounted for by the energy supplied by the group of batteries $-V_B$ feeding the system.

In dynamic conditions all that is needed is to compare the output and input signals in order to determine the properties of the amplifier stage. The diagram in Fig. 142 should make this quite clear.

A second possibility is to use transistors as switches. This is the 'switching' function. In this case the system supplies rectangular signals. The term 'generation' is however not suitable. This interrupter system borrows the available energy from a group of batteries and transforms it into rectangular signals. Other pulses are often injected into the input of such a device; their purpose is either:

(a) to determine the trigger frequency of the system;

(b) to affect the duration of the pulses.

Between these two main possibilities for the use of semiconductor devices there are important differences which we will discuss below.

2.8. Pulse amplification

In practice the amplifying function requires that all the variations applied at the input terminals of the system should be faithfully reproduced, whilst all that is required of the switching function is that the pulse applied at the input should be transferred from one level to another, whatever the amplitude of the pulse.

Figure 144 shows an amplifier stage equipped with a transistor in common emitter. This transistor is driven by a generator with a certain value of internal resistance (R_a), and simultaneously biased by a battery $-V_{BB}$. The output circuit is fed by a group of batteries ($-V_{CC}$); the load consists of resistance R_L.

We must now define the operation of a system of this type. For this purpose it is useful to know the different values of the voltage applied to the terminals of the input circuit, and also the shape of the signal which may be received at the terminals of the load resistance.

Two generators exist in fact in the input circuit:

(a) the battery $-V_{BB}$;

(b) the generator, which supplies a direct voltage equal to $-V_1$ from instant t_2 to instant t_4.

In order to determine the voltage variations effectively applied to the input terminals of the circuit we need do no more than to add at each instant the two voltages supplied by the battery and the generator; Figs. 145a, b, c represent respectively:

(a) the voltage supplied by the battery;

(b) the voltage supplied by the generator as a function of time;

(c) the voltage resulting from these two generators connected in series.

If we wish to know the shape of the output signal it is essential to define: on the one hand the variations with time of the output current (collector current), and on the other hand the voltage variations at the terminals of the load resistance. These two items of information are shown in Figs 146c and b.

From instant t_1 to instant t_2 the voltage supplied by the group of generators is equal to $-V_{BB}$, the collector current has its minimum value ($-I_{CO}$) and the collector–emitter voltage is a maximum ($-V_{RL}$).

From instant t_2 to instant t_4, ignoring the transient effects, the voltage supplied by the group of generators is equal to ($-V_1$) plus ($-V_{BB}$). In these conditions the current is a maximum ($-I_{C1}$) and the output voltage a minimum ($-V_{RL}$).

From instant t_4 the voltage supplied by the group of generators is again equal to $-V_{BB}$ (Fig. 145c). The current has returned to its original value (Fig. 146a) and the output voltage is a maximum and equal to $-V_{RL}$ (Fig. 146b).

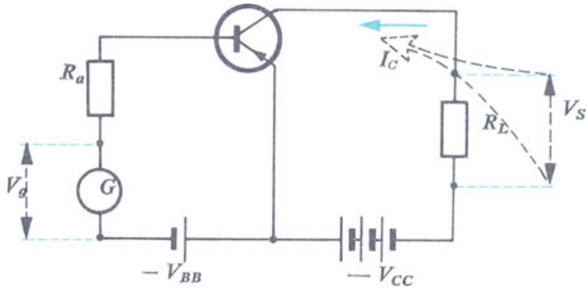


Fig. 144

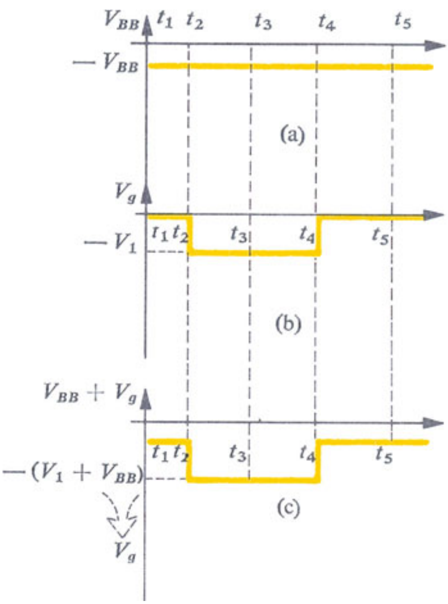


Fig. 145

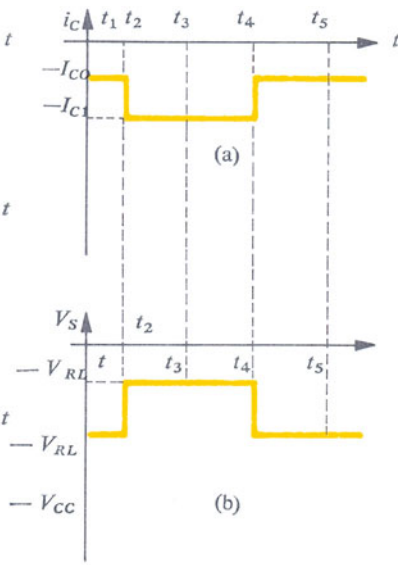


Fig. 146

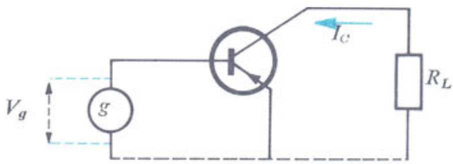


Fig. 147

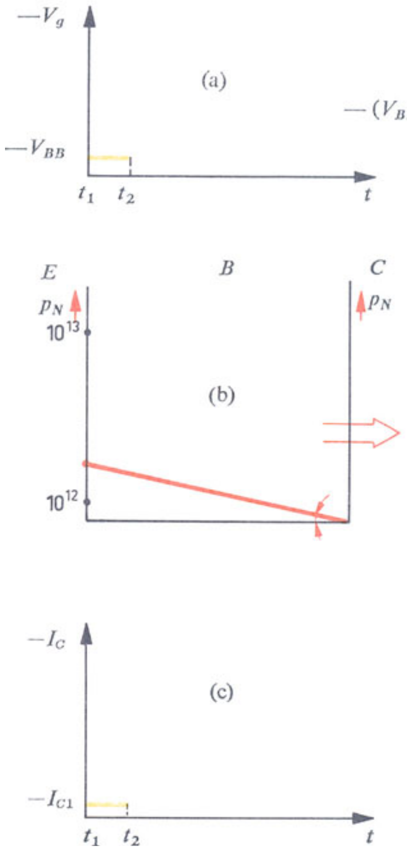


Fig. 148

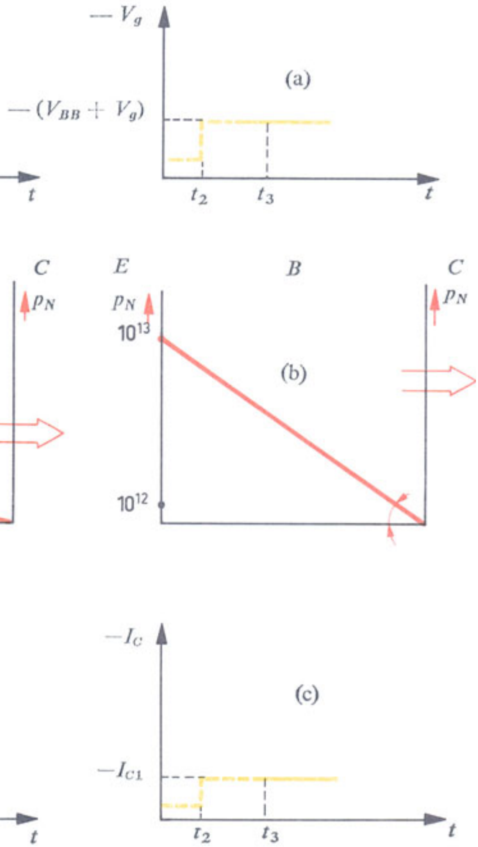


Fig. 149

Study of an amplifier stage with common-emitter configuration

The object of this section is to define the different shapes of the collector current and voltage appearing at the output terminals of a transistor driven by a generator supplying a rectangular signal.

The first part of the discussion will be limited to the behaviour of a transistor operating on a small-signal basis.

When no signal is being supplied by the generator, the transistor is already in the normal operating state; the application of a signal (generator V_G) results in a larger forward bias of the base–emitter junction.

Use of density graphs. Referring to Fig. 148a we see that from instant t_1 to the instant preceding t_2, the generator of rectangular signals does not supply any voltage. The voltage between the base and the emitter of the transistor represented by $-V_g$ is only equal to the voltage supplied by the bias battery $-V_{BB}$.

On the other hand the collector is strongly negative; the transistor is operating normally. The hole density graph is oblique, as shown in Fig. 148b. The hole diffusion current in the interior of the N region is then given by the tangent of the angle between this graph and the horizontal near the barrier on the collector side.

The corresponding value of the collector current $-I_{C1}$, shown in Fig. 148c, can be obtained from this current.

To an instant t_3 a long way from t_2, there is a new signal supplied by the two generators (V_B and V_G), Fig. 149a.

The increase in the forward bias of the base–emitter junction causes the density graph to shift upwards. The angle between this graph and the horizontal is larger and the hole current increases. The new value of the current in the collector circuit is shown at instant t_3 as equal to $-I_{C1}$ (Fig. 149).

The two limiting values of the collector current are directly dependent on the positions of the density graph of the holes in the base. The values of these two currents can be calculated only from the values of the angle between this graph and the abscissa in the neighbourhood of the base–collector junction.

Influence of the drive on the rise time of the current in the circuit. In all cases in which it has been necessary to determine the shape of the current in a circuit, account has been taken of the value of the internal resistance of the generator connected to the input of the system.

The two limiting operating states of a transistor are:

(a) voltage drive, that is for R_a very near to zero;

(b) current drive, that is for a very large value of R_a.

Voltage drive (Fig. 150a). At instant t_2, the voltage at the input terminals of the transistor changes abruptly from $-V_{BB}$ to $(-V_g)+(-V_{BB})$ (Fig. 150b).

As there is no resistance in the input circuit, the density graph of the holes in the base shifts instantaneously from position 1 to position 2 in Fig. 150c. The angle made by this graph with the horizontal (collector side) is only slightly influenced by the increase in density; in these conditions the increase in current is very small (position 2 in Fig. 150d).

The density graph of the holes in the base then assumes the different positions 3, 4 and 5 shown in Fig. 150c. Both the angle and current increase at the same rate. The variations of current in the circuit are shown in Fig. 150d, in which position 5 gives the final value (that is $-I_{C1}$). $t_2'-t_2$ gives the rise time of the current in the circuit.

Current drive (Fig. 151a). At instant t_2 the voltage increases abruptly at the input terminals of the transistor (Fig. 151b).

The barrier narrows; the holes can however no longer enter the interior of the base of the transistor 'in unlimited quantities'. The density graph shifts gradually from position 1 to position 2 (Fig. 151c).

To each of these positions corresponds a value of the angle made by the hole density graph near the barrier, on the collector side.

The collector current increases to its value $-I'_{C1}$ as a function of the increase of this angle (Fig. 151d). $t_2''-t_2$ gives the new value of the rise-time of the current in the circuit.

The rise-time of the collector current increases with the value of the driving resistance. It is therefore better to operate the transistor permanently with voltage drive when dealing with small amplitude signals.

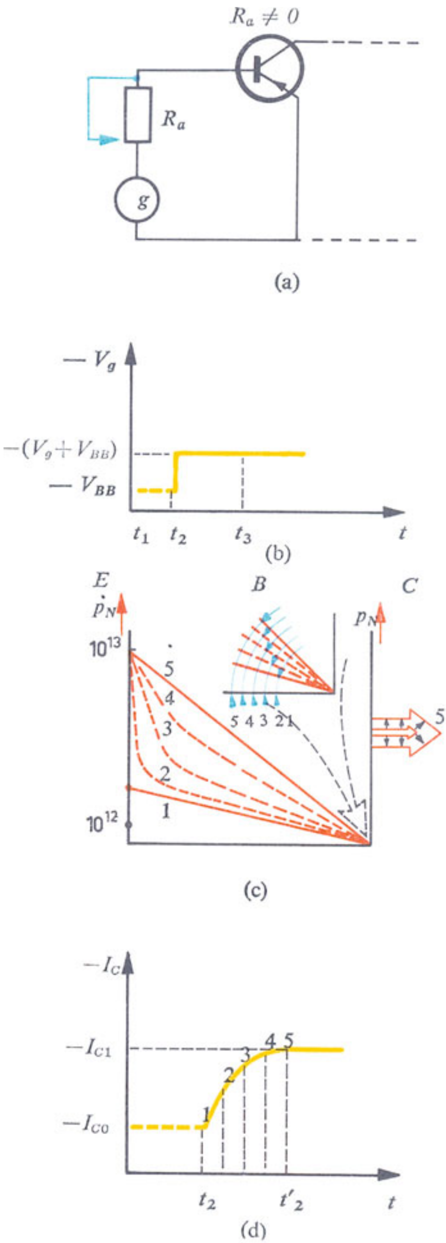


Fig. 150

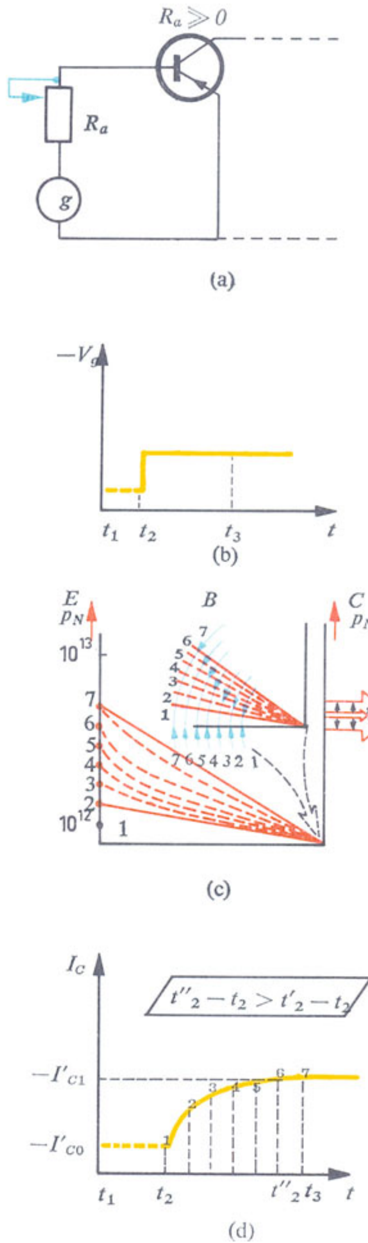


Fig. 151

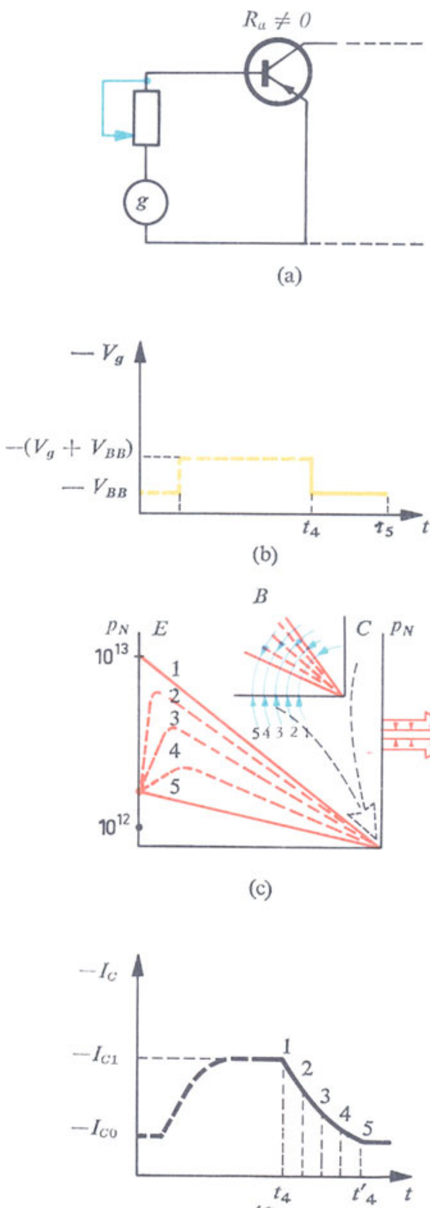


Fig. 152

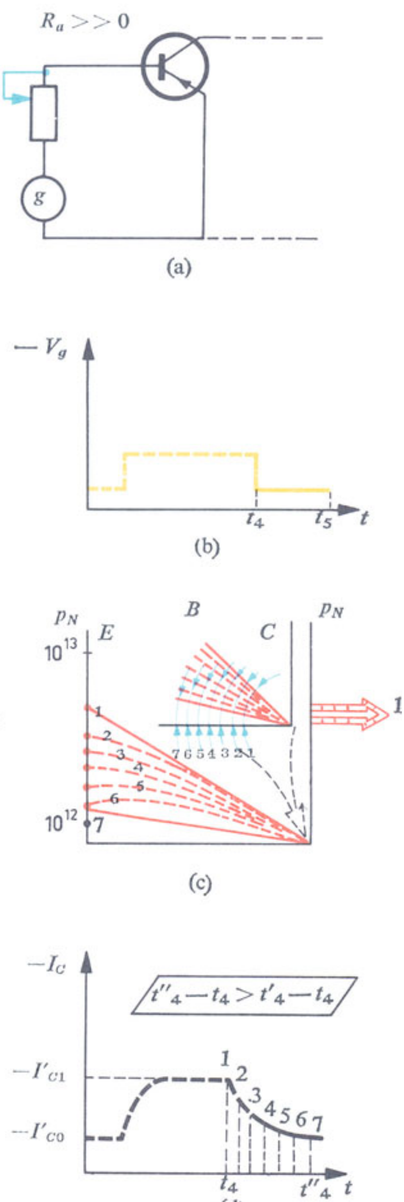


Fig. 153

Influence of the drive on the fall time of the current in the circuit. When the signal disappears the transistor reverts to the original bias conditions. Because of the importance of the type of drive on the shape of the current, we must now study the behaviour of the circuit, first with voltage drive and then with current drive.

Voltage drive. Figure 152a shows the transistor with voltage drive. At instant t_4 the signal supplied by the group of generators changes abruptly from the value $-(V_g + V_{BB})$ to the value $-V_{BB}$ (Fig. 152b).

Before instant t_4 the density graph was in the position 1 shown in Fig. 152c. At instant t_4 the density changes instantaneously, because of the absence of bias in the drive circuit, from the value 10^{13} to the value corresponding to the new bias of the input circuit. The angle between the hole density graph with the horizontal near the collector–base junction is only slightly affected by this variation. The collector current has decreased slightly (position 2 in Fig. 152d).

The density graph continues to swing until it reaches position 5. Gradually the angle gets smaller and the current decreases (positions 2, 3, 4 and 5 in Fig. 152d).

$t_4' - t_4$ gives the rate of decrease of the current in the circuit.

Current drive. With current drive (Fig. 153a), the same voltage variation (Fig. 153b) causes a much slower shift of the graph of the hole density in the interior of the N region.

In these conditions the angle between this graph and the horizontal decreases much more slowly and the current in the collector circuit decreases according to the various positions in Fig. 153d (1, 2, 3, 4, 5, 6, 7). $t_4'' - t_4$ gives the new value of the rate of decrease of current in the circuit.

Because of the slower removal of the holes from the base, this value is considerably increased with relation to the time given in the previous case.

$$t_4'' - t_4 > t_4' - t_4$$

This time is frequently designated in data sheets as 't_f', the fall time.

It is clear from the above discussions that the type of drive has a strong influence on the operation of a transistorized stage for the amplification of rectangular signals.

The two limiting types of drive are: on the one hand, voltage drive corresponding to a very low driving resistance (in practice this resistance is smaller than the input resistance); on the other hand, current drive, depending on a very high driving resistance, that is a drive resistance higher than the input resistance.

The amplifier shown in Fig. 154a is voltage driven; the internal resistance of the generator is in fact very near to zero. The variations in the voltage supplied by the group of generators, including the generator of rectangular signals and the battery, are shown in Fig. 154b. From instant t_1 to the instant preceding t_2 the voltage supplied by the group of generators is a minimum and the collector current is low ($-I_{CO}$ in Fig. 154c).

At instant t_2 the value of the voltage changes abruptly to $-V_g$, and the collector current starts to increase. It reaches in fact its final value at instant t_2', much later than instant t_2 (Fig. 154c). From instant t_2' to the instant preceding t_4 the voltage applied at the input terminals remains at $-V_g$; the current is maintained at the value $-I_{C1}$. At instant t_4' the signal reverts to its value $-i_{CO}$ at which it remains until the arrival of another pulse.

We will now consider the shape of the current in the output circuit of the transistor. A certain time $(t_2'-t_2)$ is required for the increase of this current from its minimum value $(-I_{CO})$ to its maximum value $(-I_{C1})$; similarly $t_4'-t_4$ gives the time taken for the current to change from its maximum to its minimum value. These two times give the rise time and the fall time respectively of the current in the collector circuit.

With current drive the times are much longer (Fig. 155c).

$t_2''-t_2$ gives the new value of the rise time of the current in the collector circuit and $t_4''-t_4$ the new fall time of the current:

$$t_2''-t_2 > t_2'-t_2 \quad \text{and} \quad t_4''-t_4 > t_4'-t_4$$

Quantitative expression of the triggering times. The rise time of the current (t_r) is given by the equation:

$$t_r = \tau_c \log_e I$$

τ_c represents the time constant of the rise time of the current. This approximate formula may also be used for the calculation of the fall time of the current t_f:

$$t_f = t_r.$$

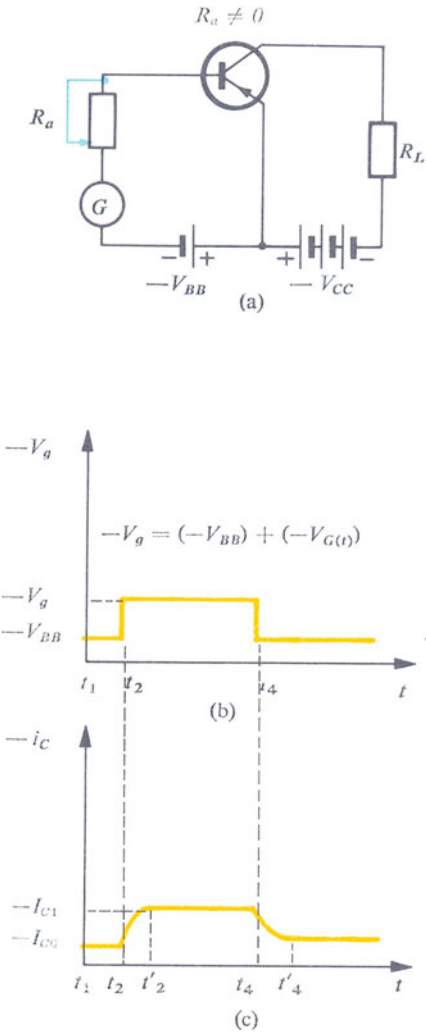


Fig. 154

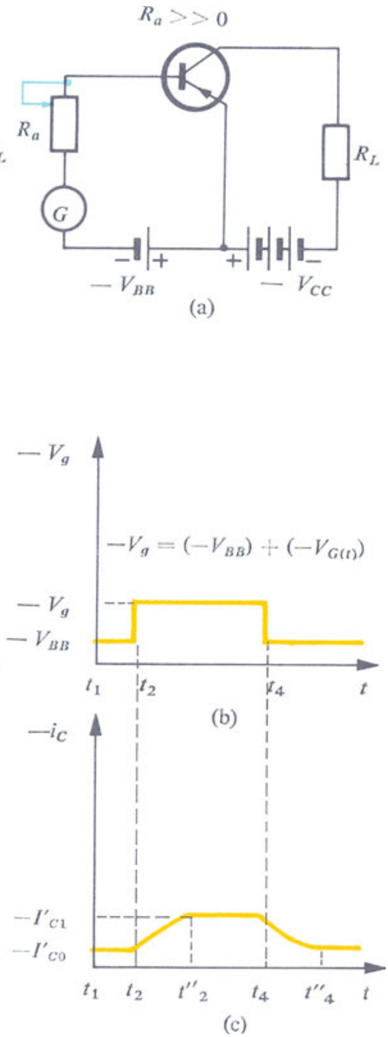


Fig. 155

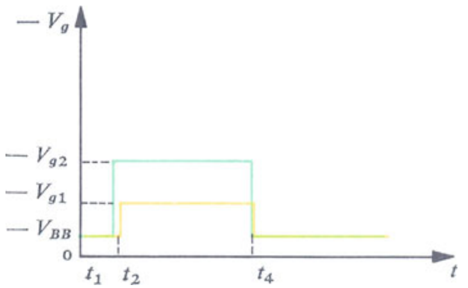


Fig. 156

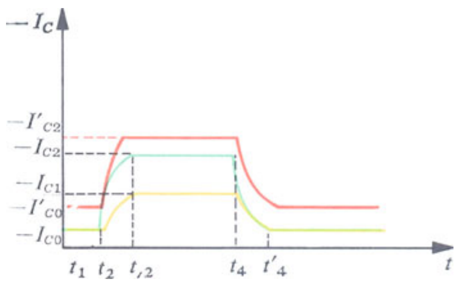


Fig. 157

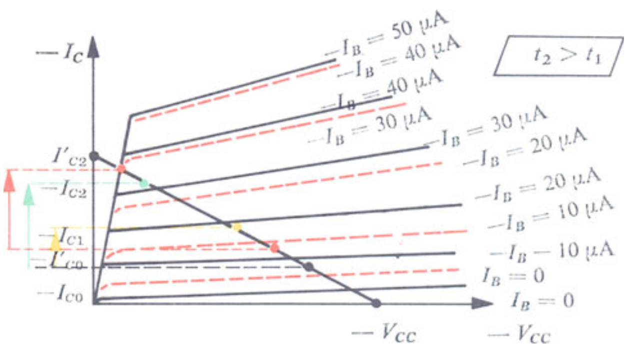


Fig. 158

2.9. Thermal instability of transistors

The study of the use of transistors in the audio or radio-frequency fields requires a specially important chapter devoted to thermal problems. Similarly, with pulse drive, it is very important to take account of instability problems and to emphasize the necessity of stabilizing the transistor against temperature variations.

We will assume two different operating conditions of the transistor used in the circuit on the previous page.

In the yellow example (Fig. 156) the variations in the voltage supplied by the generator cause the transistor to be permanently in the small-signal state. The corresponding variations in the collector current are shown in Fig. 157. The value of the current changes in fact from $-I_{CO}$ for a voltage $-V_{BB}$ to $-I_{C1}$ for a voltage $-V_{g1}$. In the network of output characteristics in Fig. 158, the working point on the load line shifts from the lowest yellow position to the highest yellow position.

Let us now assume a signal of larger amplitude, namely $-V_{g2}$, from instant t_2 to instant t_4 in Fig. 156. The corresponding variations in the collector current at a given temperature are shown in green in Fig. 157. The value of the current changes from $-I_{CO}$ to $-I_{C2}$.

On the network of output characteristics of the transistor in Fig. 158, the working point shifts from the lowest green position to the highest green position and the limiting values of the collector current are obtained by projecting these two points on to the axis $-I_C$.

For an increase in temperature (T_2 higher than T_1), a same voltage variation applied to the input terminals of the transistor causes a smaller change than before in the current in the output circuit:

$$I_{C2}'-I_{CO}' < I_{C2}-I_{CO}.$$

This is explained by the shift upwards of the network of characteristics due to the rise in temperature (the new characteristics shown in red dotted lines in Fig. 158). For the same variations in the input current the working point shifts from the 'minimum' position shown in red to the 'maximum' position also in red; the latter is responsible for the transistor entering the saturation state.

A second temperature rise causes only a new decrease in the amplitude of the collector current.

2.10. Improvement in the shape of the collector current

There are two disadvantages in using a stage of this sort as a switching element, namely:

(a) a decrease in the amplitude of the output signal due to thermal drift of the working point;

(b) a danger that the state of conduction will pass to the other state without affecting the input signal.

In addition to questions of instability, it is very important to have as rectangular a signal as possible at the output of a switch; in other words, the transistor should constitute a single or a group of perfect switches. How can the shape of the signal be improved in the light of the above discussion?

We will first consider the rise time of the current in the circuit for a signal of small amplitude at the input terminals of the circuit ($-V_{g1}$, Fig. 159). The time required for the value of the current to change from $-I_{CO}$ to $-I_{C1}$ is given by $t_2'-t_2$ in Fig. 160. This time is relatively long. The calculation of the limiting values of the current can be simplified by the use of network of characteristics in Fig. 161. The variation $(-I_{C1})-(-I_{CO})$ can be obtained by the projection of the two extreme positions of the yellow working point on to the load line.

Let a much stronger signal now be applied at the input terminals of the transistor, namely voltage $-V_{g2}$ at instant t_2 (Fig. 159).

The collector current approaches a much higher value ($-I_{C2}$ in Fig. 160). The time interval required for the value of the current to change from $-I_{CO}$ to I_{C2} is practically the same as the time applying in the previous case; it is therefore still equal to $t_2'-t_2$. On the network of output and transfer characteristics in Fig. 161 the current $-I_{C2}$ is located well above the saturation current $-I_{CS}$.

Improvement in the rise time of the current

The collector current increases up to its limiting value $-I_{CS}$. This value involves the passing of the transistor into the saturation state. In these conditions the current cannot rise any further; $t_2''-t_2$ gives the new value of the rise time of the current in the collector circuit, which is now very small.

$$t_2''-t_2 < t_2'-t_2$$

Finally, a definite reduction in the rise time of the collector current is obtained when the transistor becomes very strongly saturated.

Explanatory diagram of the improvement in the rise time. In addition to the use of the network of output and current transfer characteristics, it is interesting to use the density graphs to explain the reduction in the rise time of the current when the transistor is abruptly put into the saturation state.

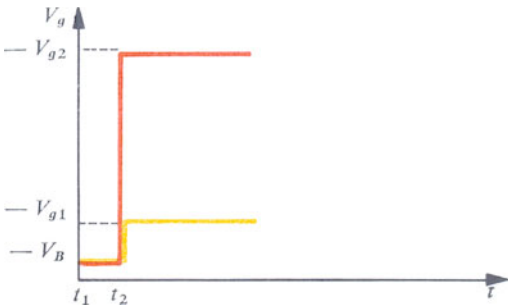


Fig. 159

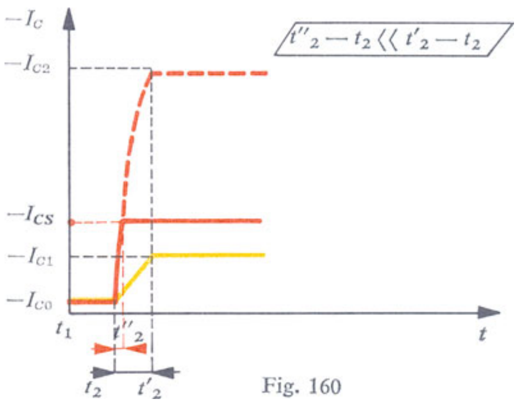


Fig. 160

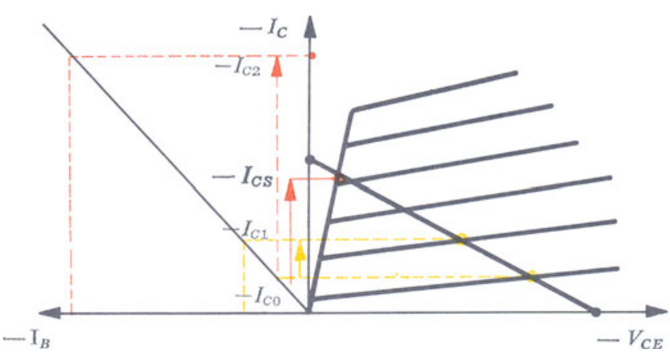


Fig. 161

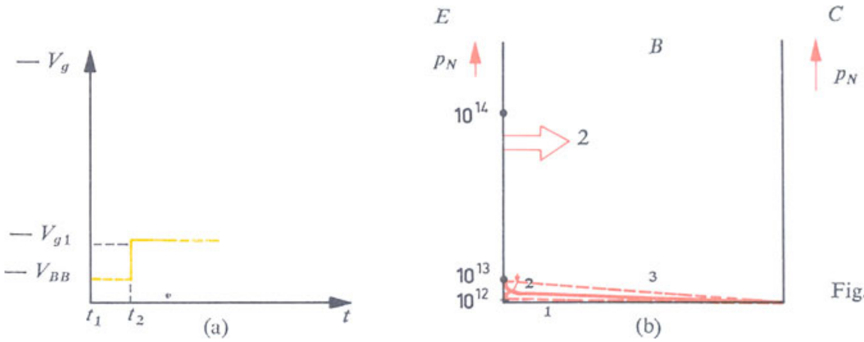


Fig. 162

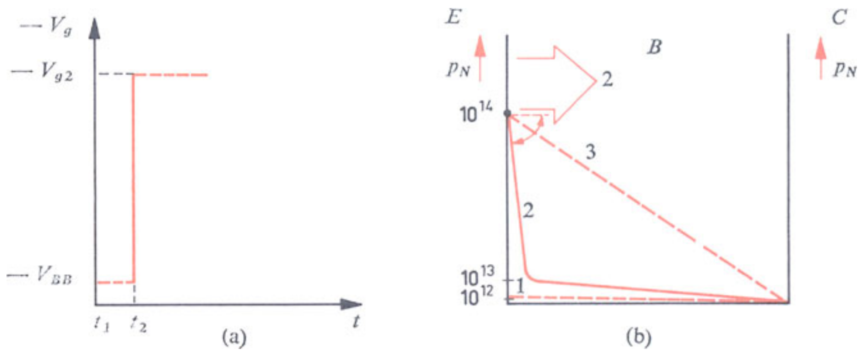


Fig. 163

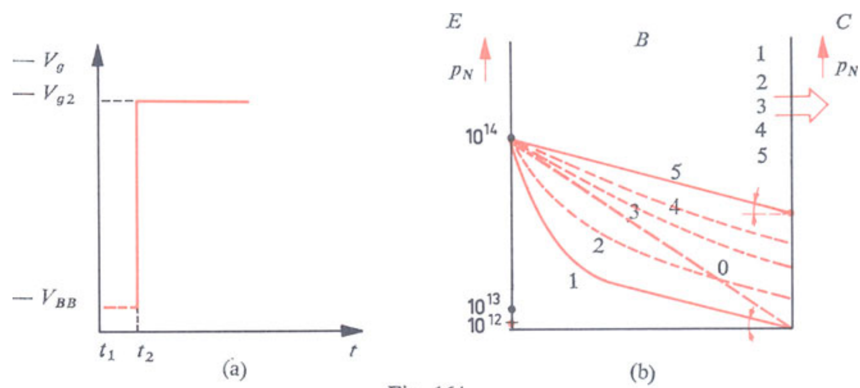


Fig. 164

Small-signal drive. With small-signal operation (Fig. 162a) the variation in the voltage applied at the terminals of the emitter–base junction causes only a small decrease in the width of the barrier, which is accompanied by relatively small flow of holes from the emitter to the base. In Fig. 162b, which shows the density graph of the holes in the interior of the base, we see four successive stages:

(a) stage 1, that is from instant t_1 to the instant preceding t_2: the density graph is slightly oblique and the current in the collector circuit is very small;

(b) stage 2, that is instant t_2 in Fig. 162a; the barrier becomes narrower, the holes penetrate into the base in relatively small numbers and the density graph is slightly oblique; the current in the collector circuit is still small;

(c) stage 3, that is instant t_2 in Fig. 162a; the barrier narrows, the holes penetrate into the base in relatively small numbers and the density graph is in a more pronounced diagonal position (2 in Fig. 162b);

(d) stage 4, the graph is in its final position (3 in the figure); the collector current has reached its final value.

Large-signal drive. With large-signal operation, namely with a voltage $-V_{g2}$ applied across the input of the transistor (Fig. 163a), the transistor becomes saturated. We will ignore for the moment this saturation state and observe the different positions of density graph of the holes in the base.

In Fig. 163b we see that:

(a) in position 1 the transistor is weakly biased, the graph is slightly oblique and the current is low;

(b) in position 2, corresponding to instant t_2, the holes penetrate in very large numbers. The diffusion current of holes into the base is then associated with a very large angle of the graph of the hole density in the neighbourhood of the barrier, on the emitter side;

(c) in stage 3 the density graph is in its final position and the collector current is at its theoretical value $-I_{C2}$.

When the drive voltage changes from the value $-V_{BB}$ to the value $-V_{g2}$, the density graph immediately takes up the position 1 shown in Fig. 164b. The transistor then reaches saturation. The collector–base junction is then biased in the forward direction and the angle on the collector side remains practically constant.

Physical explanation of the reduction in the rise time. It is possible to add to our previous discussions, either from the networks of output and current transfer characteristics, or with the aid of the graph of the hole density in the base, a physical explanation enabling us to understand the influence of the saturation of the transistor on the rise time of the current.

When the signal is only equal to $-V_{BB}$ the base–emitter junction is weakly biased in the forward direction and the base–collector junction strongly biased in the reverse direction. This condition is represented by a relatively thick (yellow) junction between the emitter and the base in Fig. 165a, and very thick (blue) junction between the base and the collector.

In the network of output characteristics of the transistor (Fig. 165) the working point is very low down on the load line; it gives a value of collector current equal to $-I_{CO}$.

At the instant between t_2 and t_2', the base–emitter junction is slightly more strongly biased in the forward direction (decrease in the width of the yellow line between the emitter and base in Fig. 166a). The collector current consequently increases, and the collector–base junction becomes less strongly biased in the reverse direction. This condition is shown in Fig. 166b.

The value of the collector current rises to $-I_{C1}$ which is higher than $-I_{CO}$ and the collector–emitter voltage $-V_{CE}$ decreases.

If the collector current continues to increase, the base–collector junction in its turn becomes biased in the forward direction.

The working point shifts upwards on the load line in Fig. 167b; any increase in the collector current is accompanied simultaneously by a decrease in the collector–emitter voltage.

At a certain moment the collector becomes less negative than the base with respect to the emitter. The collector–base junction is then biased in the forward direction and the holes diffuse from the emitter and collector towards the base of the transistor. The simultaneous forward bias of the two junctions is represented in Fig. 167a by a very narrow yellow barrier between the emitter and base, and by a slightly thicker yellow barrier between the collector and base. A slight increase in the thickness of the base collector junction is associated with a weaker bias of this junction.

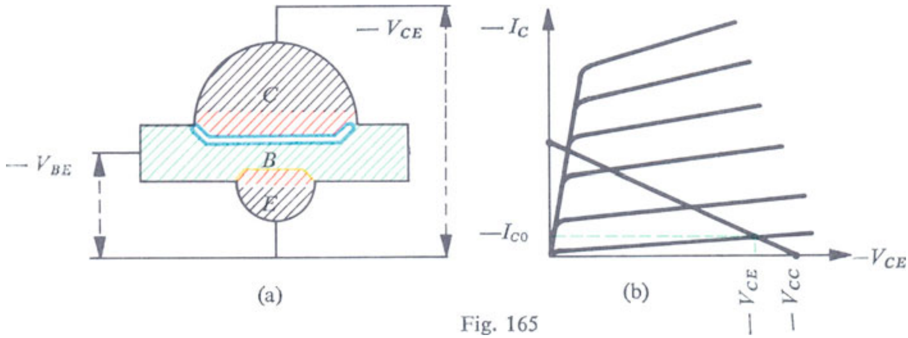


Fig. 165

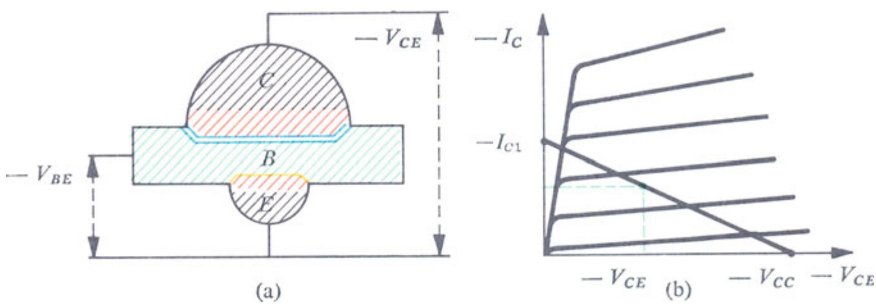


Fig. 166

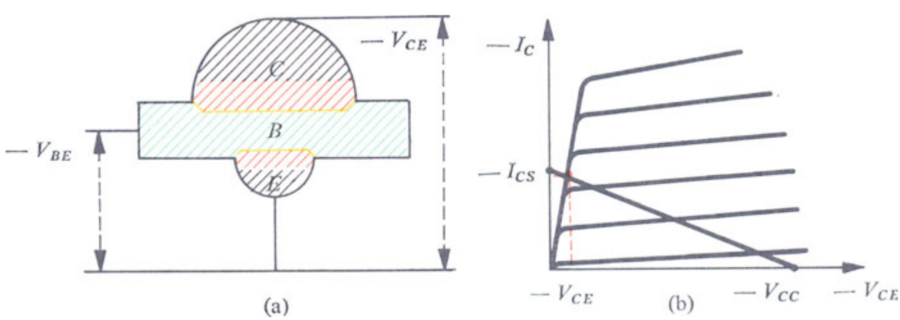


Fig. 167

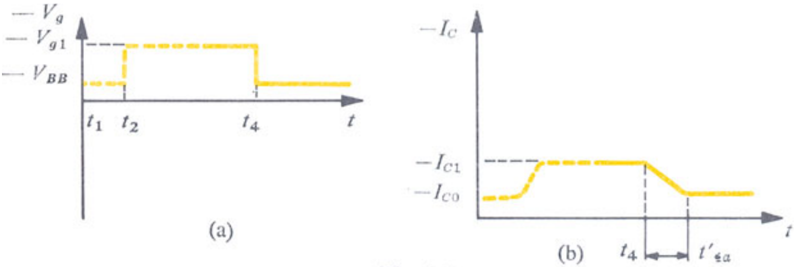


Fig. 168

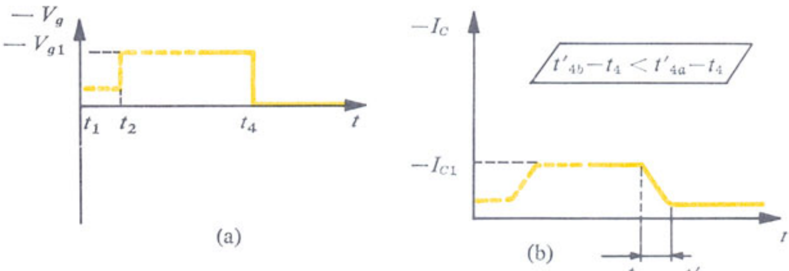


Fig. 169

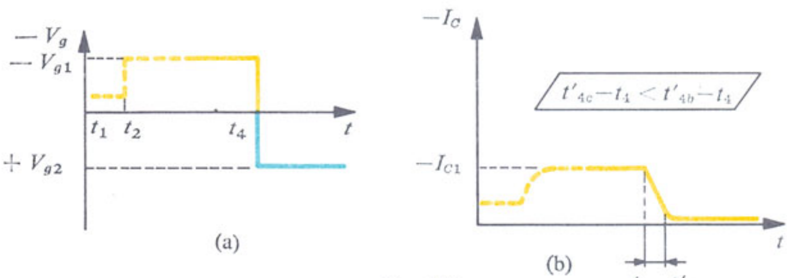


Fig. 170

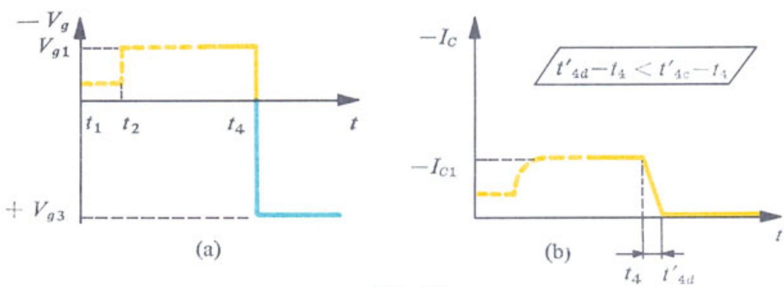


Fig. 171

Improvement in the fall time of the collector current

Now that we have dealt with the possibility of shortening the rise time of the current in the circuit, we can ask ourselves whether it is possible in the same way to improve the fall time of the current in the collector circuit.

Transition from one normal state to another normal state. In small signal conditions the variation in the input voltage at instant t_4 (Fig. 168) is such that the value of the collector current changes from $-I_{C1}$ to $-I_{CO}$ (Fig. 168). The time required for the collector current to reach the value $-I_{CO}$ is given by $t'_{4a}-t_4$ in this figure.

Transition from a normal state to the absence of bias. Let us assume that at instant t_4 there is a larger variation of voltage, namely from $-V_{g1}$ to 0 (Fig. 169a). In these conditions the collector current decreases and reaches a value lower than that shown before (Fig. 169b).

The period required for the value of the current to change from the maximum value to the minimum value is now equal to $t'_{4b}-t_4$.

The fall time of the current (t_4) is consequently reduced.

$$t'_{4b}-t_4 < t'_{4a}-t_4$$

Transition from the normal state to a cut-off state. If at instant t_4 a large current variation in the reverse direction is applied (Fig. 170a), the value of the collector current changes from its maximum value $-I_{C1}$ to a minimum value very near 0; this change takes place in a much shorter time, as shown in Fig. 170b.

The fact of having biased the input junction in the reverse direction at the moment of interruption of the current produces a definite improvement in the fall time of the current in the circuit.

Transition from the normal state to a strongly cut-off state. Let us assume at instant t_4 a still larger voltage variation V_{g3} (Fig. 171a). In these conditions the value of the current at instant t_4 changes from $-I_{C1}$ to zero.

The fall time of the current, now very short, is given by

$$t'_{4d}-t_4 < t'_{4c}-t_4 < t'_{4b}-t_4 < t'_{4a}-t_4$$

in Fig. 171b. A reverse bias of the input junction of the transistor, at the moment of the interruption of the current, causes a definite improvement in the fall time of the current in the collector circuit.

Explanation of the effect

We will now try to explain the reasons why the collector current decreases more rapidly when the transistor is switched from the normal state to a strongly cut-off state, compared with switching from one normal state to another normal state.

Weak switching. We will first consider the case of the small-signal operating condition. The value of the voltage supplied by the generator changes abruptly at instant t_4 from $-V_{g1}$ to $-V_{BB}$ (Fig. 172a).

When the drive voltage is equal to $-V_{g1}$, the density graph of the holes in the base is in position 1 (Fig. 172b); the limits of the barrier separating the emitter from the base are shown by yellow dotted lines. At instant t_4' when all the holes have been removed from the base, the density graph is in position 2 and the barrier assumes the position shown in full yellow lines. The time required for the removal of the surplus carriers in the base then gives the fall time of the current in the circuit (that is $t'_{4a}-t_4$ in Fig. 172c). The red shaded area in Fig. 172b occurs quantitatively in the fall time of the current in the circuit.

Strong switching. We will now consider the case when the transistor is strongly switched to the cut-off state. Before instant t_4 a voltage $-V_{g1}$ was applied between the base and emitter of the transistor (Fig. 173a). The density graph of the holes in the base was in position 1 (Fig. 173b) and the barrier is shown in this figure by broken yellow lines. At instant t_4 a strong reverse voltage pulse is applied between the base and emitter.

At instant t'_{4d} the density graph is in position 2 in Fig. 173b; due to the very strong bias of the junction the barrier becomes very wide.

The barrier widens instantaneously and all the holes in the region between the shaded yellow line and the blue line well on the right are immediately removed due to the influence of the electric field appearing in the space charge regions. The decrease in the collector current is completely dependent on the removal of the charges contained in the red shaded area demarcating the new thickness of the base of the transistor.

This region is now very thin (Fig. 173b) and therefore the fall time of the current in the circuit decreases (Fig. 173c).

We have just discussed the possibility of improving the rise time and fall time of the collector current of a transistor.

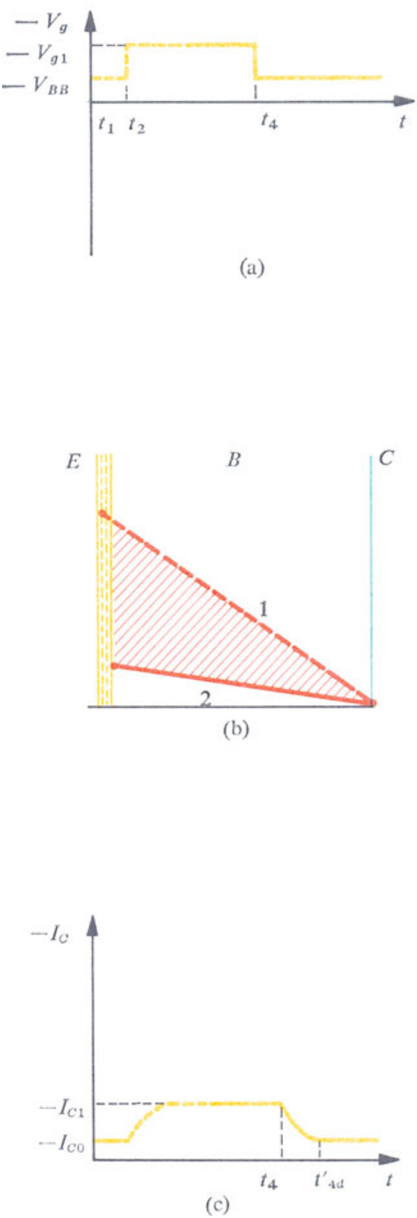


Fig. 172

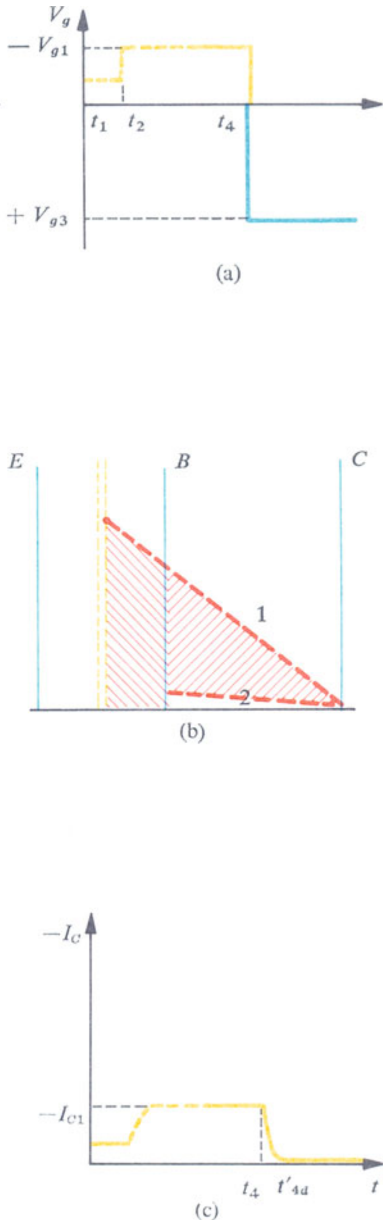


Fig. 173

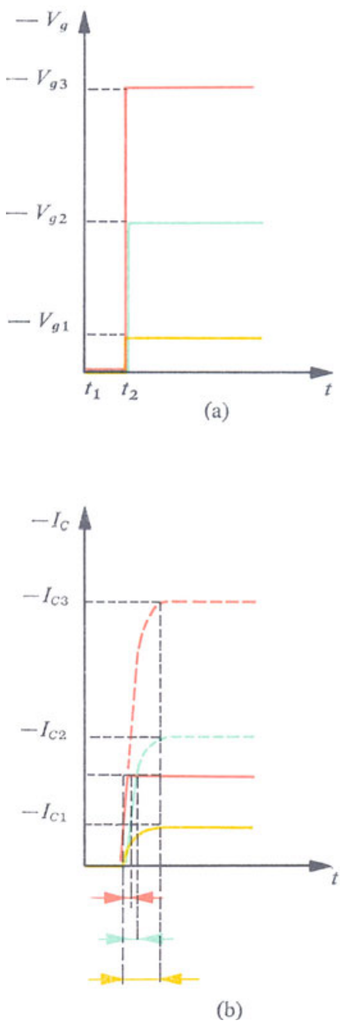


Fig. 174

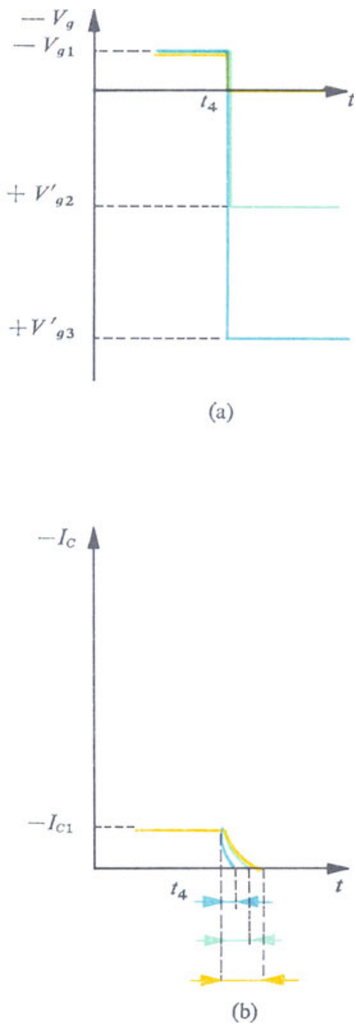


Fig. 175

Improvement in the rise time of the current

The rise time of the current is influenced by the amplitude of the signal applied to the input terminals of the transistor. Figure 174a shows three variations of the input voltage $-V_{g1}$, $-V_{g2}$ and $-V_{g3}$.

In the yellow example, i.e. when the voltage changes from zero to $-V_{g1}$, the collector current takes a certain time to increase from zero to the value $-I_{C1}$; this time is shown in yellow on the abscissa of Fig. 174b.

If we now increase the applied voltage, that is $-V_{g2}$ at instant t_2 (in green in Fig. 174a), the collector current tends to rise from zero to the value $-I_{C2}$ in the same time as that given previously; since the current is limited to its saturation value, the new rise time is represented by the green arrow, parallel to the abscissa in Fig. 174b.

For a still larger signal, namely from 0 to $-V_g$ (in red in Fig. 174a), the time taken for the current to rise from zero to $-I_{C3}$ is still shorter (red arrow parallel to abscissa in Fig. 174b).

Improvement in the fall time of the current

We will now consider the three possibilities shown in Fig. 175.

If at instant t_4 the voltage supplied by the generator falls from $-V_{g1}$ to zero, the fall time of the current in the circuit is given by the yellow arrow, parallel to the abscissa, in Fig. 175b.

Conversely, if at instant t_4, the voltage changes suddenly from $-V_{g1}$ to $+V_{g2}$, the fall time of the current shortens; it is now given by the green arrow in Fig. 175b.

A higher reverse voltage, namely $+V_{g3}$ at instant t_4, corresponds to a new reduction in the fall time (blue arrow in Fig. 175b).

Finally, in order to have available at the output of a transistorized switching stage a signal with as rectangular a shape as possible, the transistor must be switched from a very pronounced cut-off state to a very pronounced saturation state and vice versa.

2.11. Definition of the different switching times of a transistor

The use of a transistor as a switching component in an electric circuit consists, therefore, in switching it from a cut-off state to a saturation state and vice versa. Figure 176 illustrates this. In this circuit the transistor is driven by a group of generators consisting of a generator of rectangular signals in series with a battery whose positive pole is connected to the base; since the internal resistance of the generator is very low the transistor is voltage driven.

The voltage variations at the input terminals of the transistor are shown in Fig. 177a.

The collector is fed by a group of batteries $-V_{CC}$ through the load resistance R_L.

The explanation of the operation of this system consists quite simply, once we know the variations of the voltage applied to the input terminals, in the determination of the current variations in the collector circuit.

From instant t_1 to the instant preceding t_2, the generator G supplies no signal, and it is only the battery that biases the base–emitter junction in the reverse direction. Further, the collector is negative with respect to the emitter and the transistor is suddenly made nonconducting. A practically zero collector current is associated with this state (Fig. 177b).

At instant t_2, the generator G supplies a large negative voltage variation; the resultant voltage between base and emitter tends to place the transistor in a strong saturation state. Will the collector current immediately start to rise?

It does in fact remain at practically zero until instant t_2'. A delay time appears in the rise of the collector current, namely:

$$t_2' - t_2 = t_d$$

After t_2' the current rises and very quickly reaches the saturation value $-I_{CS}$. In Fig. 177b $t_2'' - t_2'$ gives the rise time of the current in the circuit (t_r).

From instant t_2'' until instant t_4 the current remains at its saturation value.

At instant t_4 the voltage at the input terminals of the transistor changes abruptly from a negative value to a very strong positive value. The collector current remains at the value $-I_{CS}$ until instant t_4'.

This new delay is known as the 'delay time of the fall of current' or desaturation time, namely: $t_4' - t_4 = t_s$.

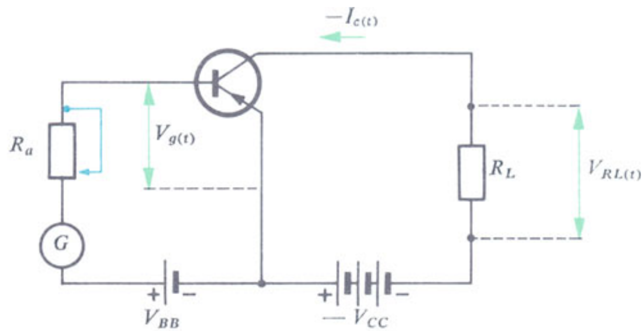


Fig. 176

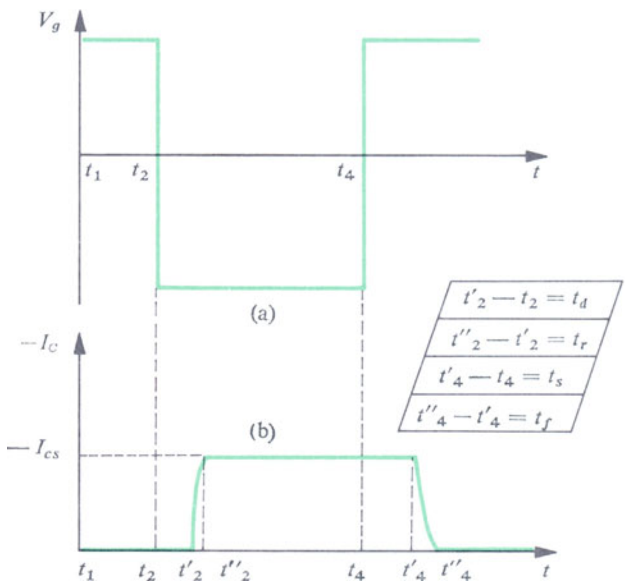


Fig. 177

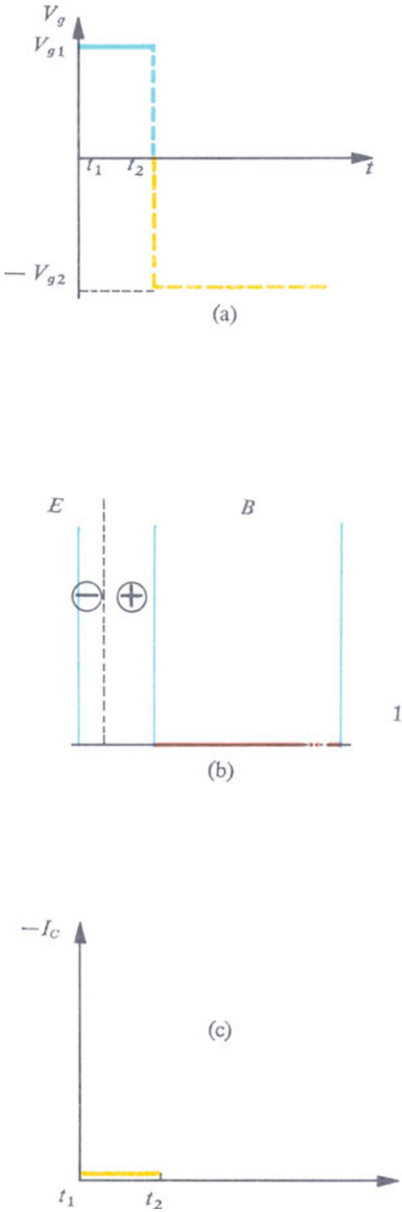


Fig. 178

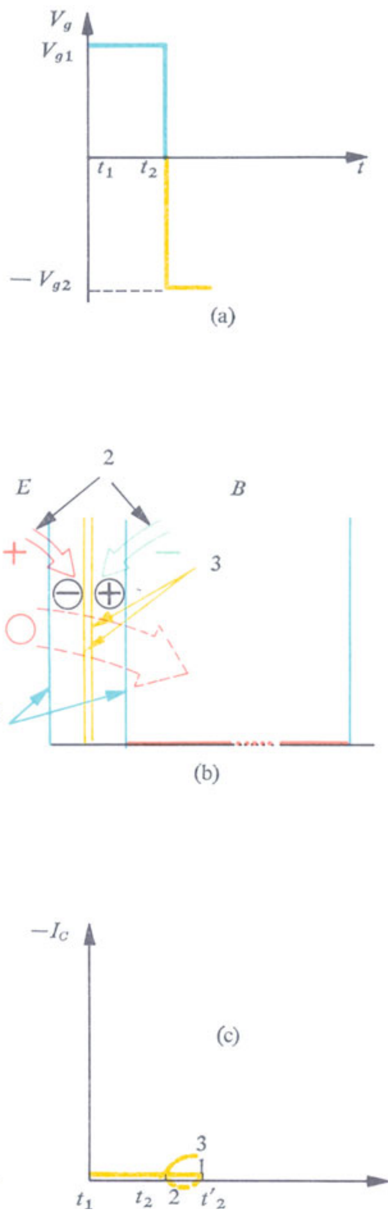


Fig. 179

Physical explanation of the switching time

We must now pass to an examination of the various time intervals discussed above, and for that purpose we will first consider the behaviour of the transistor in the cut-off state.

From instant t_1 to the instant preceding t_2 the voltage supplied by the group of generators is positive and equal to V_{g1} (Fig. 178a). The transistor is in the cut-off state. The density graph of the holes in the base is horizontal in the neighbourhood of the abscissa (Fig. 178b); the two barriers separating the emitter from the base and the base from the collector are very wide. They consist of two space charge regions, the one negative on the P region side, and the other positive on the N region side. Because of the position of the density graph of the holes in the base the current in the collector circuit is nearly zero.

Causes of the delay time of the current (t_d). At the instant t_2 a strong voltage variation tends to cause the transistor to pass from the cut-off state to a strong saturation state (voltage $-V_{g2}$ in Fig. 179a).

In order that the holes may diffuse from the emitter into the base and so cause the density graph to swing, it is essential that the barrier between the emitter and base should narrow. This narrowing is not instantaneous; it takes place in accordance with the different stages shown in Fig. 179b:

(a) stage 1, the transistor is in the cut-off state, and the barrier is very wide;

(b) stage 2, the forward bias is applied; electrons leave the base and neutralize part of the positive space charge region on the right; holes leave the emitter and neutralize part of the negative space charge region;

(c) stage 3, the arrival of positive and negative charges in the space charge regions cause a narrowing of the barrier, and holes start to diffuse into the base.

The time required for the reduction in the width of the barrier gives the delay time of the collector current t_d. This time is given by $t_2'-t_2$ in Fig. 179c. The narrowing of the barrier can have no influence on the collector current; on the other hand for certain types of transistors or circuits, current variations (marked 2 or 3 in Fig. 179c) appear in the output circuit.

The delay time of the current in the collector circuit is exclusively dependent on the width of the emitter–base junction, and thus on the cut-off state of the transistor.

Significance of the rise time of the current (t_2'). Since from instant t_2 the emitter–base barrier is narrower, the holes start to diffuse from the emitter into the base, and the density graph swings (refer to earlier explanations).

The voltage supplied by the group of generators is equal to V_{g2} well after instant t_2' and corresponds to the initiation of the saturation state of the transistor (Fig. 180a).

From instant t_2' the holes penetrate into the base and cause the density graph to move from position 1 to position 5 (that is 1, 2, 3, 4 and 5 in Fig. 180b).

When the graph moves upwards the barrier between the base and collector becomes even narrower. In fact, any increase in the collector current causes an increase in the voltage drop across the load resistance and causes the collector to become less and less negative with respect to the base. At the instant corresponding to position 5 in Fig. 180b the base–collector junction is no longer biased.

During the whole time during which the density graph is being established, the angle between it and the horizontal increases and the collector current rises until it reaches the saturation value. $t_2'' - t_2'$ gives the rise time of the current in the circuit (t_r) shown in Fig. 180c.

After instant t_2'' the signal supplied by the generator remains equal to $-V_{g2}$ and the transistor is still influenced by a voltage tending to place it in the saturation state (Fig. 181a).

The base–collector junction is now biased in the forward direction and the density graph moves parallel to position 1 to its final position 5 corresponding to the bias condition chosen. The change from position 1 to position 5 takes place through stages 2, 3 and 4 in Fig. 181b. During this whole period, the angle between the density graph and the horizontal is constant and the collector current is equal to the saturation value ($-I_{CS}$ in Fig. 181c).

The collector current reaches its saturation value ($-I_{CS}$) at the instant t_2'' and remains at this value as long as the voltage supplied by the group of generators does not vary ($-V_{g2}$ in Fig. 181a).

An increase in the applied voltage at instant t_2 will cause an upward movement of the density graph of the holes in the N region, but this increase in the saturation will, on the other hand, not cause any change in the final value of the collector current.

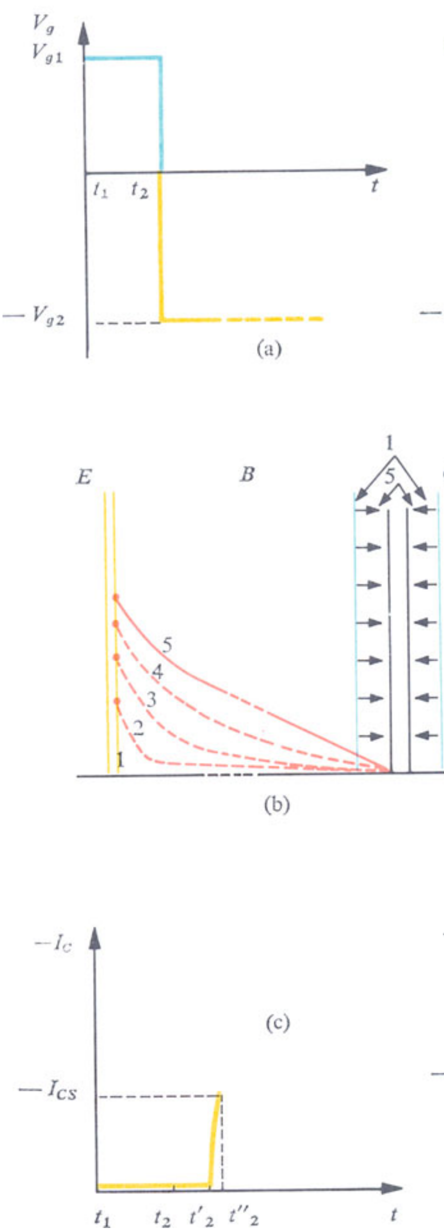


Fig. 180

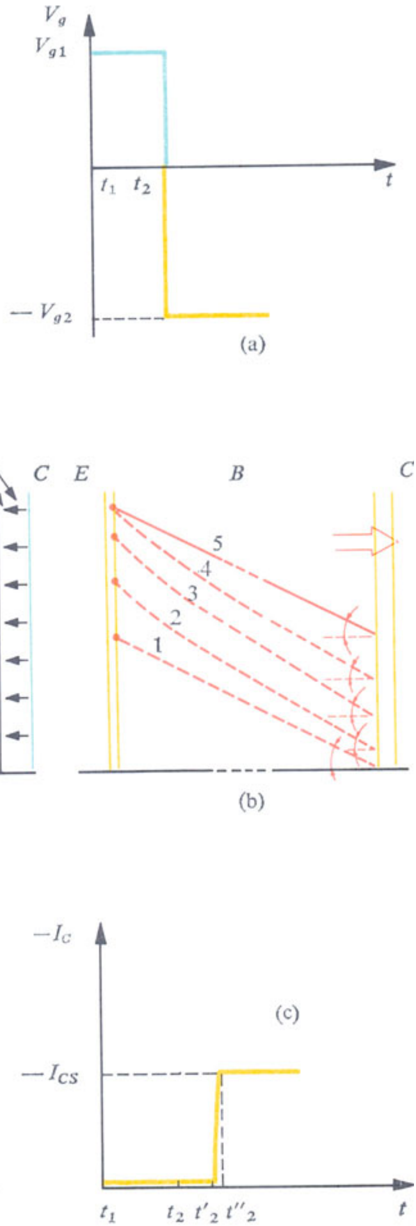


Fig. 181

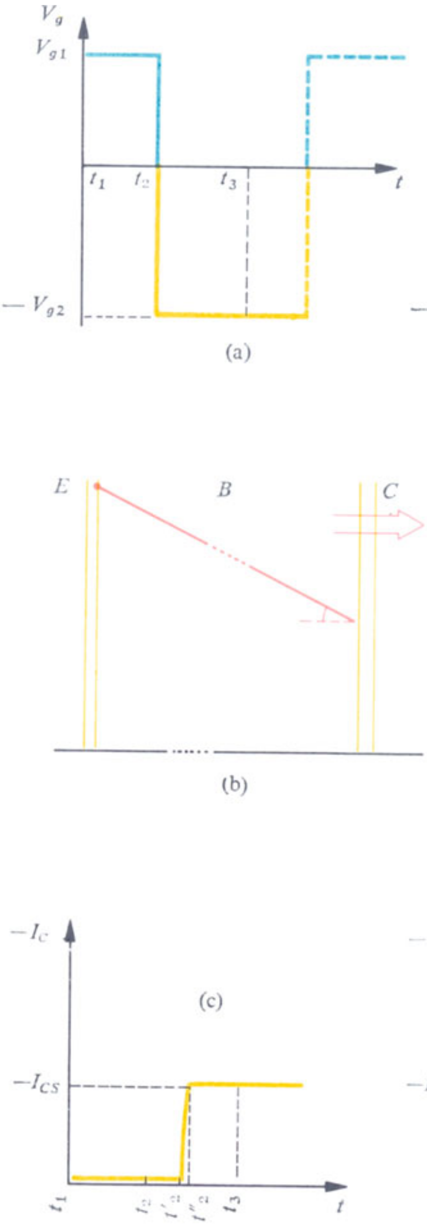


Fig. 182

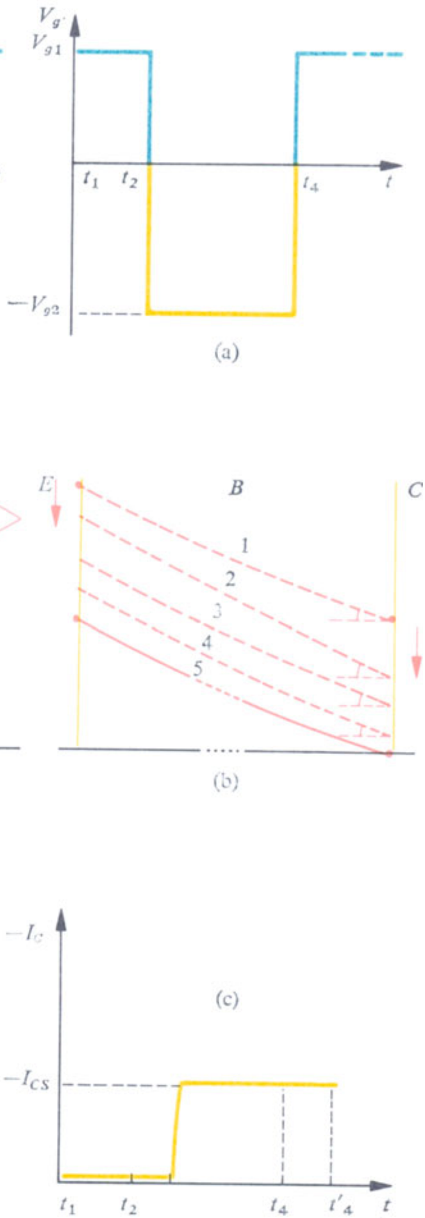


Fig. 183

At instant t_3 the voltage supplied by the group of generators is equal to $-V_{g2}$ (Fig. 182a). This voltage saturates the transistor.

The density graph of the holes in the base is consequently in the position shown in Fig. 182b. The tangent of the angle between this graph and the horizontal at the right of the base (that is in the neighbourhood of the base–collector junction) determines the value of the hole current flowing from the base to the collector. Only the tangent of this angle occurs in the expression for the collector current.

Since the transistor is saturated, the collector current is equal to $-I_{CS}$ (Fig. 182c). At instant t_3, whatever the degree of saturation of the transistor, the value of this current remains equal to $-I_{CS}$.

Explanation in the delay in the fall time of the current (t_s). At instant t_4 the signal supplied by the generator G disappears and the battery alone supplies a voltage to the input terminals of the transistor. Since this positive voltage is high, the transistor tends to be in a cut-off state (Fig. 183a).

From instant t_4 the reverse bias of the emitter–base junction causes the interruption of the strong diffusion of holes from the emitter to the base. The holes stored in this region will therefore continue to move; simultaneously the density graph moves downwards, and takes up the different positions shown in Fig. 183b. Position 5 determines the end of the saturation state of the transistor.

The time required for the hole density graph to move from position 1 to position 5 corresponds to the period of the removal of the surplus holes in the base of the transistor.

Whatever the position considered, the angle between this graph and the horizontal in the neighbourhood of the base–collector barrier remains practically constant. The diffusion current of holes has therefore not changed and the collector current remains at a constant value equal to $-I_{CS}$ (Fig. 183c).

The time interval $t_4' - t_4$ corresponds to the maintenance of the collector current at the saturation value and is defined in data sheets as the 'delay of the fall time of the current in the circuit' (t_s). This time increases when the transistor is strongly in the saturation state.

This is sometimes known as 'time of desaturation', and this definition has the advantage that it draws attention to the connection of the time with the effect causing it.

Influence of the type of drive on the time of desaturation

In the above discussion no account was taken of the influence of variations in the thickness of the barriers or of the drive system in defining the behaviour of the transistor. Figure 184a is a reminder of the process of the movement of the hole density graph. The value of the angle being constant, the collector current does not vary, but remains at the saturation value $-I_{CS}$ (Fig. 184b).

In practice there are two limiting possibilities for the drive of a transistor. What happens for example when the transistor is voltage driven? Does the drive system modify the shape of the signal arriving at the output?

Voltage drive. When the transistor is voltage driven, the density graph of the holes in the base moves from position 1 to position 5 from the instant when the command has been given to it to become non-saturated. The variation of the voltage applied between the base and the emitter has a direct effect on the bias conditions of this junction and very rapidly causes the widening of the barrier, which approaches a final position directly associated with the bias conditions imposed.

In position 1 the transistor is in the saturation state and the two barriers are very narrow; the current is at its saturation value ($-I_{CS}$ in Fig. 185b).

The hole density graph then starts to fall more rapidly on the left of the base than on the right. The corresponding angle has not changed and the current maintains its value $-I_{CS}$ (position 2 in Fig. 185b).

At stage 3 the density graph has swung again and the emitter–base barrier has widened; the angle between the graph and the horizontal on the collector side has not changed. The collector current is still equal to $-I_{CS}$.

From position 4 to position 5 the density level of the holes near the barrier, on the emitter side, is near to zero; this abrupt reverse bias of the junction causes a slight increase in the collector current (positions 4 and 5 in Fig. 185b).

Current drive. With current drive the applied voltage has a far slower influence on the thickness of the emitter–base junction; the variations of the collector current with time are much more like the purely theoretical shape mentioned above.

Figure 186a shows the different positions of the density graph of the holes in the base, and Fig. 186b the variations of the collector current.

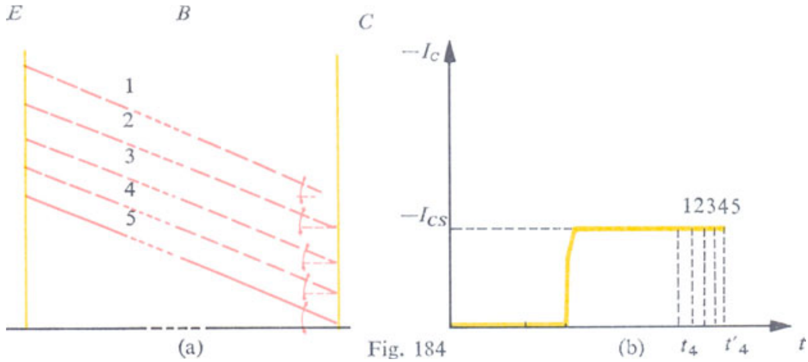


Fig. 184

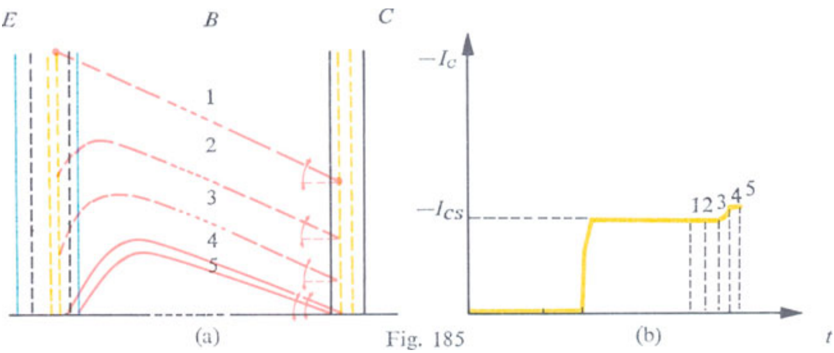


Fig. 185

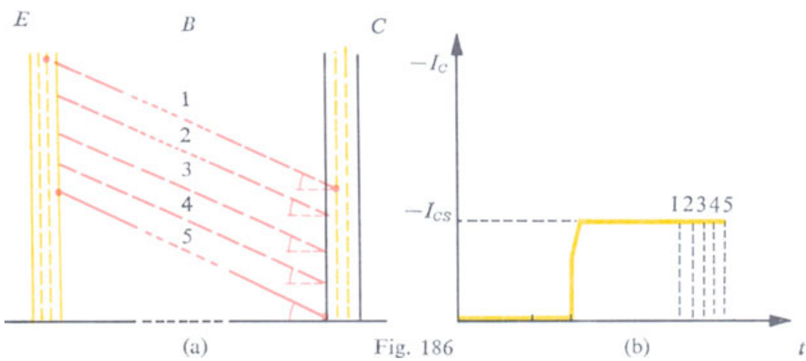


Fig. 186

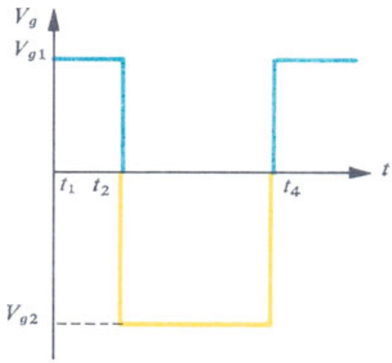


Fig. 187

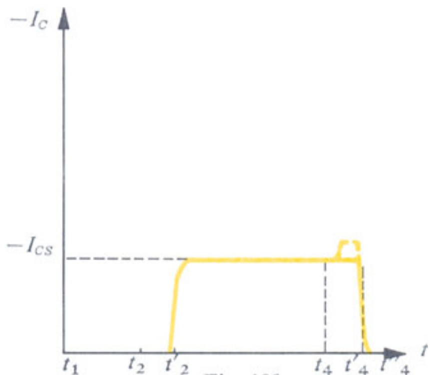


Fig. 188

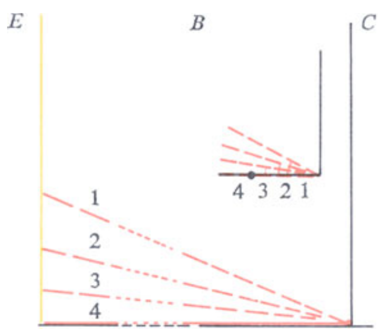


Fig. 189

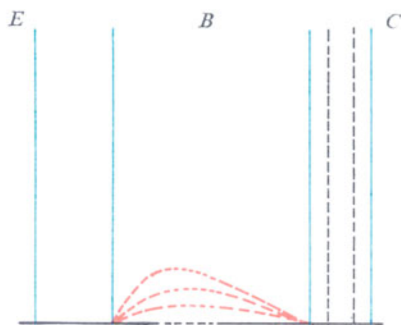


Fig. 190

At instant t_4 the interruption of the voltage supplied by the generator G produces a strong positive voltage between the base and the emitter of the transistor, which receives the command to operate in the cut-off state. The collector current may either remain at the value $-I_{CS}$ or exceed it slightly for a very short time.

At instant t_4' the transistor is no longer saturated. What happens at this instant?

Explanation of the fall time of the current (t_f). In Fig. 187 instant t_4' is associated with position 1 of the density graph of the holes in the base. The holes continue to diffuse towards the collector, and at the same time the graph swings towards the horizontal.

Position 2 immediately follows position 1, and the angle between this graph and the horizontal near the base–collector barrier decreases. This decrease causes a reduction in the current in the collector circuit of the transistor.

Position 2 is followed by positions 3 and 4, which are accompanied by a new decrease in the angle until it disappears completely. The collector current decreases and disappears at instant t_4'' in Fig. 188.

$t_4''-t_4'$ gives the time required for the current to change from its saturation value to zero; it is known as the 'fall time of the current' in the circuit (t_f).

In practice the fact of having placed a transistor in the cut-off condition is accompanied, as already mentioned, by a large widening of the base–emitter barrier. In this case the left end of the density graph of the holes in the base reaches zero level well before the other end. The graph then assumes the different positions shown in Fig. 190.

Whatever the type of reasoning chosen, the graph swings with relation to the point on the extreme right from a higher position associated with the value $-I_{CS}$ to a horizontal position corresponding to the absence of current.

From instant t_4' the current decreases fairly rapidly to zero. In position 1 in Fig. 189 the base–collector junction is no longer biased; it then becomes more strongly biased in the reverse direction as the collector current decreases. Any reduction in the collector current is accompanied by an increase in the negative voltage applied between the collector and emitter of the transistor.

When the transistor is switched from cut-off to saturation and from saturation to cut-off, the variations in the collector current with time are very different from the applied signal.

In fact, from instant t_1 to the instant preceding t_2, a strong positive voltage is applied between the base and emitter of the transistor. Since the transistor is then cut off, no current flows in the collector circuit.

At instant t_2 a strong negative voltage pulse tends to put the transistor in the saturation condition. From this instant the thickness of the emitter–base barrier decreases and the time $t_2' - t_2$ required for this narrowing determines the delay in the rise time of the current in the circuit.

After instant t_2' the applied voltage maintains the transistor in the saturation condition. The current begins to rise and the transistor reaches saturation; the collector current then assumes the value $-I_{CS}$ shown in Fig. 192.

From instant t_4 the voltage supplied by the generator tends once more to switch the transistor to cut-off. In fact, it is first necessary to remove the surplus holes in the base of the transistor, and the time necessary for this removal is given by $t_4' - t_4$ on the abscissa of Fig. 192 (it gives the time during which the current remains at its saturation value).

The hole density graph may then swing and the current decreases until it stops. $t_4'' - t_4'$ gives the limits of the fall time of the collector current.

Quantitative expression for the triggering time

(a) *Delay time of the current*

$$t_d = [R_B(C_{Te} + C_{TC}') + R_L C_{TC}'] V_{g2}/(-V_{g1} + V_{g2}) + R_L C_{TC}$$

(b) *Rise time of the current*

$$t_r = \tau_C' \log_e [\mathrm{h}_{FE}(-I_{B1})/\{\mathrm{h}_{FE}(-I_{B1}) + (-I_{CS})\}]$$

(c) *Desaturation time*

$$t_s = \tau_S \log_e [(-I_{B1} + I_{B2})/(-I_{CS}/\mathrm{h}_{FE} + I_{B2})]$$

(d) *Fall time of the current*

$$t_f = \tau_C' \log_e [1 + (-I_{CS})/\mathrm{h}_{FE} I_{B2}]$$

knowing that: $\tau_C' = \tau_C + \mathrm{h}_{FE} R_L C_{TC}$.

In these equations τ_C represents the time constant of the rise time of the current, and τ_S the time constant of desaturation. The capacitive effects represented by C_{Te}, C_{TC} and C_{TC}' will be studied in detail in the paragraphs entitled 'Equivalent circuits' and 'Switching parameters' of the transistor.

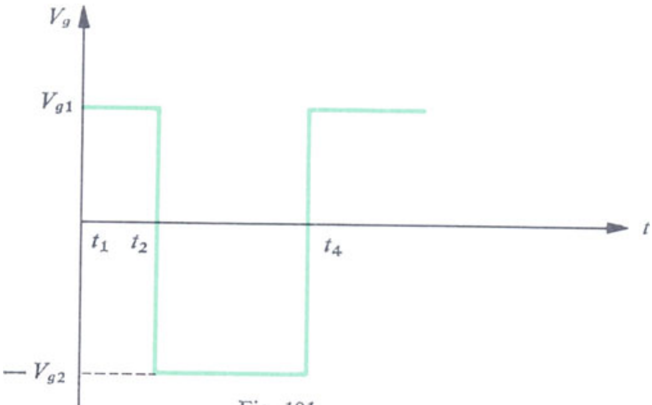


Fig. 191

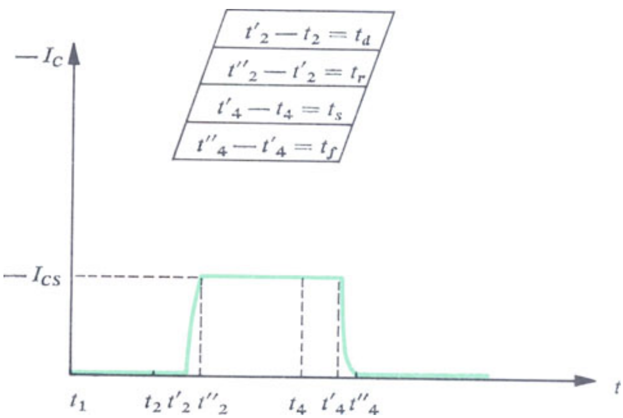


Fig. 192

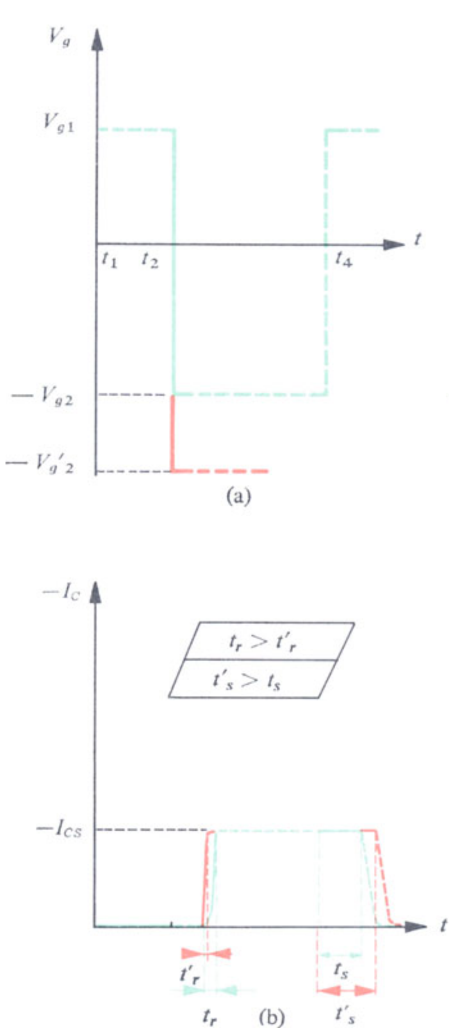


Fig. 193

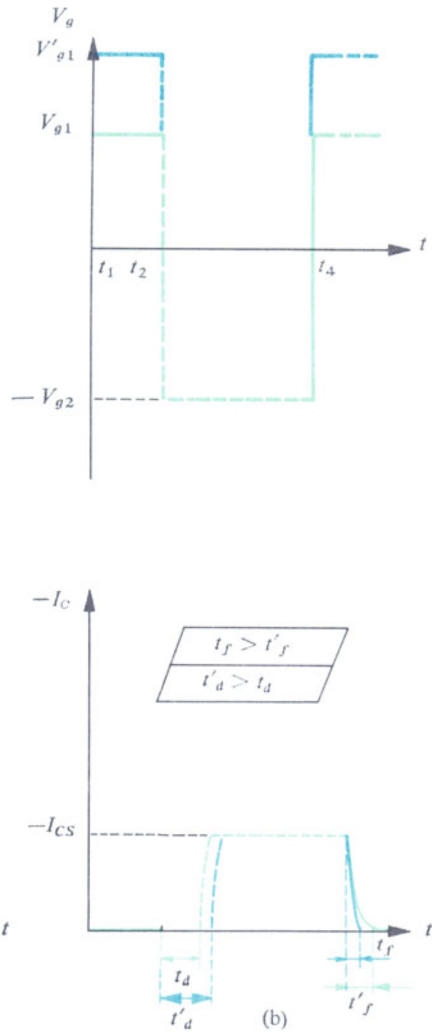


Fig. 194

We have already emphasized the advantage of bringing the transistor abruptly into saturation and switching it towards a very strong cut-off state. We will now attempt to determine the advantages and disadvantages of saturation on the one hand and cut-off on the other.

Advantages and disadvantages of a strong saturation. If at instant t_2 (Fig. 193a) a voltage $-V_{g2}$ is applied to the input terminals of the transistor, a certain time is required for the current to rise from zero to $-I_{CS}$; this time is shown in Fig. 193b in green. To this saturation level corresponds a certain storage of holes in the base of the transistor, resulting in the appearance at instant t_4 of a saturation time necessary for the removal of these carriers. Two periods can be then defined:

(a) the rise time of the current in the circuit (t_r);

(b) the desaturation time (t_s).

In Figs 193a and b the voltage $-V_{g2}$ is shown in green.

The use of a higher voltage at instant t_2 ($-V_{g2}'$ in red in Fig. 193d) involves a shorter rise time of the collector current due to the stronger saturation (namely t_r in red in Fig. 193b); conversely the removal time of the stored charges t_s is longer because of the higher position of the density graph.

The increase in the saturation state of the transistor has led to:

(a) a decrease in t_r (rise time of the current in the output circuit;

(b) an increase in t_s (delay in the fall time of the collector current).

Advantages and disadvantages of a strong cut-off. When the transistor passes from saturation to cut-off there are two possibilities:

(a) variation of the input voltage from $-V_{g2}$ to V_{g1} at instant t_4 (Fig. 194 in green);

(b) change in the voltage from $-V_{g2}$ to $+V_{g1}'$ at instant t_4 (in blue in the figure).

In the first case, that is for a voltage V_{g1}, the fall time of the current is given by t_f in green in Fig. 194b, and the delay time of this current by t_d in green.

For a higher voltage, namely V_{g1} (in blue in Fig. 194b) the fall time of the current decreases, on the other hand the delay time of the current t_d increases.

The accentuation of the cut-off of the transistor involves:

(a) a reduction of t_f (fall time of the current);

(b) an increase of t_d (delay in the rise time of the current).

Influence of the fall time t_d and t_s. We know the importance of the rise time of the current t_r and the fall time of the current t_f on the shape of the signal received in the output circuit. These times should be as short as possible.

On the other hand, what influence have the fall times t_d and t_s?

The use of transistors in switching circuits is equivalent to associating them with multivibrator circuits or blocking oscillators.

Let us consider, for example, a multivibrator. This type of stage in principle contains two transistors: the one is in the cut-off state while the other is saturated and vice versa.

There are various types of multivibrator; for example a bistable multivibrator embodying two transistors T_1 and T_2; a synoptic representation is given in Fig. 196.

Delay time t_d. If a negative pulse is available to trigger this circuit, it must be applied to the base of the transistor T_1 which is cut-off. This transistor then becomes saturated.

When the pulse reaches the base of the transistor T_1, the emitter–base barrier is very wide. Before the collector current can increase, the width of the barrier must be reduced, whence a delay time t_d associated with the cut-off state chosen (Fig. 197).

Delay time t_s. Conversely, if a positive voltage pulse is available, it should be used to switch the transistor T_2 initially to the saturated state, in order to bring it in a cut-off state. When the pulse is applied to the base of the transistor T_2, the corresponding collector current cannot decrease immediately. It is first necessary to remove the surplus holes in the base of the transistor. The time associated with the removal of these charges was previously represented by t_s.

If the duration of the pulse is shorter than the desaturation time of the transistor, the collector current cannot start to fall, the transistor T_2 remains saturated and the transistor T_1 is cut-off.

When the transistor is triggered by positive pulses, it is very important to take account of its desaturation time (Fig. 199).

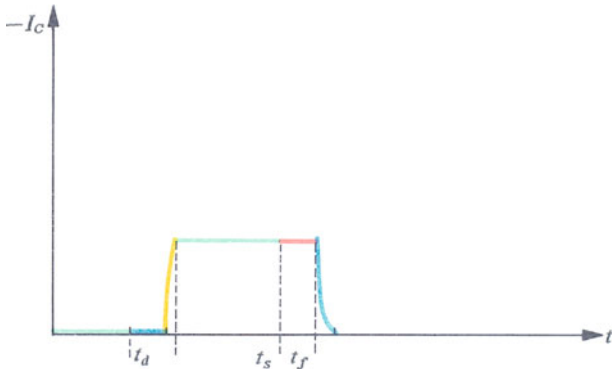


Fig. 195

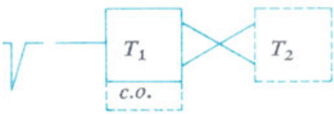


Fig. 196

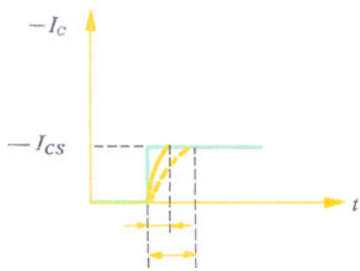


Fig. 197

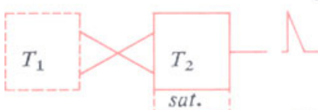


Fig. 198

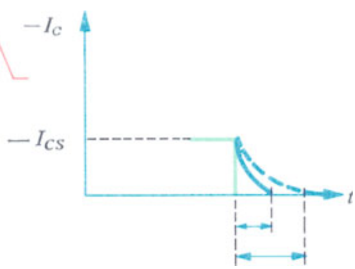


Fig. 199

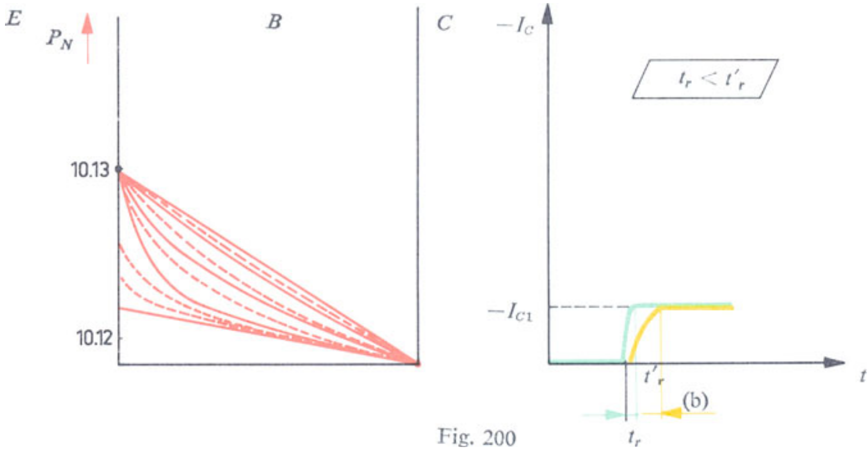


Fig. 200

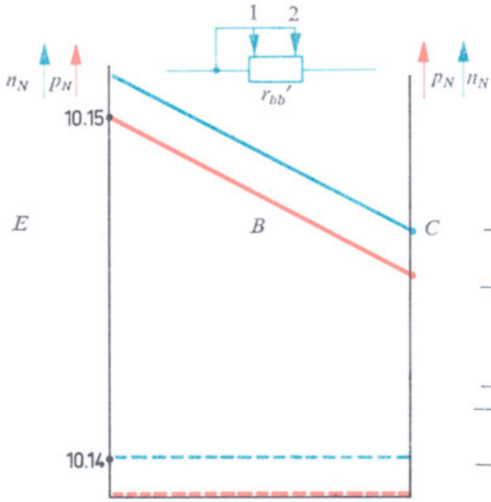


Fig. 201

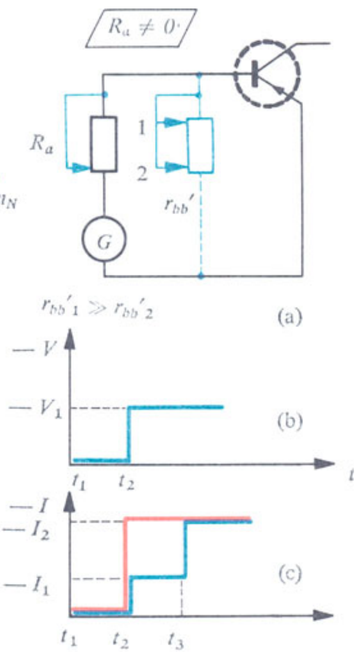


Fig. 202

2.12. Influence of the type of drive in the large-signal state on the shape of the collector current

We have dealt at some length with the influence of the type of drive. Referring to explanations already given, we must ask whether it is possible to associate with current drive a longer rise time than with voltage drive. The reader is referred to Fig. 200a, in which will be seen various density graphs for the holes in the base:

(a) voltage drive is shown in full line;

(b) current drive in dotted lines.

With voltage drive, the larger instantaneous arrival of holes in the base causes a shorter rise time of the collector current; consequently, the current variations (in green in Fig. 200b) are much more rapid than the corresponding variations (in yellow in the figure) which are associated with current drive.

Effect of saturation on the input resistance of a transistor

When a transistor becomes very strongly saturated, it would be of interest to determine the effect of the new positions of the majority and minority carrier density graphs on the intrinsic value of the resistance of the base. This resistance is in this case one of the principal components of the input resistance of the transistor.

With zero bias (Fig. 201), the majority and minority carrier density graphs are at their equilibrium level and the value of the intrinsic resistance of the base is relatively large.

In the saturation state the two density graphs become much higher, causing a big decrease in resistance r_{bb}' (position 2 of the slider of the potentiometer in Fig. 201).

With voltage drive the value of the input voltage at instant t_2 will be $-V_1$ (Fig. 202), and the value of the current will be $(-V_1)/r_{bb}'$.

At instant t_3 with the transistor saturated, the new value of the current is given by the relation:

$$-I_2 = (-V_1)/r_{bb2}'.$$

Voltage drive in these conditions becomes a delayed current drive and it is then preferable to inject a current pulse directly into the input circuit (Fig. 202c).

To recapitulate: a switching stage equipped with semiconductor devices consists of a transistor associated with a load unit and a drive system (Fig. 203). The latter may be represented by a generator with a certain internal resistance R_a, connected to the positive pole of a battery, with the negative pole connected directly with the emitter. A connection of this type is used in such a way as to switch the transistor to a cut-off state in the absence of a signal from the generator.

The load is represented by a resistance R_L connected between the collector and the emitter in dynamic conditions, and the supply to the collector is from a group of batteries $-V_{CC}$.

One of the very important factors in the determination of the type of drive of the stage is the input resistance of the transistor, namely r_{ie} in Fig. 203. The definition of the type of drive depends on the ratio between this value and the drive resistance.

In order to improve the rise time of the current (t_r) all we have apparently to do is to bring the transistor very abruptly into the saturation state whilst it is conducting. It will be seen from the networks of current output and transfer characteristics, shown in Fig. 204, that making the transistor strongly saturated involves the injection of a current $-I_{B2}$ into the input circuit, which should correspond to the flow of a current $-I_{C2}$ in the output circuit. In practice the projection on to the ordinate of the point of intersection of the load line and the destruction characteristic of the transistor ($-V_{CEK}$) gives the limiting value of the collector current $-I_{CS}$. The total variation of this current when the transistor changes from the cut-off state to saturation, is then given by the difference between $-I_{CS}$ and $-I_{CO}$.

Let us remind ourselves very briefly of the graphical representation of the improvement in the rise time of the current. When the transistor becomes strongly saturated the initial injection of holes into the left side of the N region, consisting of the base, is very large and causes the corresponding density graph to approach the most oblique position in Fig. 205 (chain-dotted lines). From the instant when the current reaches its saturation value, that is when the working point is in the red position in Fig. 204, the collector–base junction in its turn becomes biased in the forward direction. The holes then pass from the collector into the base and the density graph shown in dotted lines in Fig. 205 moves parallel to its original position to the position shown by a full line. The collector current no longer changes, but on the other hand very large numbers of holes are stored in the base.

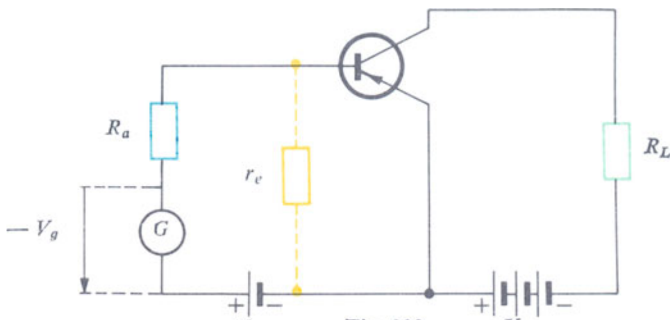


Fig. 203

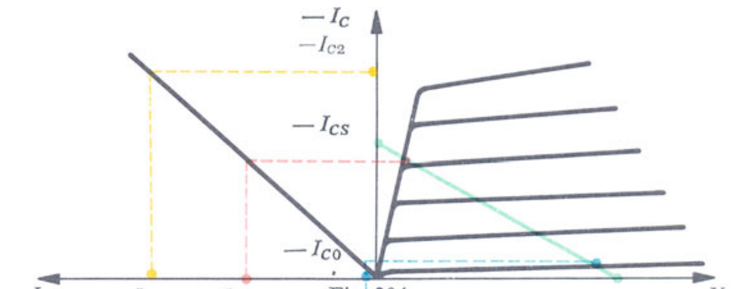


Fig. 204

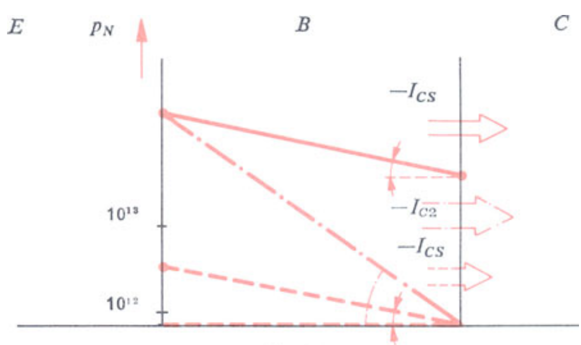


Fig. 205

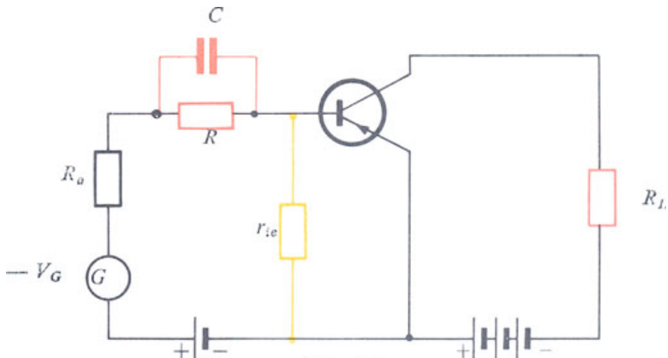


Fig. 206

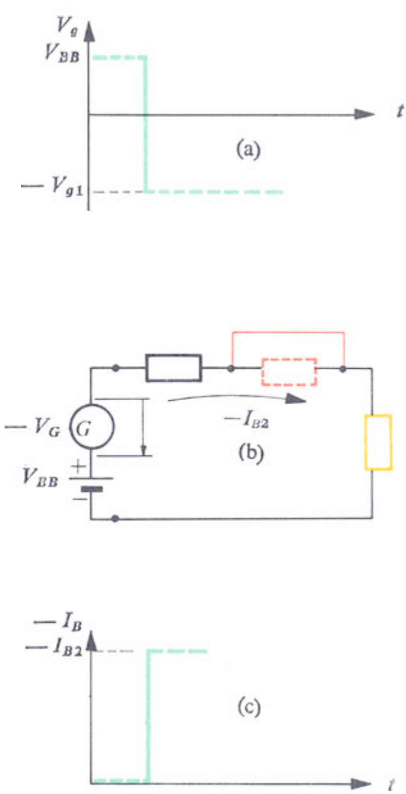


Fig. 207

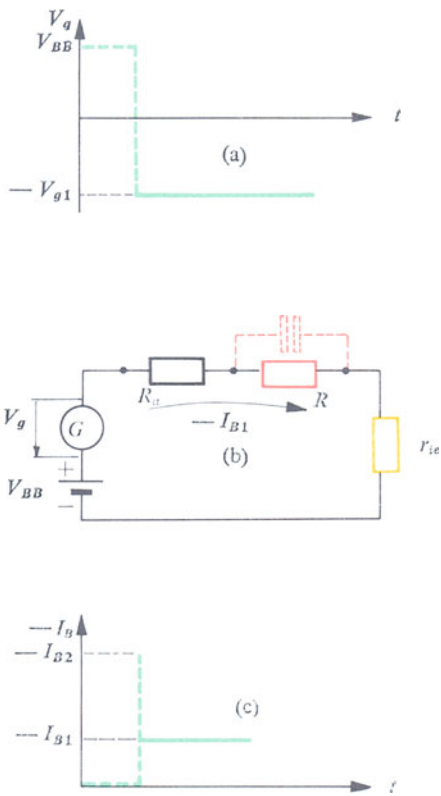


Fig. 208

2.13. Speed-up circuit

The chief drawback of a strong state of saturation is an increase in the time of desaturation. The desirability of limiting this time to a minimum has been emphasized earlier. It is therefore necessary constantly to seek a compromise between the shortest possible rise time and a very short desaturation time.

The simultaneous reduction in these two times seems to contradict the earlier explanations. There is however a circuit, known as the speed-up circuit, which provides for a very strong saturation during the transient period followed by a limitation of this state of operation to a relatively low level.

Composition of speed-up circuit

See Fig. 206. In addition to the standard components of a switching stage we find a resistance R associated with a capacitor C. The value of this resistance is higher than the sum of the input and driving resistances of the stage.

There are two stages to be noted:

(a) the transient stage: when the voltage changes from $+V_{BB}$ to $-V_{g1}$;
(b) the static stage: the voltage is maintained at the value $-V_{g1}$.

Influence of the speed-up circuit on the transient period

During the transient period the capacitor behaves practically as a short circuit, and the total resistance of the input circuit is a direct function of the drive and input resistances of the transistor, namely:

$$R_2 = R_i + r_{ie}.$$

We can then calculate the current changes as a function of time. The limiting value of this current is given by the relation:

$$-I_{B2} = \frac{-V_{g1}}{R_2} = \frac{-V_{g1}}{R_i + r_{ie}}$$

The shape of the base current, in the absence of any inertial effect, is shown in Fig. 207c by an instantaneous variation from zero to the maximum value $-I_{B2}$.

Behaviour of the speed-up circuit in the static condition

In normal state the capacitance presents an infinite impedance to the flow of the current, and the resistance R is automatically switched into the input circuit (Fig. 208b).

The value of the base current is given by the expression:

$$-I_{B1} = \frac{-V_{g1}}{R_1} = \frac{-V_{g1}}{R_i + R + r_{ie}} \quad \text{since} \quad R_1 = R_i + R + r_{ie}.$$

Since the resistance R has a very high value with respect to R_i and r_{ie}, the value of the associated current $-I_{B1}$ in normal state is much lower than the previous value $-I_{B2}$.

The fact of having connected a system of this nature in series with the generator causes a very definite increase in the number of holes injected into the base during the transient period, and then limits the saturation to a much lower value, namely to a very much reduced storage of holes.

Use of networks of characteristics and density graphs to indicate the advantages of the speed-up circuit

A full explanation of these effects can be obtained either from the network of transfer characteristics in the forward direction, or with the aid of the density graphs.

During the transient period, when the signal changes from $+V_{BB}$ to $-V_{g1}$ (Fig. 209a), the input current approaches the limiting value of $-I_{B2}$; the collector current should then reach the value $-I_{C2}$. Given the position of the load line on the network of the output characteristics of the transistor, the maximum possible value for the collector current is $-I_{CS}$ in Fig. 209b. In these conditions the rise time of the current, that is the time required for the collector current to rise from zero to the value $-I_{CS}$, is extremely short. Conversely, if the input current $-I_{B2}$ was maintained at this value there would be a risk that the hole density graph would move upwards, corresponding to a very large storate of holes in the base of the transistor.

We see on the density graph of the holes in the base (Fig. 209c) that:

(a) during stage 1 the current rises very rapidly; this rise should be associated with the final position of the density graph marked 3 in the figure;

(b) from position 1, the collector–base junction also becomes forward biased; the graph therefore moves from the two sides of the base simultaneously to position 2 in the figure.

A very large storage of holes in the base of the transistor is associated with this position 2.

After the transient period, that is when the input voltage is constant and equal to $-V_{g1}$ (Fig. 210a), the large decrease in the limiting value of the base current gives a final value of the collector current $-I_{CS}$ on the ordinate of the network of the transfer of current characteristics of Fig. 210b. A much lower degree of saturation than in Fig. 209c is associated with this value of the collector current.

The position 1 of the density graph of the holes in the base determines the start of the saturation of the transistor; this position is in fact reached very rapidly since it is located in the transient period.

The stabilization of the collector current at a relatively low value with respect to $-I_{C2}$ limits the movement of the hole density graph to position 2 in Fig. 210c. The storage of holes in the base is then smaller.

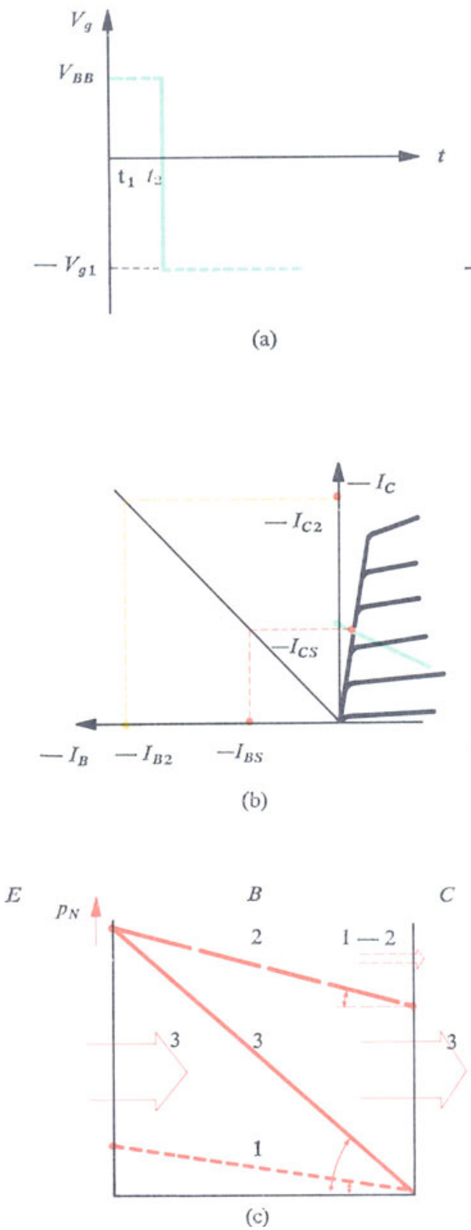


Fig. 209

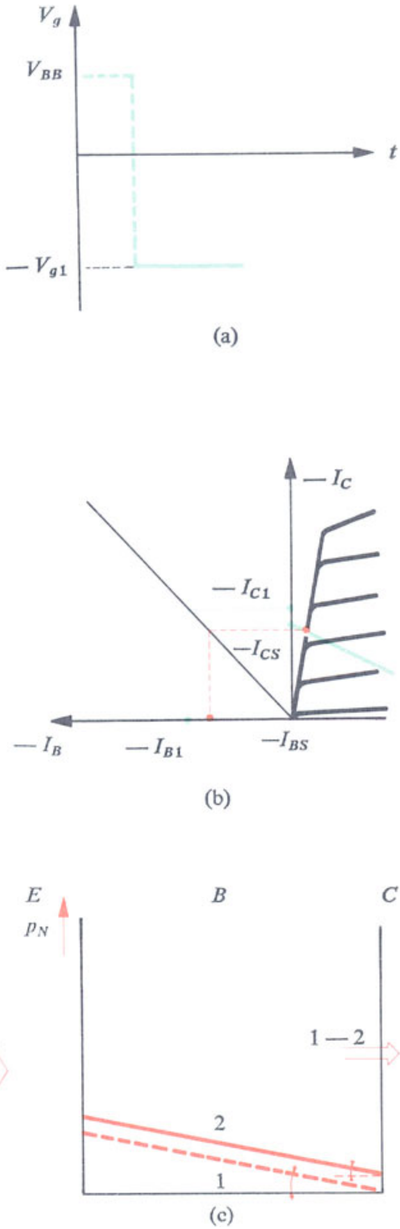


Fig. 210

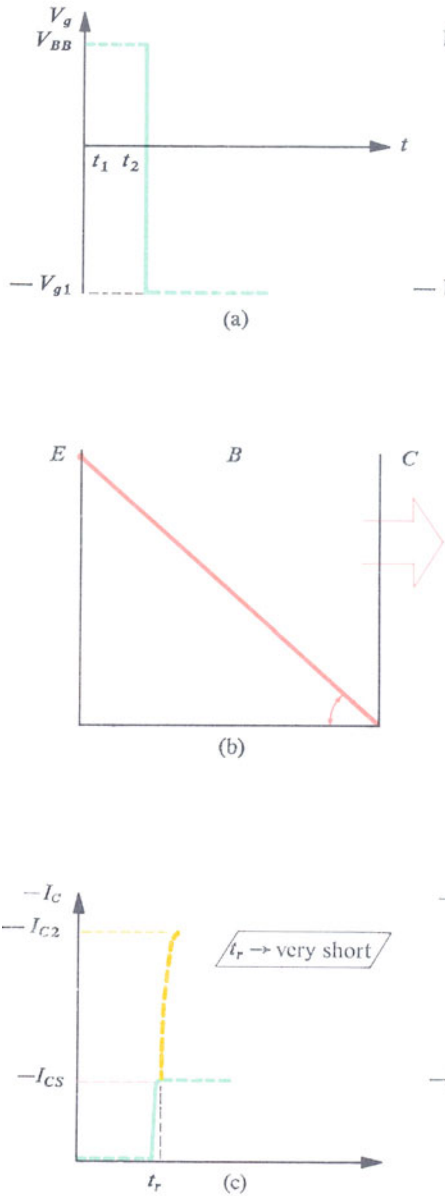


Fig. 211

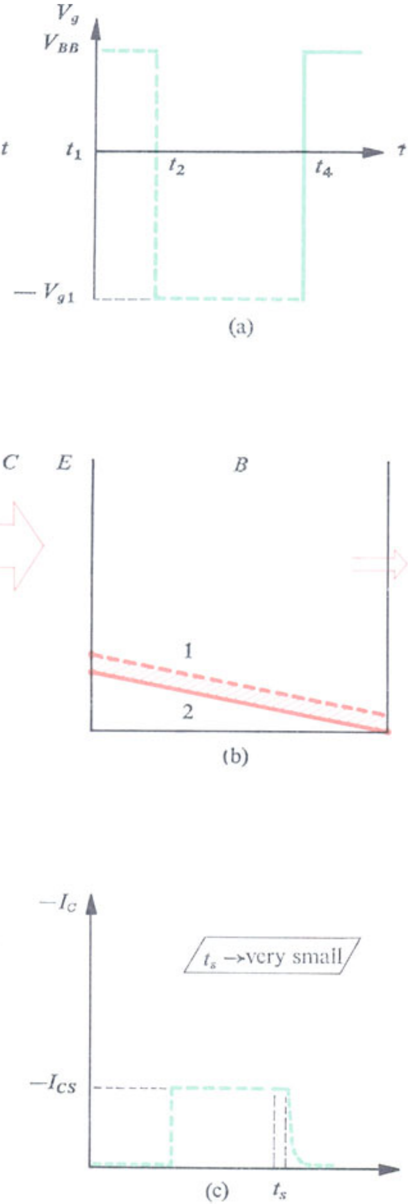


Fig. 212

The conclusions to be drawn from an examination of the hole density graph may be confirmed by the representation of the effective variations of the collector current with time in the following two cases:

(a) transient period;

(b) stable period, and return to the cut-off state.

Establishment of the collector current (*change to saturation*). When the voltage changes from $+V_{BB}$ to $-V_{g1}$ (Fig. 211a) the final value of the collector current is dependent on the position of the density graph corresponding to the absence of saturation (Fig. 211b).

The variation of the collector current is shown in yellow in Fig. 211c; the rise time of the current is relatively short. When the transistor becomes saturated the value of the collector current is limited to $-I_{CS}$: the rise time of the current in green in the circuit (t_r) now becomes very short; this time in fact represents only a very small variation in the total increase in the collector current.

During the stable period, the density graph is near the point when saturation sets in (position 1 and 2 in Fig. 212b).

Desaturation and fall time of the current (*initiation of cut-off*). At instant t_4 the voltage at the input terminals of the transistor approaches the value $+V_{BB}$ fed from the battery. The transistor receives the command for the cut-off state to set in. Before there is any sign of a possible decrease in the collector current, the level of the hole density in the base, near the barrier on the collector side, should first return to a zero value. The time required for the graph to move from position 1 to position 2, in Fig. 212b, is a function of the shaded area representing the surplus holes stored in the base of the transistor. This area is small and the desaturation time t_s is therefore short.

Figure 212c shows the variations of the collector current as a function of time; the desaturation time t_s is short in spite of the fact that the rise time of the current (t_r) is also very short. The density graph swings from position 2 towards the horizontal. The collector current then decreases very rapidly, because of the cut-off state chosen.

The speed-up circuit then has an influence on the rise time and desaturation time of the current; we must however not lose sight of the improvement in the delay time of the current in the circuit. We will not attempt to explain this improvement, but by means of the formulae on the next page we can gain a better understanding of the influence of a circuit of this type on the shape of the current in the collector circuit.

Expression for the triggering time (with speed-up circuit)

We have given considerable space to the principles underlying the use of semiconductor devices for switching. A stage of this nature consists primarily of a transistor associated on the one hand with a driving circuit, and, on the other hand, with a load circuit. With the initial position in the cut-off state, the positive pole of the battery is then connected to the base of the transistor through the generator and its internal resistance. The load circuit consists of a resistance associated with a group of supply batteries $-V_{CC}$ (Fig. 213).

In practice, in the majority of monostable or bistable multivibrators speed-up circuits are used as the connecting system, as shown in red in Fig. 214. This system consists of a high resistance and capacitor in parallel.

The operating principle of this system is very simple, and can be explained quite briefly:

(a) in the transient state the resistance R is short-circuited by the capacitor; the final current variation in the base then tends to become very large, and the degree of saturation of the transistor apparently very high;

(b) in the normal state the value of the resistance presented to the base current is much lower, and the transistor becomes much less saturated.

The circuit of this type has a very strong influence on the different times t_d, t_r, t_s and t_f; which can be calculated very easily in the present case by means of the following equations:

$$t_d = \left[(R_i + r_{bb}')(C_{TE} + C_{TE}') + R_L C_{TC}' \cdot \frac{V_{g2}}{-V_{g1} + V_{g2}} + R_L C_{TC}\right]$$

$$t_r = \tau_C' \frac{-I_{CS}}{h_{FE}'(-I_{B1})}$$

$$t_s = \tau_s \log_e \left[\frac{-I_{B1} + I_{B2}}{-I_{CS}/h_{FE}' + I_{B2}}\right]$$

$$t_f = \tau_C' \log_e [1 + (-I_{CS})/h_{FE} I_{B2}]$$

knowing that:

$$-I_{B1} = \frac{-V_{g1}}{R}; \qquad \tau_C' = \tau_C + h_{FE} R_L C_{TC}.$$

2.14. Dynamic equivalent circuits of a transistor

It is not possible to design a transistorized switching stage simply by the application of a network of characteristics or by means of parameters

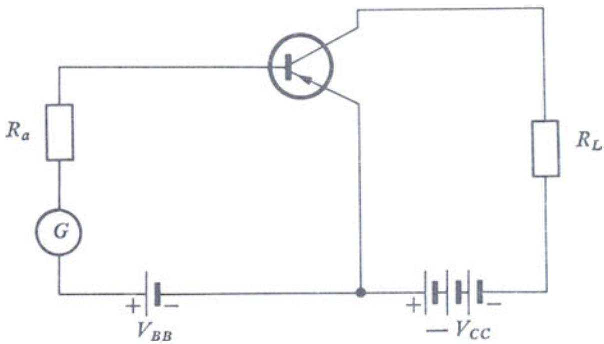


Fig. 213

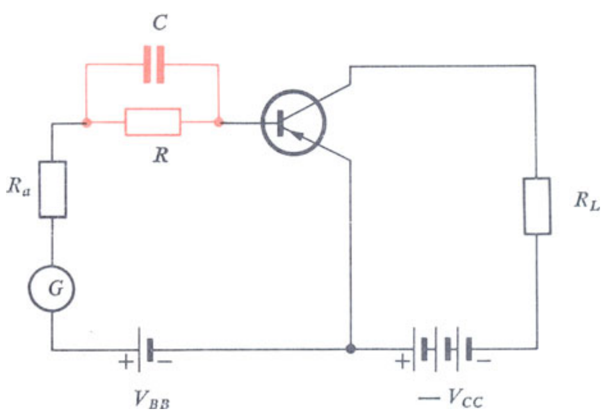


Fig. 214

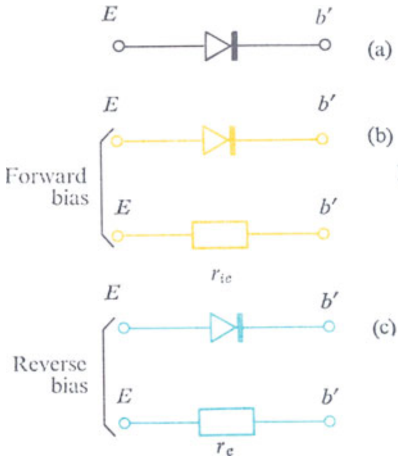


Fig. 215

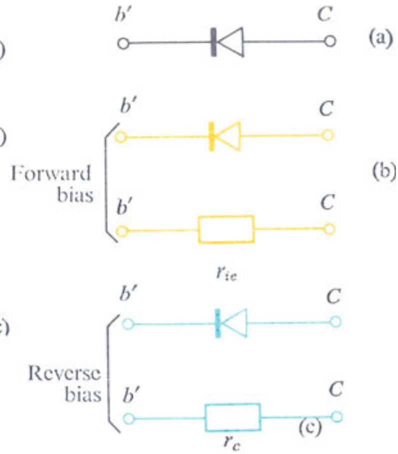


Fig. 216

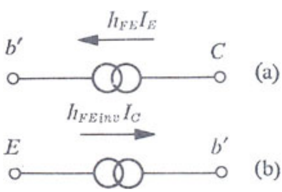


Fig. 217

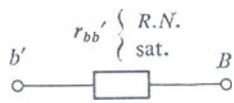


Fig. 218

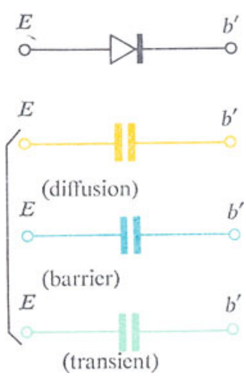


Fig. 219

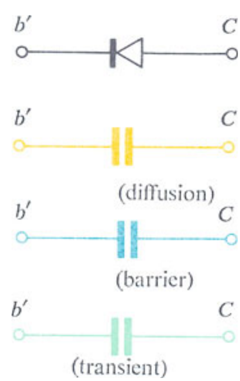


Fig. 220

such as those mentioned in connection with audio-frequency applications. Although these characteristics are of great interest, especially for the study of stable periods (cut-off and saturation), they are useless for the transient periods. We must therefore have recourse to dynamic equivalent circuits of the transistor, known also as 'equivalent electrical switching circuits'.

They contain a certain number of data, which we will now attempt to describe.

A transistor consists essentially of two junctions:

(a) The emitter–base junction (between the emitter and the imaginary point on the base) shown in Fig. 215a. This junction may be: either biased in the forward direction as in Fig. 215b, (associated in this case with a low value of resistance r_{ie}); or reverse biased (Fig. 215c), when the value of the corresponding resistance is much greater: r_e.

(b) The collector–base junction (between the collector and the imaginary point on the base, Fig. 216a). This junction may be biased in the forward direction (Fig. 216b) giving a low value of resistance r_{ic}, or in the reverse direction (Fig. 216c), with a very high resistance r_c.

Two current generators are associated with the forward bias conditions:

(c) the generator $h_{FE}.I_E$ associated with the forward bias of the emitter–base junction;

(d) the generator $h_{FE\,inv}.I_C$, dependent on the forward bias of the collector–base junction.

Between the imaginary point on the base and the corresponding electrode there is a block of weakly-doped N type material; we must therefore show the intrinsic resistance r_{bb}' of this material, the value of which changes as a function of the bias conditions chosen (Fig. 218).

The information thus made available enables us to represent the static equivalent circuit of the transistor. The various inertial effects discussed above cannot be defined by means of purely resistive elements or by the presence of current generators. It is therefore essential to introduce the capacitive effects.

There are three possible values of capacitance associated for each junction with three basic effects.

Consider for example the emitter–base junction (Fig. 219): this junction may present three different capacitive effects:

(a) in yellow, C_{eb}' diffusion; (b) in blue, C_{eb}' barrier; (c) in green C_{eb}' transient.

Similarly for the collector–base junction (Fig. 220) we have:

(a) in yellow, C_{cb}' diffusion; (b) in blue, C_{cb}' barrier; (c) in green, C_{cb}' transient.

Explanation of the capacitive effects peculiar to transistors

Before constructing the 'dynamic equivalent switching circuit' of the transistor it would be useful to define what we mean by capacitive effects and especially the possibility of distinguishing between them.

Let us consider, for example, the base–emitter junction of the transistor; it has three capacitive effects:

(a) the diffusion capacitance;

(b) the barrier capacitance;

(c) the transient capacitance.

Diffusion capacitance. When a junction is biased in the forward direction (Fig. 221b), the capacitance associated with this junction has a very high value; this is the diffusion capacitance (C_{eb}' in Fig. 221a).

If we assume a junction biased in the forward direction, this bias is associated with a very precise position of the density graph of the holes in the base (position 1 in Fig. 221c), and a certain value of collector current is associated with this graph. For a stronger bias the graph moves from position 1 to position 2 and the current increases; if the bias decreases, the graph moves in the opposite direction. The bias of a junction causes the graph to move from one position to another (either a movement of holes from position 1 to position 2 or a withdrawal from position 2 to position 3). This effect of supply or withdrawal of charges corresponds to the charging or discharge of a capacitor, hence the term diffusion capacitance of the junction.

Barrier capacitance. The concept of barrier capacitance is easier to understand; it is associated with the reverse bias of the junction (Fig. 222a and b).

When a junction is biased in the reverse direction, the barrier widens and it contains very few carriers.

The absence of carriers results in a big increase in resistivity; this region may be likened to the dielectric of a capacitor. The expression for this barrier capacitance is easily obtained by the formula $C = \varepsilon \cdot S/d$ in which S represents the area of the junction and d its thickness.

Transient capacitance. Finally, we have to define the transient capacitance of the junction, known also as the integrated barrier capacitance. This effect is dependent on the change from a reverse bias (position 1 in Fig. 223a) to a forward bias (position 2 in the same figure). In position 1 the junction is biased in the reverse direction, and the two space charge regions are very large (in blue in Fig. 223b). In position 2 the barrier is very narrow (in yellow in the figure). The arrival of charges required for the neutralization of these two regions may be likened to the charging of a capacitor. This effect occurs only during a transient period, and it is consequently known as the junction transient capacitance.

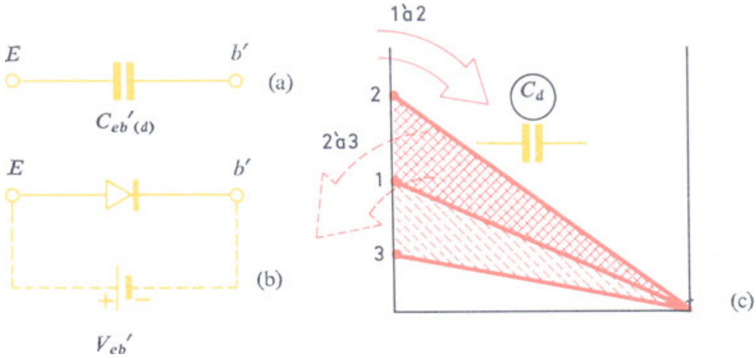


Fig. 221

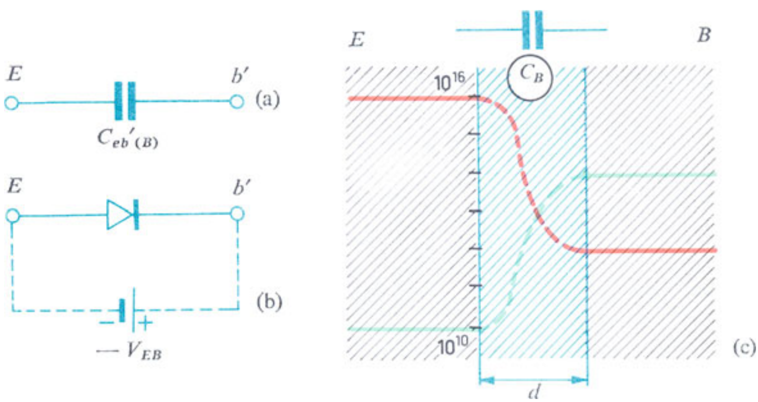


Fig. 222

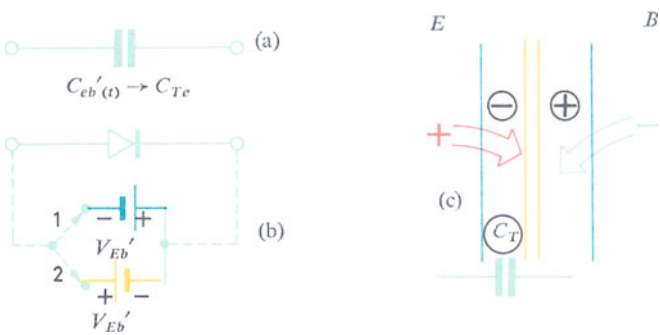


Fig. 223

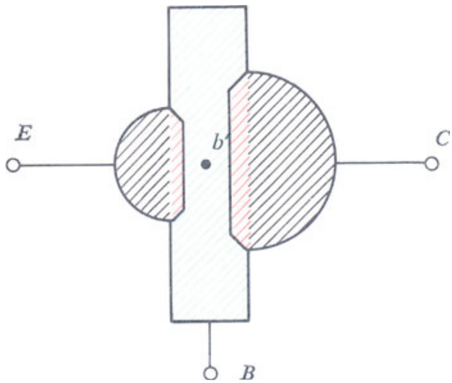


Fig. 224

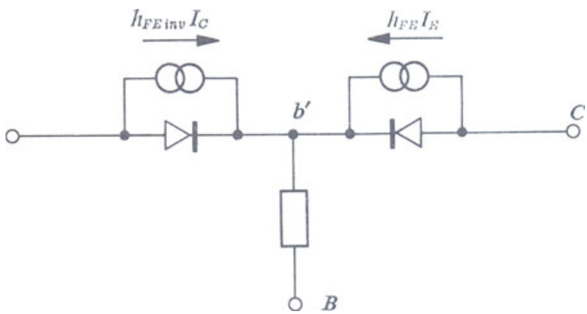


Fig. 225

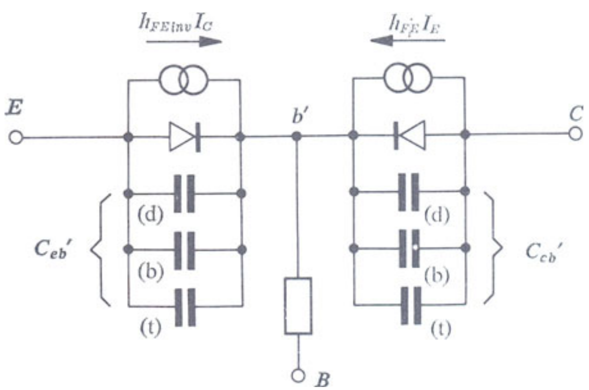


Fig. 226

Static equivalent circuit

The best way to define the equivalent electrical circuit of a transistor is to consider it directly from the physical point of view.

An alloy transistor is shown in Fig. 224. It contains two junctions:

(a) the emitter–base junction, located between the emitter terminal and the imaginary point in the base b';

(b) the collector–base junction, located between the collector and the imaginary point on the base b'.

The base electrode is separated from the point b' by a block of N type germanium.

We can now replace each of these junctions by their corresponding symbols in the simplified equivalent circuit in Fig. 225.

A circuit of this type should provide all the information necessary for the study of any type of operation of the transistor; we must therefore associate with these two junctions the current generators $h_{FE} \cdot I_E$ and $h_{FE\,inv} \cdot I_C$. Their presence is explained by the forward bias of the two junctions and they represent the transistor effect.

Dynamic equivalent circuit

To pass from the simplified static equivalent circuit to a dynamic equivalent circuit, we need only consider the same elements and associate them with the capacitive effects representing the transient behaviour of the two junctions.

Consequently the dynamic equivalent 'switching' circuit of the transistor includes:

(a) between the emitter and the imaginary point on the base, a junction with a current generator ($h_{FE} \cdot I_E$) and three capacitors C_{eb}' connected in parallel with it: the first represents the diffusion effect, the second the barrier capacitance and the third the transient capacitance;

(b) between the collector and the imaginary point on the base, a second junction with a current generator $h_{FE\,inv} \cdot I_C$ and three capacitors connected in parallel with it, which represent respectively: the diffusion effect, the barrier capacitance and the transient behaviour;

(c) between the imaginary point on the base and the base electrode, a resistance r_{bb}' which may have different values (value of r_{bb}' corresponding to the normal state and a much smaller value of r_{bb}' associated with the initiation of saturation of the transistor).

This equivalent circuit is not the exact representation of a dynamic equivalent circuit as the junctions are represented only by their symbols; in practice the bias conditions of these junctions enable them to be replaced by resistances of approximate value.

When consulting data sheets of transistor terms we often find equivalent circuits identical with that in Fig. 227. Our problem then is, knowing the origin of the components of this circuit, to construct circuits for the different operating conditions of transistors. In other words, how can we use such an equivalent circuit?

We have so far mentioned four possible types of operation of a transistor:

(a) the cut-off state;

(b) the normal state;

(c) the reverse state;

(d) the saturation state.

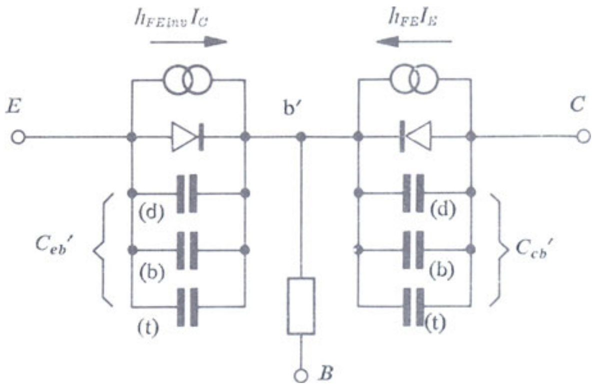
These four states cannot be discussed simultaneously: in fact, at a given instant, the transistor is in a well defined state. We will therefore try to represent the dynamic equivalent circuits of the transistor for each of these states and even, if necessary, to emphasize in passing the modifications of the values dependent on transient effects.

Equivalent circuit in the cut-off state. In the cut-off state the two junctions are biased in the reverse direction. The representation of the equivalent electrical circuit of the transistor in this state involves first the preparation of a figure containing the two junctions and the resistance r_{bb}'.

Because of the reverse bias condition of these junctions we can replace the emitter–base junction by a resistance of high value r_e; the collector–base junction may be also represented by a resistance r_c of high value (Fig. 228).

No current generator is associated with reverse bias; on the other hand the widening of the two barriers in the emitter–base and collector–base junctions requires the representation of the capacitive effects of the barrier. This equivalent circuit therefore consists of the capacitance C_{eb}' (barrier) between the emitter and the imaginary point on the base, and the capacitance C_{cb}' (barrier) between the collector and this same point.

Equivalent circuit in the normal state. In the normal state the emitter–base junction is biased in the forward direction, and it is therefore represented in Fig. 229 by a very low resistance r_{ie}; the collector–base junction is biased in the reverse direction and is therefore represented by a resistance r_c of high value. The forward bias of the emitter–base junction involves the representation of the current generator $h_{FE} \cdot I_E$. In addition to the resistance r_{bb}', the diffusion capacitance $C_{eb}'_{(d)}$ must be included for the emitter–base junction, and the barrier capacitance $C_{cb}'_{(b)}$ for the collector–base junction.



Fig, 227

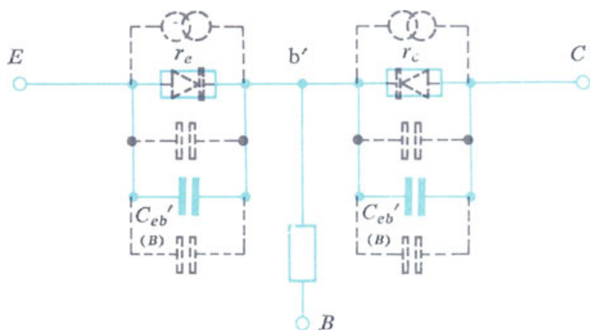


Fig. 228

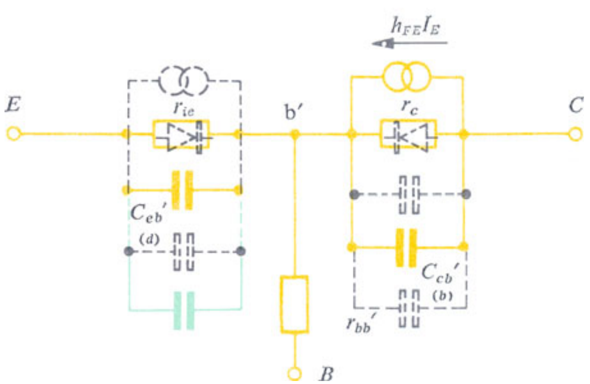


Fig. 229

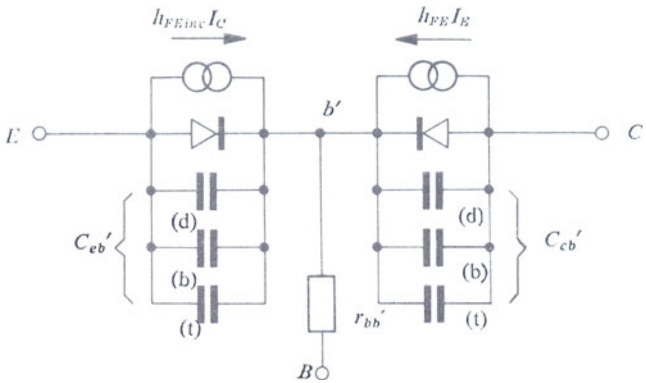


Fig. 230

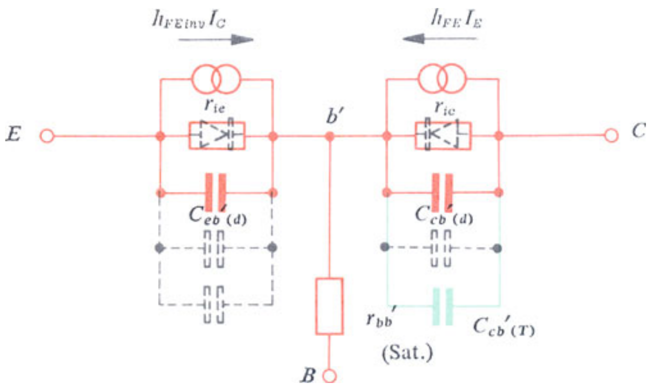


Fig. 231

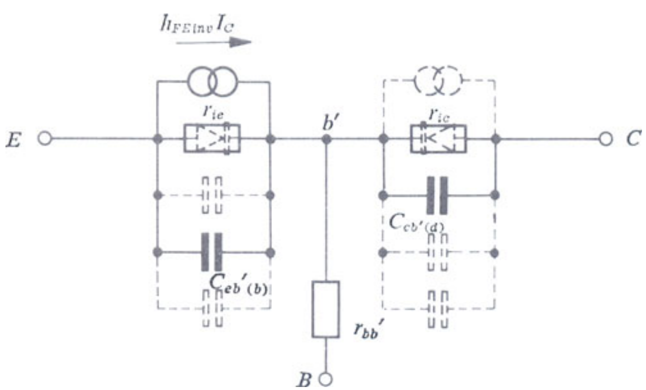


Fig. 232

Change from cut-off to the normal state. When the transistor changes from the cut-off to the normal state the biasing of the emitter–base junction in the forward direction causes the barrier to narrow considerably; the transient period calls thus for a modification of the equivalent circuit, which includes, (in green in Fig. 229) the capacitance $C_{eb}'_{(T)}$ (transient capacitance of the junction).

The representation of the complete equivalent circuit of a switching transistor, unlike the two states discussed above, requires the inclusion of all the known components (Fig. 230).

Equivalent circuit in the saturation state. The emitter–base junction is biased in the forward direction and it can, therefore, be represented by a low resistance r_{ie} (Fig. 231).

The collector–base junction also is biased in the forward direction, resulting in the low value resistance r_{ic}.

The forward bias of the two junctions accounts for the simultaneous presence of the two current generators $h_{FE} \cdot I_E$ and $h_{FE\,inv} \cdot I_C$. In addition to the resistance r_{bb}', the value of which is a function of the degree of saturation, various capacitances must be added; the forward bias of the emitter–base junction results in the diffusion capacitance C_{eb}' and the forward bias of the collector–base junction in the diffusion capacitance C_{cb}'.

When the transistor passes from the normal state, dependent on a reverse bias of the collector–base junction, to the saturation state caused by a forward bias of this junction, it is essential for passing from one state to another that the width of the collector–base barrier should be decreased and that consequently the transient capacitive effect, shown in green in Fig. 231, should be associated with the equivalent electrical circuit.

Equivalent circuit for inverse state. The fourth operating state, an exclusively transient one, requires the representation of the dynamic equivalent circuit of the transistor for inverse state.

The forward bias of the collector–base junction accounts for the use of a resistance r_{ic} of low value in Fig. 232 and the connection of a current generator $h_{FE\,inv} \cdot I_C$ in parallel with the emitter–base junction.

The reverse bias of the emitter–base junction requires the use of a high value resistance r_{ie} associated with the corresponding barrier capacitance (C_{eb}' barrier).

We must remember that in addition to the resistance r_{bb}' (normal state), it is essential to represent the diffusion capacitance associated with the forward bias of the collector–base junction (C_{cb}' diffusion).

In practice, if the design of a multivibrator requires a definition of the triggering time and the delay time of the circuit, it will be necessary to associate with the various components external to this multivibrator the components peculiar to the 'dynamic switching equivalent circuit' of the transistor just described.

There is a dynamic equivalent circuit corresponding to each transistor configuration. The circuit with the transistor in common base is shown again in Fig. 233. We need not again emphasize the justification of the different components of this circuit, since their origin and the method of use were dealt with earlier at some length.

Equivalent switching circuit of a transistor in common emitter

In switching, on the other hand, it is much more common to use the common emitter configuration. The reasons why our initial study was not based on this configuration are purely academic. It is in fact simpler to discuss the various components of the equivalent circuit of the common base configuration knowing that the two circuits, input and output, are completely different.

Let us now try to establish the equivalent circuit of the transistor in common emitter.

The base constitutes the input electrode and it is separated from the imaginary point of the base by resistance r_{bb}' (Fig. 234).

The collector–base junction is located between the imaginary point b' on the base and the collector; it is represented by the corresponding symbol and it could be replaced by a resistance of value r_{ic} or r_c depending on the bias conditions affecting this junction. The current generator $h_{FE} \cdot I_B$ and the three capacitors associated with the three capacitive effects $C_{b'c}$ diffusion, $C_{b'c}$ barrier and $C_{b'c}$ transient are connected in parallel with this junction.

Between the imaginary point on the base b' and the emitter electrode (electrode common to the input and output) the emitter–base junction is located; it may also be replaced by resistances, the value of which will be a function of the bias conditions of the junction. In parallel with this junction, the current generator $h_{FE\,inv} \cdot I_B$ and the three capacitors dependent on the following capacitive effects, are connected:

(a) the diffusion capacitance $C_{b'e(d)}$;

(b) the barrier capacitance $C_{b'e(b)}$;

(c) the transient capacitance $C_{b'e(T)}$.

The equivalent circuit of a transistor in common collector could be represented similarly, but in view of the small amount of interest in this configuration in standard applications we will not pursue the subject.

It would however be quite easy, starting from the equivalent circuit of the transistor in common base, to reproduce it by moving the electrodes and associating with them components the values of which would have been recalculated if required.

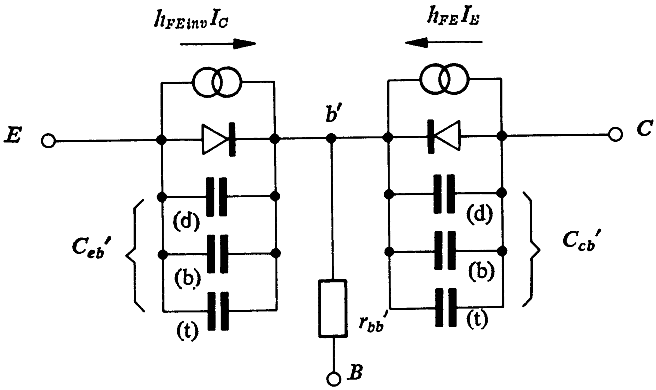


Fig. 233

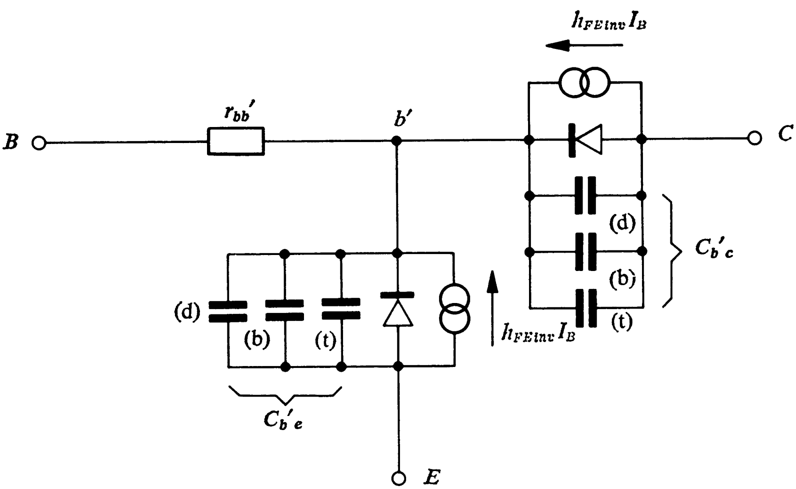


Fig. 234

Current amplification factor (normal state)

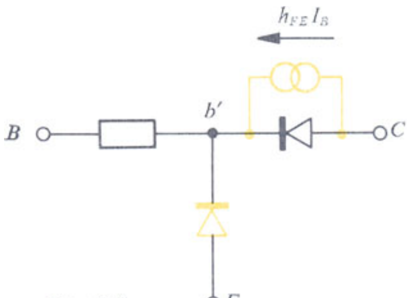


Fig. 235

Current amplification factor (inverse state)

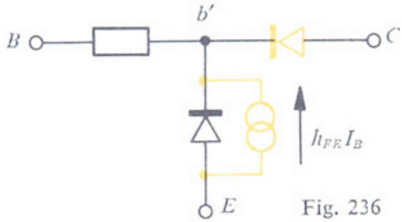


Fig. 236

Intrinsic base resistance

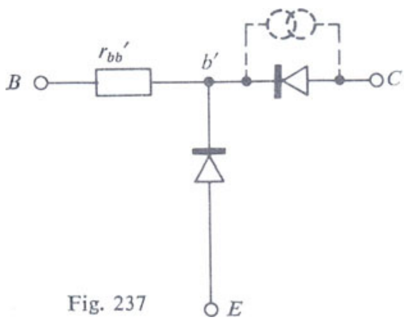


Fig. 237

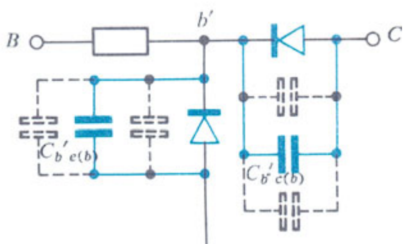


Fig. 238

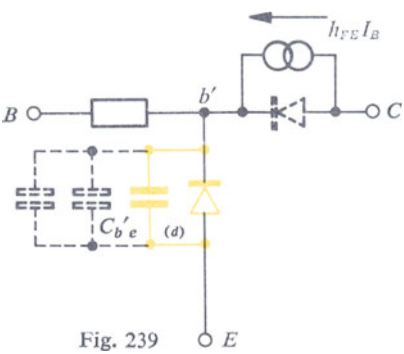


Fig. 239

Capacitative effects

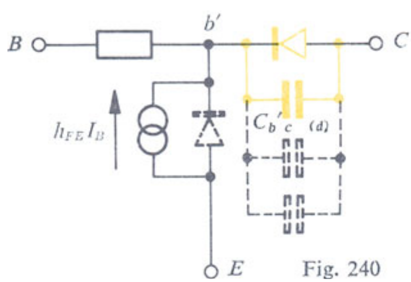


Fig. 240

2.15. Switching parameters of the transistor

In addition to the characteristics which are of great interest for the study of static states and equivalent circuits, which may sometimes be difficult to use and of comparatively little interest for slow-switching state, it will be useful to consider the parameters which are specific to 'switching' of the transistor.

It very often happens, when we are dealing with transistors designed for pulse techniques, that the manufacturers supply data on the delay, rise and fall times of the current.

We may however sometimes use other transistors for which this information is not available. It is then useful to be able to calculate the data from the normal parameters which are essential for the study of the behaviour of the transistor in audio- and radio-frequency applications.

The 'switching' parameters may be grouped into two main categories: (a) static parameters; (b) dynamic parameters.

Static parameters

Amplification factor of the current (normal state). In the expression for the different time constants which we will be defining shortly, the amplification factor of the current in common emitter (h_{FE}) always occurs.

This factor appears in the simplified equivalent circuit in Fig. 235. It is in fact associated with the forward bias of the emitter–base junction.

Current amplification factor (inverse state). Another very important parameter which is indispensible for the calculation of the desaturation time constant is the current amplification factor with the transistor in the inverse state (common emitter configuration). It is associated quantitatively with the current generator connected to the terminals of the emitter–base junction in Fig. 236 and is accounted for by the forward bias of the collector–base junction.

Intrinsic resistance of the base. A third static element occurring in the limiting operating conditions of the transistor (from the point of view of rapidity of switching) is defined by the intrinsic resistance of the base represented by r_{bb}' in Fig. 237.

Dynamic parameters

Capacitive effects. When the transistor is cut-off, the barrier capacitances $C_{b'e(b)}$ and $C_{b'c(b)}$ affect the behaviour of the circuit (Fig. 238).

The diffusion capacitive effects, which are dependent on the forward bias conditions of the junctions, appear in the study of the three other operating states of the transistor.

For example, a forward bias of the emitter–base junction involves the use of a capacitor whose value is equal to $C_{b'e(d)}$ (Fig. 239). Similarly, the forward bias of the base–collector junction implies the representation of a capacitor of value equal to $C_{b'c(d)}$ in Fig. 240.

Cut-off frequency of the current gain (*normal state*). The principle underlying the calculation of the cut-off frequency of the current gain of a transistor has been dealt with fully in a book entitled 'Radio-Frequency Transistors'. The determination of this cut-off frequency will be made easier by reference to Fig. 241. The forward bias of the emitter–base junction is shown diagrammatically by the association in parallel of a low value resistance r_{ie} and a large capacitance C_{eb}'. The cut-off frequency is the frequency at which the emitter current is divided into two equal parts, namely:

$$r_{ie} = 1/\omega C_{eb}'_{(d)}.$$

It is easy to obtain from this expression, on the one hand the cut-off frequency of the current gain (common base configuration):

$$f = 1/2\pi r_{ie} C_{eb}'_{(d)}$$

and on the other hand, the cut-off pulsatance:

$$\omega = 1/r_{ie} C_{eb}'_{(d)}.$$

Cut-off frequency (*inverse state*). If the cut-off pulsatance in normal state is of great interest in the quantitative expression for the time constant of the rise time of the current, the cut-off pulsatance in inverse state is no less important for the calculation of the time constant of desaturation. In this connection the reader is referred once more to the simplified equivalent circuit in Fig. 242.

The forward bias of the collector–base junction is shown there by the connection in parallel of a low resistance r_{ic} with a large capacitance $C_{cb}'_{(d)}$. The cut-off frequency of the current gain, with inverse state of the transistor, corresponds to the frequency for which the current is divided into two equal parts:

$$r_{ic} = 1/\omega_{inv} \cdot C_{cb'}$$

This expression will give us the cut-off frequency of the current gain with inverse state:

$$f = 1/2\pi r_{ic} C_{cb}'_{(d)}$$

And it also gives us the value of the cut-off pulsatance of the current gain with inverse state:

$$\omega_{inv} = 1/r_{ic} C_{cb}'_{(d)}.$$

Time parameters

Time constant of the rise time of the current (*current drive*): $\tau_C = h_{FE}/\omega$.

Time constant of the rise time of the current (*voltage drive*)

$$\tau_v = r_{bb}'/\omega \cdot q(-I_C)/kT$$

q electron charge, k Boltzmann constant, T temperature in degrees Celsius.

Time constant of desaturation

$$\tau_s = (1/\omega + 1/\omega_{inv})/(1/h_{FE} + 1/h_{FE\,inv})$$

Calculation of the capacitances

$$C_{TE} \neq 2C_{be}'_{(b)}; \qquad C_{TC} \neq 2C_{bc}'_{(b)}; \qquad C_{TC}' \neq C_{bc}'_{(b)}$$

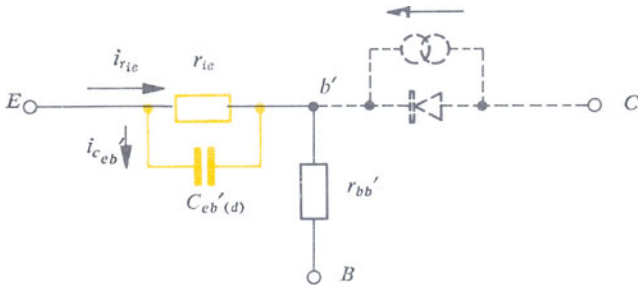


Fig. 241

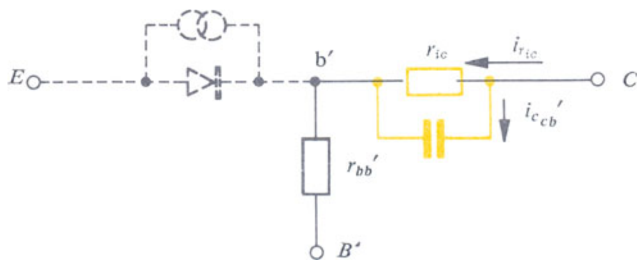


Fig. 242

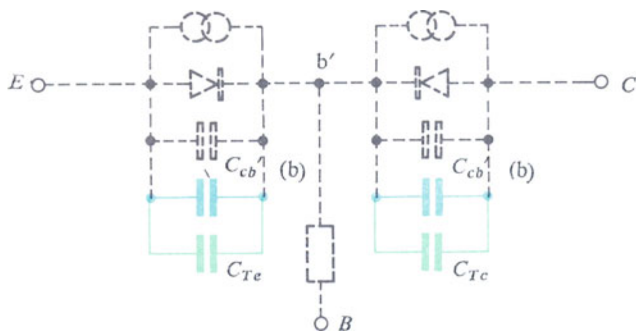


Fig. 243

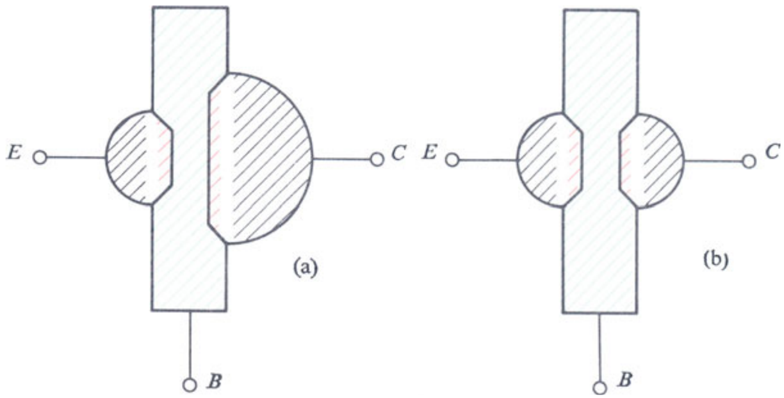


Fig. 244

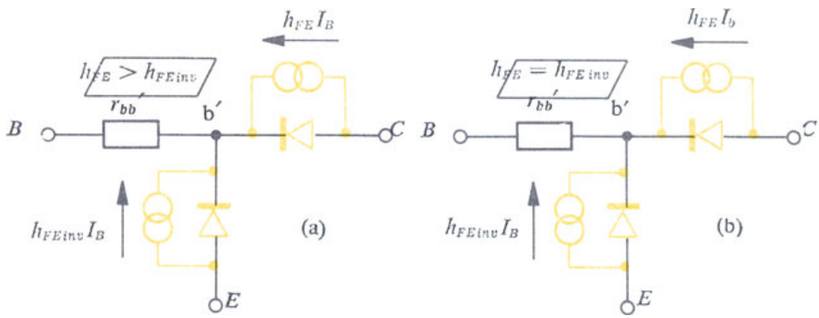


Fig. 245

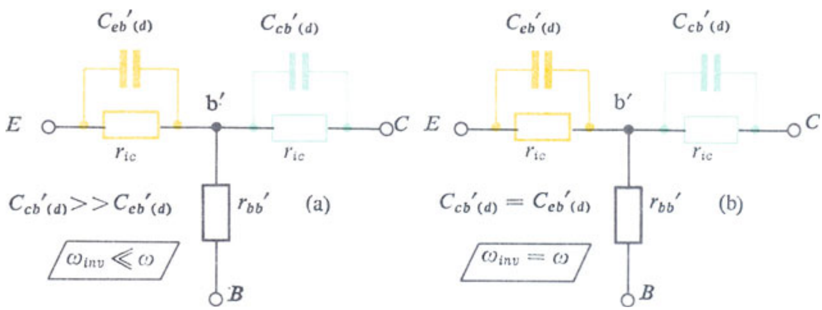


Fig. 246

2.16. Symmetrical transistors

The use of symmetrical transistors is very often preferable to that of asymmetric transistors. How do these two types differ?

Principle of the symmetrical transistor

An asymmetric transistor is shown in Fig. 244a; the purpose of the larger surface of the base–collector junction is to improve the amplification factor when the transistor is operating normally. On the other hand, in an inverse state when the holes diffuse from the collector to the base, they reach the emitter in large numbers and the amplification factor in the reverse direction is considerably smaller.

In a symmetrical transistor (Fig. 244b) the surfaces of the two junctions, emitter–base and collector–base, are identical. In normal state a large number of the holes from the emitter recombine on the surface of the base and the corresponding amplification factor is relatively small. In inverse state the recombinations in the base are identical and the current amplification factor in inverse state is exactly equal to the previous one.

Advantages and disadvantages of the symmetrical transistor. What are the advantages and disadvantages of the symmetrical transistor compared with the asymmetric transistor?

The collector current can be calculated from those expressions in which we find the time constant of the rise of the current (τ_c) and the time constant of desaturation (τ_s).

The desaturation time constant is given by the relation:

$$\tau_s = (1/\omega + 1/\omega_{inv})/(1/h_{FE} + 1/h_{FE\,inv})$$

If we assume that:

$$\tau_c = h_{FE}/\omega$$

It is possible to express the value of τ_s with respect to τ_c:

$$\tau_s = \tau_c[(1 + \omega/\omega_{inv})/(1 + h_{FE}/h_{FE\,inv})]$$

Disadvantage of the symmetrical transistor. In a symmetrical transistor the two current amplification factors in inverse and forward states are equal:

$$h_{FE} = h_{FE\,inv}.$$

At the same time the two cut-off pulsatances (common base configuration) and almost equal:

$$\omega \neq \omega_{inv}.$$

In an asymmetric transistor the current amplification factor in normal state is much larger than the current amplification factor in inverse state:

$$h_{FE} > h_{FE\,inv}.$$

Further, the diffusion capacitance of the collector–base junction is much larger than the corresponding capacitance of the emitter–base junction; in these conditions the cut-off pulsatance in inverse state is lower:

$$\omega > \omega_{inv}.$$

In a symmetrical transistor the resistances and capacitances are equal, but this is not true for the two cut-off pulsatances (Fig. 246b).

In these conditions, if we again consider the expression for the desaturation time constant, we see that in a symmetrical transistor:

$$\tau_s = \tau_c.$$

The saturation time constant of a symmetrical transistor is always much greater than the corresponding time constant of an asymmetric transistor.

Advantages of the symmetrical transistor

Example with a low load resistance. The fact that the desaturation time constant and the time constant of the rise of the collector current are identical is an advantage with a low load resistance. In this case we may write:

$$\tau_s = \tau_c[(2+\omega/\omega_{inv})/(1+h_{FE}/h_{FE\,inv})]$$

for:

$$\tau_c' \neq \tau_c.$$

The expression for the total triggering time (that is from saturation to cut-off):

$$t_s + t_f = f(-I_{B1}/I_{B2})$$

is independent of the current amplification factor in normal state of the transistor and of the permanent value of the collector current. The following relation will be sufficient to illustrate this:

$$t_s + t_f = \tau_c \log_e [1+(-I_{B1})/I_{B2}].$$

In these conditions we may say that the total reconduction time of the transistor does not depend on the degree of saturation.

Example of a high load resistance. When the load resistance is high (in blue in Fig. 249) the condition for saturation (Fig. 248) is given by the relation:

$$[(-I_{CS})/(-I_{B2})] < h_{FE}.$$

If when the transistor is saturated there is a sudden variation in the load resistance (load line in green in Fig. 249) the same value of the collector

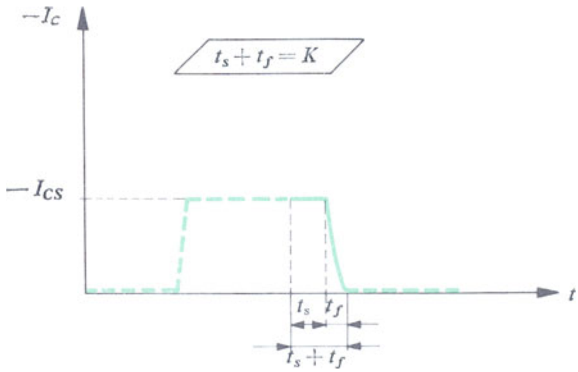


Fig. 247

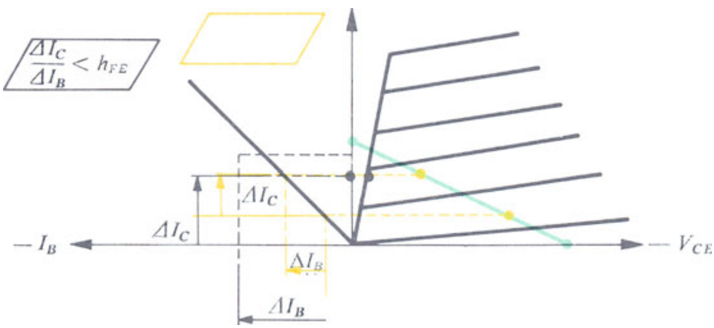


Fig. 248

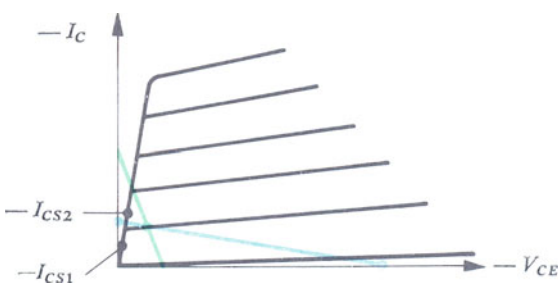


Fig. 249

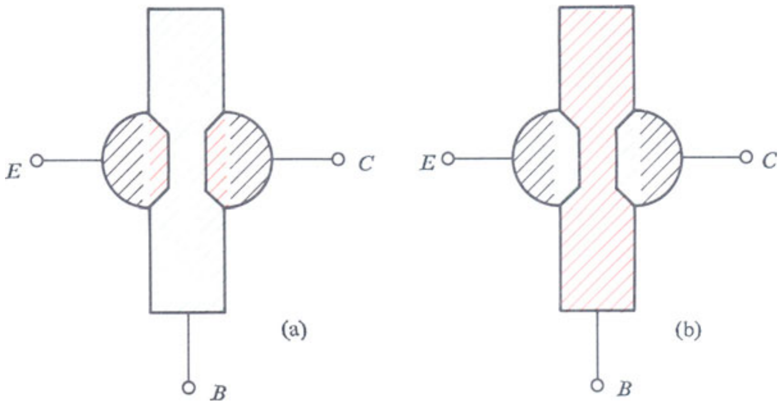


Fig. 250

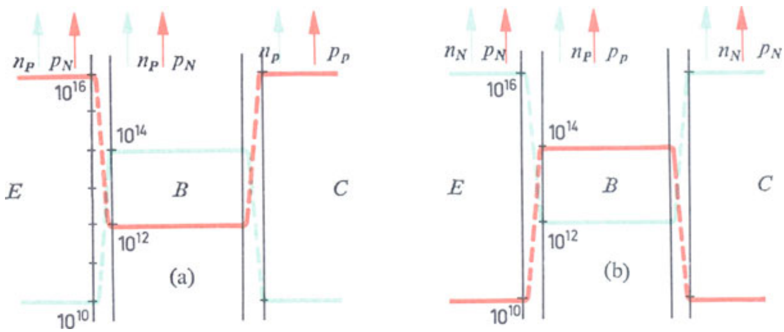


Fig. 251

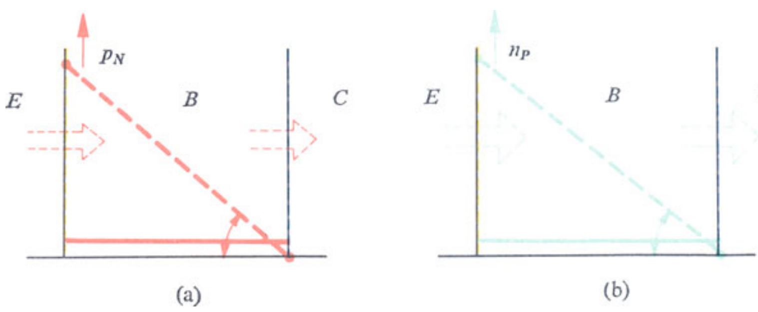


Fig. 252

saturation current $(-I_{CS1})$ no longer corresponds to the initiation of saturation in the transistor. This condition is now respected for a value of collector current equal to $-I_{CS2}$. In this case a base current pulse must be injected such that:

$$-I_{B1} \geqslant \frac{(-I_{CS})}{h_{FE}} \cdot \frac{\tau_c}{\tau_s}.$$

We can then define a current amplification factor in the transient state by:

$$h_{FE(t)} = h_{FE} \cdot \frac{\tau_s}{\tau_c}.$$

In a symmetrical transistor:

$$\tau_s = \tau_c.$$

The two current gains in the static and transient states are practically equal:

$$h_{FE} = h_{FE(t)}.$$

With a high load resistance there is no risk that any decrease in this resistance will, if the switching circuit consists of a symmetrical transistor, cause the saturation to disappear even if it originated in the static state.

2.17. *NPN* transistors

NPN transistors are often met with in switching circuits, and the often consist of symmetrical junctions. How does the *NPN* transistor differ from the *PNP* transistor?

Figure 250a shows a symmetrical *PNP* transistor in which the emitter and collector form the two *P* regions and the base the *N* region. The *NPN* transistor (Fig. 250b) on the other hand contains an emitter and an *N* type collector separated by a *P* type base.

On the majority and minority carrier density graphs in Fig. 251a it will be seen that the two emitter and collector regions contain 10^{16} holes and 10^{10} electrons respectively, whilst the base contains only 10^{14} electrons (majority carriers) and 10^{12} holes (minority carriers). In an *NPN* transistor the emitter and collector contain respectively 10^{16} free electrons (majority carriers) and only 10^{10} holes (minority carriers); on the other hand the base contains 10^{14} holes (majority carriers) and 10^{12} electrons (minority carriers). The barriers separating each region (in the *PNP* or in the *NPN* transistor) are of the same width; the density graphs increase or decrease depending on the direction of observation.

The quantitative study of the currents in the *PNP* transistor were, in the normal state facilitated by the observation of the density graph of the holes in the base of the transistor (Fig. 252a); the hole diffusion current, that is the collector current, is a function of the tangent of the angle between this graph and the horizontal in the neighbourhood of the base–collector barrier. Similarly for an *NPN* transistor we can determine the currents with the aid of the density graph of the electrons in the base of the transistor (Fig. 252b).

In the absence of bias, the density graph is horizontal (full line in the figure); the angle made by this graph is zero and there is no current in the external circuits.

In normal state the emitter–base junction is biased in the forward direction; electrons from the emitter penetrate into the base. Reverse bias of the base–collector junction causes the corresponding barrier to widen, and allows only the minority carriers to flow.

The collector current is thus a function of the number of electrons flowing from the base into the collector, that is of the tangent of the angle between this graph and the horizontal in the neighbourhood of the base–collector barrier.

These two transistors differ not only in the network of density graphs and their physical representation (Fig. 253a and b); we must also know how to use their equivalent circuit and their network of output characteristics.

Equivalent dynamic 'switching' circuits

Figure 254a shows the equivalent dynamic circuit of a *PNP* switching transistor (common base configuration). We have dealt at length with this circuit earlier in this work. It consists of two junctions which may be replaced by resistances depending on their bias condition. It consists also of two current generators associated with the forward bias of the junctions, and three capacitors for each junction (diffusion capacitance, barrier capacitance and transient capacitance). The imaginary point on the base b' is separated from the corresponding electrode by the resistance $r_{bb'}$.

The same components occur in an *NPN* transistor but with an inversion of symbols. In Fig. 254b we see that the two emitter–base and base–collector junctions are reversed with respect to the previous case. Further, the two current generators, which are tied to the forward bias conditions of the junctions, are quantitatively defined by identical expressions but the direction of the arrows is now reversed.

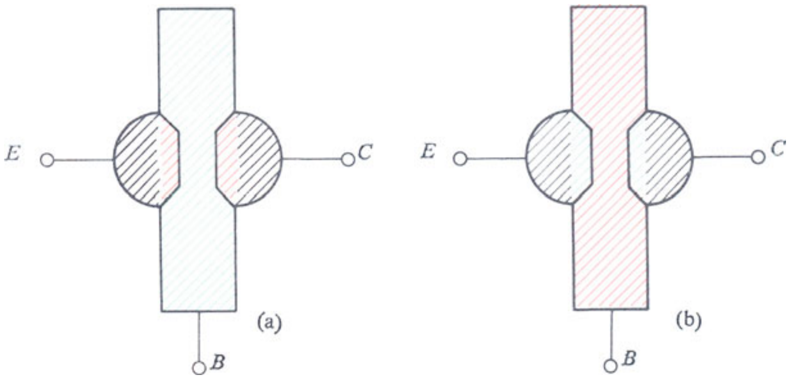


Fig. 253

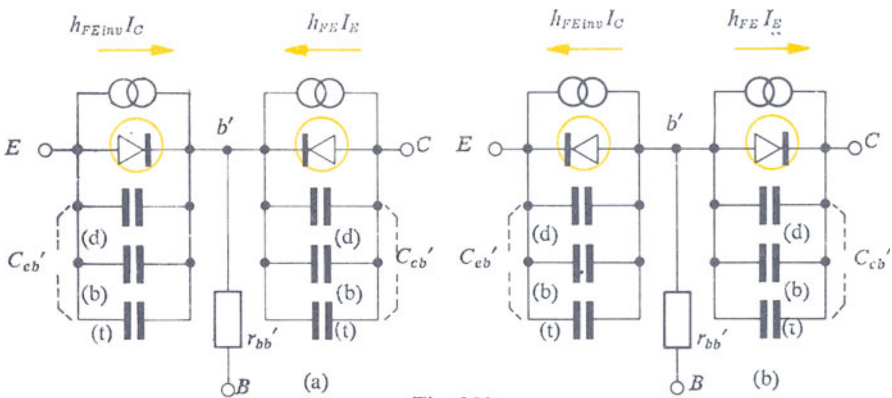


Fig. 254

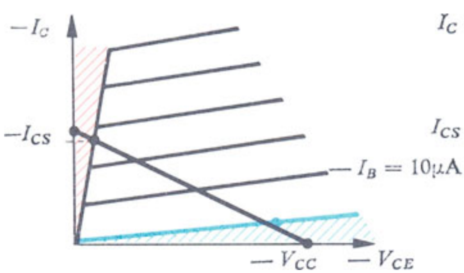


Fig. 255

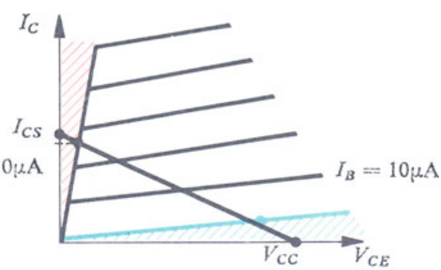


Fig. 256

It should be noted that the normal state of an *NPN* transistor corresponds to two polarities opposite to those in the equivalent state of the *PNP* transistor. Note that in a *PNP* transistor the emitter is positive with respect to the base and the collector is negative, while in an *NPN* transistor the emitter is negative with respect to the base and the collector positive.

Figure 255 shows the network of output characteristics of a *PNP* transistor. The cut-off and saturation states are shown respectively by the 'blue' and 'red' shaded areas.

In an *NPN* transistor the collector current I_C is shown on the ordinate, and the collector–emitter voltage V_{CE} on the abscissa. The two regions, saturation and cut-off are, as before, shown by the shaded 'red' and 'blue' regions in Fig. 256.

Further the symmetrical *NPN* transistor differs from the symmetrical *PNP* transistor in that in the first case the collector current consists exclusively of holes, and in the second of electrons.

Given the greater mobility of free electrons compared with that of mobile holes, it would be logical to consider that they are geometrically identical, and that a *NPN* transistor might imply higher cut-off frequencies. However, it is technically possible to produce *PNP* transistors with very high cut-off frequencies, and this frequency aspect alone constitutes the basic parameter for use in rapid switching.

3. Influence of the Load

The type of drive of a transistor affects the shape of the current in the collector circuit; current drive and voltage drive have different effects on the output current of the transistor with both small and large signals.

Types of load

The type of the load of a transistor affects also the shape of the signal received in the output circuit. To enable us to form a complete picture of the possible uses of a transistor with rectangular signals, it is very important that the effect of the loads should be discussed.

There are three basic types of load for a switching stage:

(a) a purely resistive load;

(b) a capacitive and resistive load;

(c) an inductive and resistive load.

A resistive load involves the use of a pure resistance in the output circuit of the transistor, that is between the collector and the negative pole of a group of batteries $-V_{CC}$ (Fig. 257).

The transistor is driven by a generator and its internal resistance R_i; it is biased by a battery $-V_{BB}$.

A definition of the operation of a system of this sort is equivalent to determining the variations of the collector current with time (provided we know the signal supplied by the group of generators consisting of the generator of rectangular signals and the battery V_{BB}).

A purely resistive load is rarely met with in industrial circuits. The actual load of a transistor often consists of a system of diodes feeding connecting units associated with the inputs of the stages immediately following the transistor under study. These different units, diodes and transistors which make up the load, may act capacitively or inductively. We will first consider capacitive behaviour.

Figure 258 shows a transistorized switching stage, the output circuit of which is loaded by the connection in parallel of a capacitor C and a resistance R_L.

The drive conditions are the same as in the previous case, and the problem therefore is to define the output signal, to determine the currents flowing on the one hand through the load resistance, and on the other hand through the capacitor, and thus to obtain the collector current of the transistor.

In the third case (Fig. 259) the transistor is loaded by an inductance in series with a resistance. There is now no need to determine the different values of the current (since the same current flows through the resistance and inductance) but it is very important to be able to calculate, in addition to the shape of the collector current, the instantaneous values of the voltage at the terminals of the resistance and inductance, which constitute the output voltage of the stage, that is the driving voltage of the following stage.

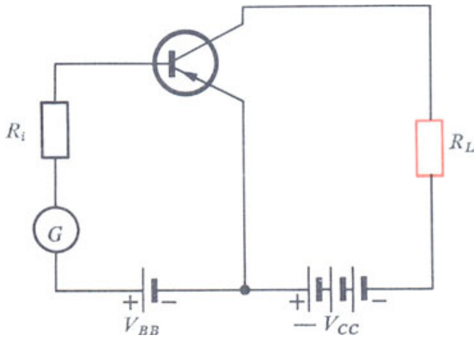


Fig. 257

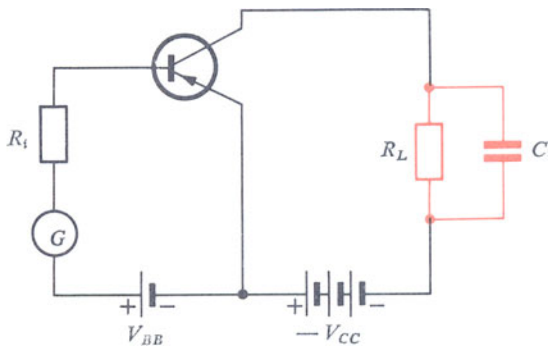


Fig. 258

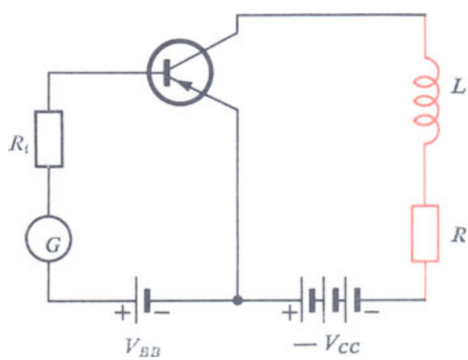


Fig. 259

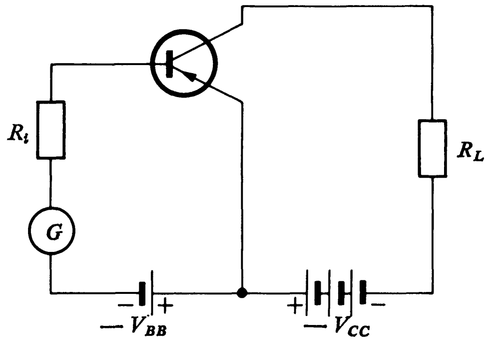


Fig. 260

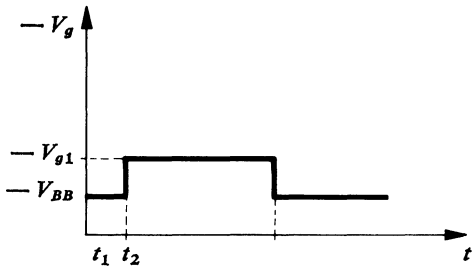


Fig. 261

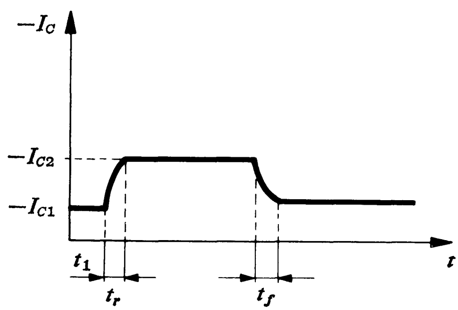


Fig. 262

3.1. Resistive load

In order to describe the behaviour of a transistorized switching stage (Fig. 260), knowing the shape of the input signal, we must determine the shape of the output signal by calculating the variation of the collector current with time.

This circuit consists of a transistor driven by a generator of rectangular signals and its internal resistance, the combination being connected in series with a biasing battery $-V_{BB}$.

On the output side the transistor is loaded by a resistance R_L and fed by a group of batteries $-V_{CC}$.

A full study of the different shapes of the current liable to flow in the output circuit of the stage must include discussion of three possible conditions:

(a) the transistor operating permanently in the normal state;
(b) the transistor switched from the cut-off state to the normal state and back to the cut-off state;
(c) the transistor switched from the cut-off state to saturation and back to the cut-off state.

Permanent normal state

In this case, given the bias conditions of the input junction of the transistor (base–emitter junction), the stage operates permanently in the normal state.

From instant t_1 to the instant preceding t_2 a voltage $-V_{BB}$ is applied between the base and emitter of the transistor (Fig. 261). The collector current assumes the value $-I_{C1}$ shown in Fig. 262.

At instant t_2 the value of the voltage changes suddenly from $-V_{BB}$ to $-V_{gi}$, and the collector current increases very slowly from $-I_{C1}$ to $-I_{C2}$ as shown in Fig. 262. The rise time of the current is shown by t_r on the abscissa.

From instant t_2 to instant t_3 the signal is maintained and the current remains equal to $-I_{C2}$.

At instant t_3 the voltage supplied by the group of generators decreases; the current starts to fall and reaches the final value a certain time later. The time required for the fall of this current is given by t_f on the abscissa in Fig. 262.

The rise time of the current (t_r) and the fall time (t_f) have been described above. But it is not necessary to recall here the already known quantitative expressions.

When the transistor operates permanently in the normal state the triggering times are taken as being practically equal.

Change from the cut-off to the saturation state

The circuit conditions of the transistor shown in Fig. 263 are practically identical with those in the previous case.

The input circuit consists of a battery V_{BB} (the positive pole of which is connected to the base side of the transistor) connected in series with a generator G. The load circuit consists of the resistance R_L connected between the negative pole of the group of batteries $-V_{CC}$ and the collector of the transistor.

From instant t_1 to the instant preceding t_2, the generator G delivers no signal; the battery supplies a positive voltage between the base and the emitter of the transistor, which is in the cut-off state, and the collector current is zero (Fig. 264).

At instant t_2, the generator supplies a high negative voltage, and the resultant voltage at the input terminals ($-V_{g1}$) is such that the transistor approaches saturation. It is however necessary at this instant, in order to allow the diffusion of holes from the emitter into the base, for the base–emitter barrier to become narrower; this accounts for the existence of a delay time of the current t_d shown in Fig. 265.

The hole density graph then starts to turn from the horizontal position; the current rises very rapidly up to the value $-I_{CS}$. t_r then gives the rise time of the current in the circuit. The current remains at this value until instant t_3.

At instant t_3 the signal supplied by the generator disappears and the transistor approaches the cut-off state.

The surplus holes stored in the base should be removed: the current remains constant and equal to $-I_{CS}$ during the time t_s shown in Fig. 265.

The graph may then finally assume its horizontal position and the current falls. The fall time of the current t_f is then very short.

The formulae for the calculation of the delay and triggering times can be derived from the above discussion or from data sheets on transistors.

There are now four of these times:

(a) the delay time of the current, t_d;

(b) the rise time of the current, t_r;

(c) the desaturation time, t_s;

(d) the fall time of the current, t_f.

We must therefore include the two delay times, in addition to the triggering time, when switching a transistor from the cut-off to the saturation state and vice versa.

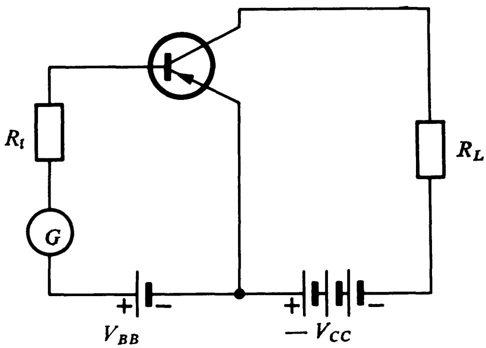


Fig. 263

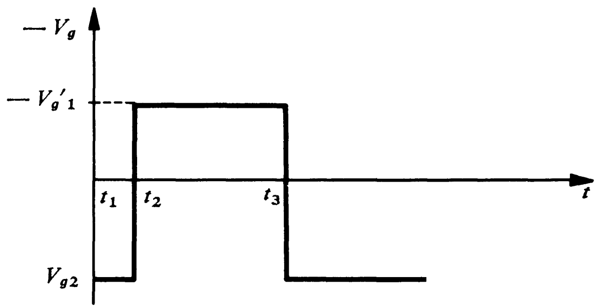


Fig. 264

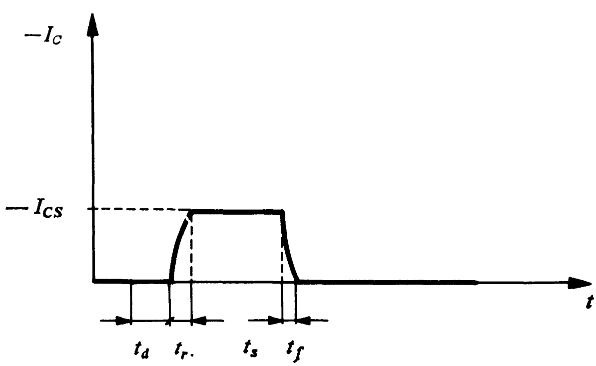


Fig. 265

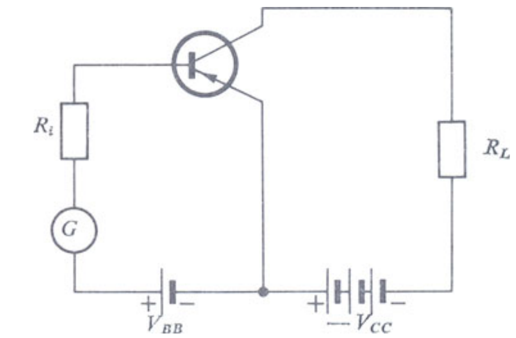


Fig. 266

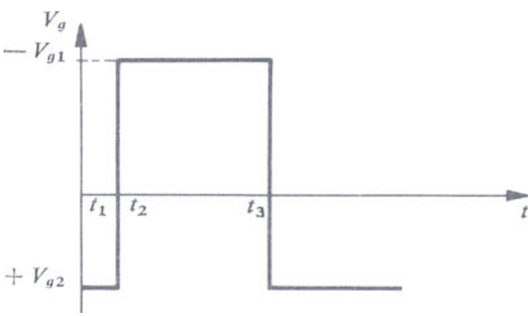


Fig. 267

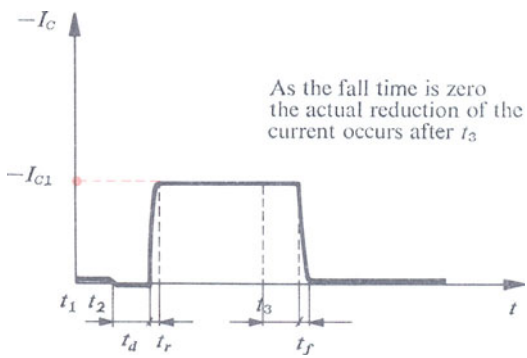


Fig. 268

Change from the non-conducting to the normal state

The second possible study, corresponding to a purely resistive load, consists in the examination of the behaviour of a switching transistor when the voltage supplied by the generator causes the transistor to pass from the cut-off to the normal state and vice versa.

In Fig. 266 the transistor is driven by a generator and its internal resistance (R_i) and the positive pole of the battery. The load resistance R_L is still in the collector circuit in series with the group of batteries $-V_{CC}$.

From instant t_1 to the instant preceding t_2 the generator supplies no signal; the base–emitter junction is biased in the reverse direction only by the battery, and the transistor is non-conducting (Fig. 264).

The current in the collector circuit is zero (Fig. 268). At instant t_2 a strong negative voltage pulse supplied by the generator tends to cause the transistor to pass to its normal operating state.

The collector current does not rise immediately. The thickness of the emitter–base barrier must first be reduced. The time required for this process is given by t_d in Fig. 268.

From the instant when the thickness of the barrier is at its new value, the holes penetrate into the base and the collector current rises to reach its final value (Fig. 268). The collector current then remains at a constant value equal to $-I_{C1}$.

When at instant t_3 the transistor is triggered from the normal to the cut-off state, the collector current immediately starts to decrease and very soon disappears; t_f in Fig. 268 gives the fall time of the current in the circuit.

As in the previous case, it is not necessary to enter into the detail of the quantitative expressions for the calculation of the different delay and triggering times.

We must however remember that there are three of these times:

(a) the delay time of the current, t_d;

(b) the rise time of the current, t_r;

(c) the fall time of the current, t_f.

When the transistor is switched from the cut-off state to normal operation, we must include the delay time of the current as well as the triggering time.

3.2. Capacitive and resistive load

The load on a transistor is rarely purely resistive; it often consists of the input of the following stage which often acts as a capacitor associated with a resistance. We will now consider the operation of a transistor driven by a rectangular signal when its load is equivalent to a resistance R_L and a capacitor C in parallel (Fig. 269).

The driving circuit is as before defined by the connection of a battery $-V_{BB}$ and a generator of rectangular signals in series with its internal resistance.

For the study of the operation of a circuit of this type it would be interesting to determine:

(a) the current in the load resistance $-R_L(-I_{RL})$;

(b) the current in the capacitance section $(-I_C)$;

(c) the collector current $(-I_{col})$ which is equal to the sum of the above two currents.

In order to obtain the variations with time of the current in the load resistance we must first know the variations of the voltage applied between the base and the emitter of the transistor.

For a voltage equal to $-V_{BB}$ the current in the load resistance is a minimum. At instant t_2 a complementary signal is applied to the input terminals of the stage, and the collector current tends to increase, resulting in an increase of current in the load resistance (Fig. 270b).

This increase in the current and in the corresponding rise time (that is the time required for the current to increase from $-I_{CO}$ to $-I_{C1}$) are marked in yellow in the figure.

How does the capacitance affect the shape of the current in the load resistance?

An increase in capacitance results in a change in the rise time of the current. When C_3 is greater than C_2 which is itself greater than C_1, t_{r3} is longer than t_{r2}, which is itself longer than t_{r1}.

Similarly when the voltage changes at instant t_2 from $-V_{BB}$ to $-V_{g1}$, a charging current appears in the capacitive section until instant t_C (Fig. 271). The current then decreases to zero.

The charging of this capacitor is a simple function of the tangent of the angle in Fig. 271, which is represented by tan φ.

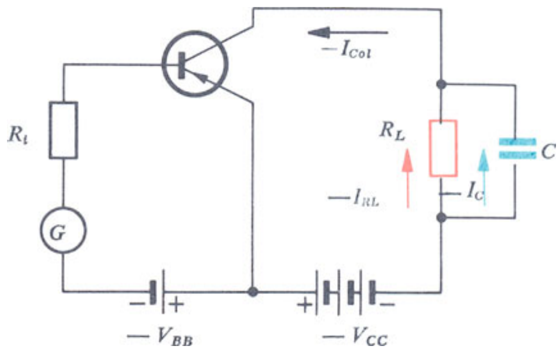


Fig. 269

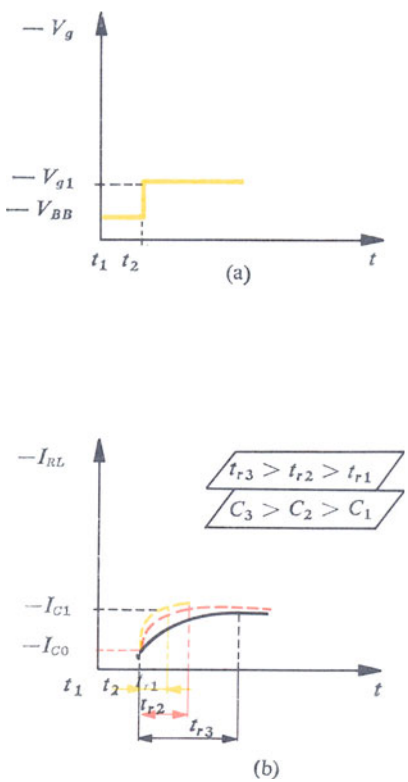


Fig. 270

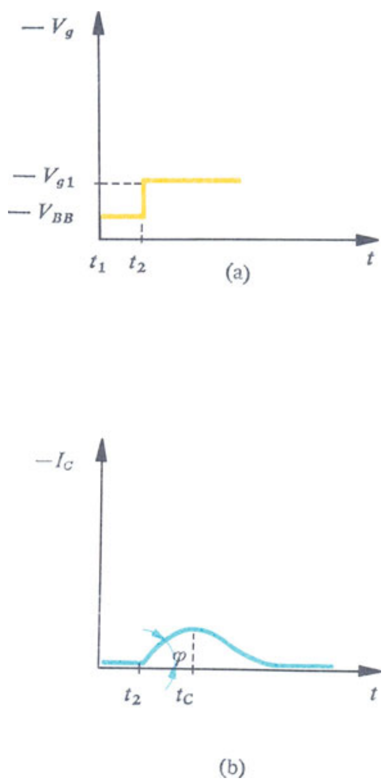


Fig. 271

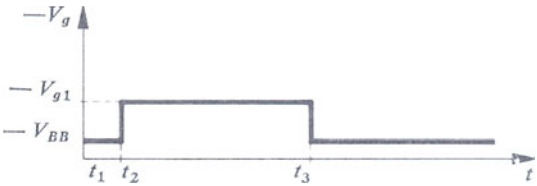


Fig. 272

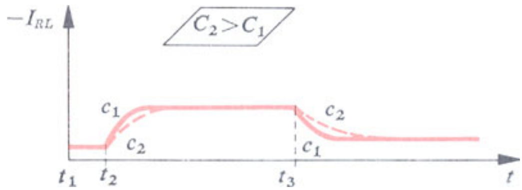


Fig. 273

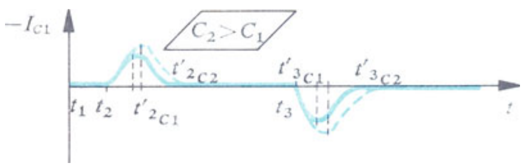


Fig. 274

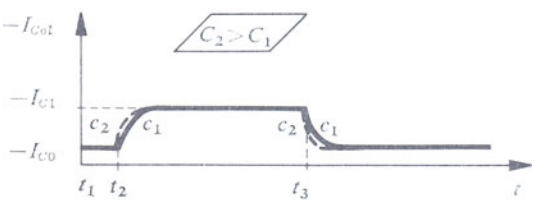


Fig. 275

Normal state

In this type of state the transistor operates permanently in the small-signal condition. When the voltage supplied by the group of generators varies from $-V_{BB}$ to $-V_{g1}$, the operation of the transistor changes from the normal state to a more pronounced state.

Figure 272 shows the variations with time of the voltage at the terminals of the group of generators.

From instant t_1 to the instant preceding t_2, the generator supplies no signal and the resultant voltage at the terminals of the group of generators is equal to the battery voltage $-V_{BB}$.

The current in the load resistance R_L (Fig. 273) is a minimum. No current flows through the capacitive section, as the system is in continuous state (Fig. 274). The current in the collector circuit is therefore exactly equal to the current flowing through the resistance R_L, namely $-i_{CO}$ in Fig. 275.

At instant t_2 the signal changes suddenly from its minimum value to $-V_{g1}$ (Fig. 272). The current in the load resistance starts to rise more or less rapidly as a function of the capacitance.

The rise time of the current in the resistance becomes shorter as the value of the capacitance decreases (Fig. 273). The variation of the current in the circuit produces a charging current which increases very rapidly until instant t_2' and then decreases until it reaches zero (Fig. 274).

An increase in the capacitance causes an increase of this current pulse. The rise time of the collector current, given by the sum of the currents flowing through the resistive and capacitive sections, increases as the value of the capacitance increases (Fig. 275).

When the collector current reaches its final value ($-I_{C1}$) and for the same applied voltage, the current in the resistance remains constant; its value in the capacitive section is zero.

At instant t_3, the voltage supplied by the generator is interrupted (Fig. 272). The current in the load resistance starts to fall at a rate proportional to the decrease in the value of the capacitance C (Fig. 273).

The variation of the collector current in the other direction causes the appearance of a reverse current pulse in the capacitive section passing through its maximum at instant t_{3C1} (Fig. 274).

An increase in capacitance causes an increase in the reverse current pulse and at the same time in the time corresponding to the maximum value of this current ($t_3'_{C2}$ longer than $t_3'_{C1}$).

Transition from the cut-off to the normal state

When the transistor is switched from the cut-off to a normal state and vice versa, the shapes of the different currents in the resistive and capacitive sections are very different from those in the previous example.

Let us consider the circuit in Fig. 276. It consists of a transistor driven by a generator associated with its internal resistance; the bias is obtained from a battery, the positive pole of which is connected to the base; the load still consists of a resistance R_L and a capacitor C in parallel; further, the current supply comes from a group of batteries $-V_{CC}$.

The changes to the signals caused by the switching of the transistor from a cut-off to a normal state occur:

(a) when the current starts to flow in the circuit;

(b) when this current disappears.

At instant t_2 the voltage supplied by the group of generators changes suddenly from the value V_{BB} to the value $-V_{g1}$ (Fig. 277); no current flows in the resistance because the holes cannot yet diffuse into the interior of the base of the transistor.

The time required for the narrowing of the barrier is then given by the delay time $(t_2'-t_2)$ in Fig. 277b. Since there is no current in the collector circuit, the capacitor C has no influence and the current through this section remains practically at zero (from t_2 to t_2' in Fig. 277c).

At instant t_3 the signal changes suddenly from $-V_{g1}$ to V_{BB} (Fig. 278a) The transistor is then switched to the cut-off state. In this case the current in the load resistance tends to decrease as the collector current decreases; it decreases relatively slowly, reaching the final value after a certain time (Fig. 278b). At instant t_3 the variation of the collector current causes the appearance of a strong current pulse in the capacitive section; the first part of this variation is practically linear.

The maximum value is reached at instant t_3' and the current in the capacitive section then varies exponentially (Fig. 278c).

When a transistor is switched from a cut-off to a normal state and back to the cut-off state (the transistor being loaded capacitively), the output signal is affected:

(a) at the establishment of the collector current by the appearance of a delay time;

(b) at the disappearance of the collector current by the presence in the capacitive section of a current pulse whose shape is practically linear during the rise time.

Any variation in the capacitance causes a modification of the various rise and interruption times of the current in the circuit.

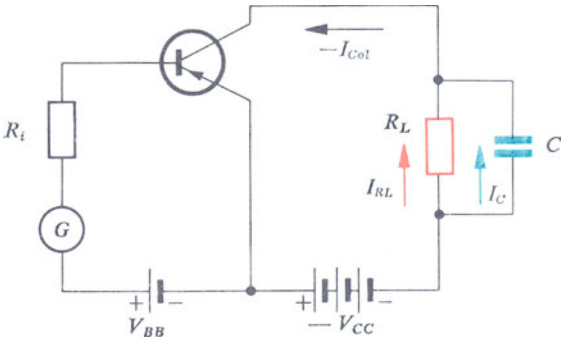


Fig 276

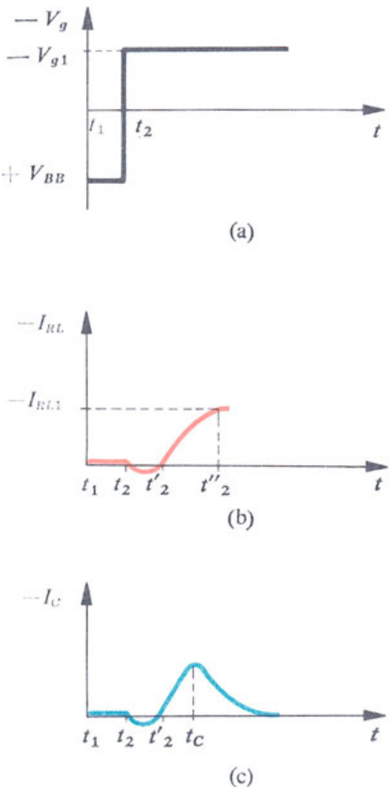


Fig. 277

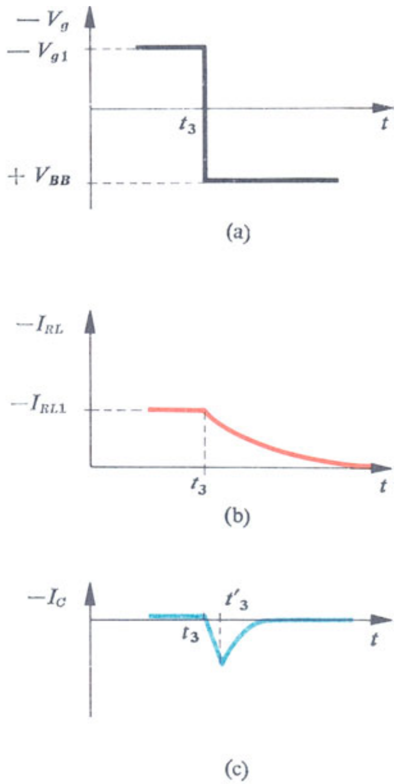


Fig. 278

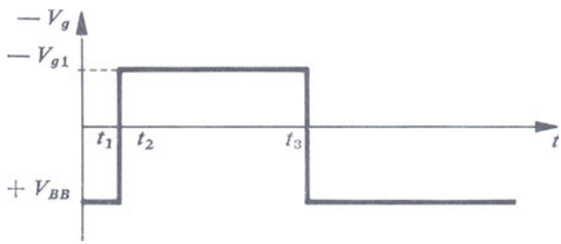


Fig. 279

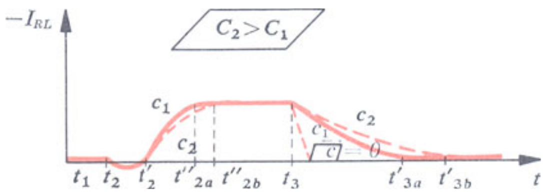


Fig. 280

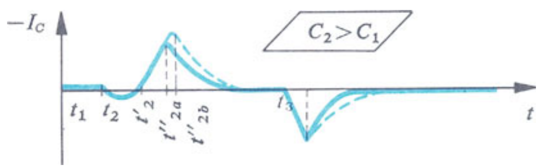


Fig. 281

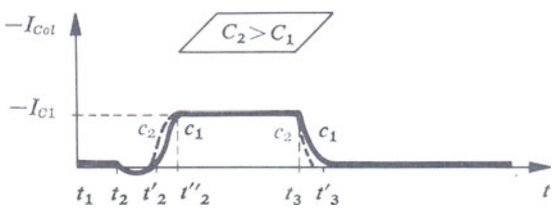


Fig. 282

Shape of the currents. From instant t_1 to the instant preceding t_2 the generator of rectangular signals supplied no voltage; the transistor is therefore non-conducting (Fig. 279), and the current in the load resistance is zero (Fig. 280). The same applies to the currents in the capacitive section (Fig. 281) and in the collector of the transistor (Fig. 282).

At instant t_2 the signal supplied by the group of generators changes suddenly from the value V_{BB} to the value $-V_{g1}$ (Fig. 279). As the transistor was in the cut-off state, no current flows in the output circuit, and the width of the barrier should at first decrease. A delay time then appears when the current in the load resistance R_L is established (namely from t_2 to t_2' in Fig. 280); the current in the capacitive section is zero ($t_2'-t_2$ Fig. 281) and the current in the collector circuit is also near to zero (Fig. 282).

The current in the collector circuit starts to rise from instant t_2'; the corresponding variation occurs from zero to $-I_{C1}$, its final value shown in Fig. 280. The capacitance affects the rise time of the current; any increase in its value (namely C_2 larger than C_1) is associated with an increase in the rise time of the current in the load resistance (t_{2b}'' larger than t_{2a}'').

A current pulse appears in the capacitive section and passes through a maximum value at instant t_{2C1}''; the current then approaches zero. An increase in the capacitance (namely C_2 larger than C_1) causes a strengthening of the current pulse and at the same time a lengthening of the time dependent on its maximum value. As the collector current is at every instant defined by the sum of the two former currents, it will become established more rapidly as the value of the capacitance C_2 increases (Fig. 282).

Until instant t_3, the currents remain at the same value in the load resistance and in the collector of the transistor, but there is no current in the capacitive section.

At instant t_3, the sudden change of voltage from $-V_{g1}$ to the value $+V_{BB}$ (Fig. 279) causes a larger decrease in the current in the load resistance as the value of the capacitance decreases (Fig. 280). A current pulse appears at the same time in the capacitive section, but it rapidly decreases and finally disappears. The speed of this decrease is greater as the value of the capacitance decreases. Consequently the speed of decrease of the collector current of the transistor from its value $-I_{C1}$ to 0 (that is C_2 larger than C_1 in Fig. 282), decreases as the value of the capacitance increases.

Change from the cut-off to the saturation state

The switching of a transistor from the cut-off to the saturation state represents one of the most commonly occurring applications in logic circuits and in industrial electronics.

The circuit shown in Fig. 283 is not very different from the previous one except for the amplitude of the input signal during the period of conduction.

How does the switching of the transistor to saturation affect the shape of the currents in the collector of the transistor, in resistance R_L and in the capacitance C?

From instant t_1 to the instant preceding t_2, the generator G supplies no signal and the transistor is cut-off signal; the currents in the resistance R_L, in the capacitor and in the collector of the transistor are zero.

At instant t_2 the voltage changes suddenly from a value corresponding to the cut-off state to a value at which the transistor will be saturated.

This change does not cause, however, any variation of the current in the load resistance R_L at this instant; the current in the resistance is closely associated with the flow of holes from the emitter to the base of the transistor, but this could not occur except for a much narrower emitter–base barrier (Fig. 285).

On the other hand, a current pulse appears immediately in the capacitive section; it approaches a maximum value near to the saturation value of the collector current ($-I_{CS}$ in Fig. 285). This current then decreases until it disappears.

If we wish to define the variation with time of the collector current and if we remember that it is equal to the sum of the two previous currents, it is obvious from Fig. 287 that the collector current exactly follows the variations of the current in the capacitive section.

From instant t_2' (with $t_2'-t_2$ defining the delay time of the current in the load resistance) the current in the resistance increases with time and rapidly reaches a maximum value equal to the saturation current ($-I_{CS}$ in Fig. 285).

At instant t_2' the current in the capacitive section is rising steadily. If we add the currents in the resistive and capacitive sections we see that the value of the collector current exceeds the saturation value and reaches a maximum at the instant when the currents in the two sections are equal (Fig. 287).

The large decrease in the current in the capacitive section is no longer compensated by the increase in current in the resistive section, and the collector current decreases to reach its final value at the saturation value.

In addition to the effect of the delay in the fall time of the current (change from saturation to cut-off), the transient from the cut-off to the saturation state of a transistor loaded by a capacitor associated with a resistance causes a current pulse to appear in the collector, the value of which may exceed the saturation value.

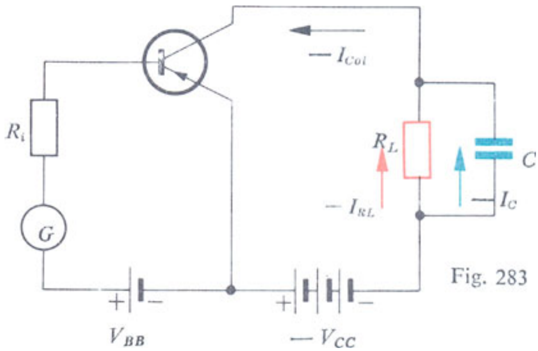


Fig. 283

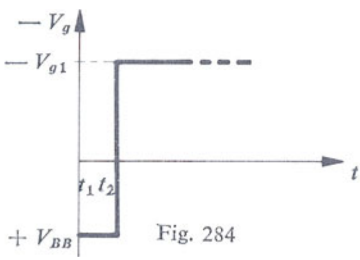


Fig. 284

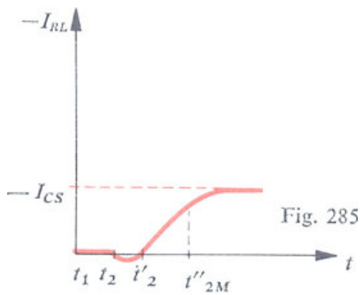


Fig. 285

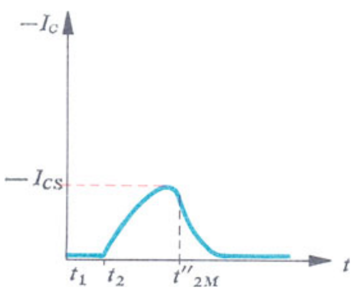


Fig. 286

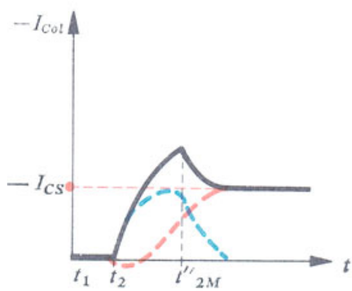


Fig. 287

Shape of the currents. From instant t_1 to the instant preceding t_2, the voltage supplied by the group of generators causes the transistor to be in

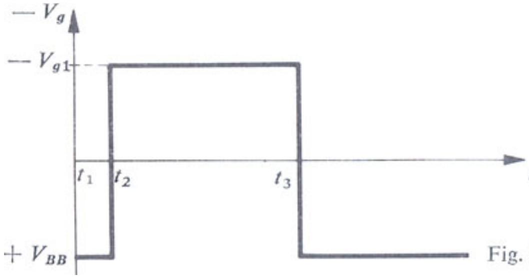


Fig. 288

the cut-off state ($+V_{BB}$ Fig. 288). The current in the load resistance is zero (Fig. 289), and the same applies to the currents in the capacitive section

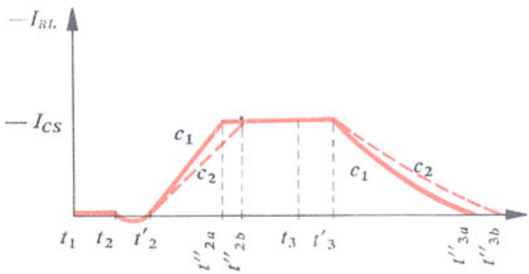


Fig. 289

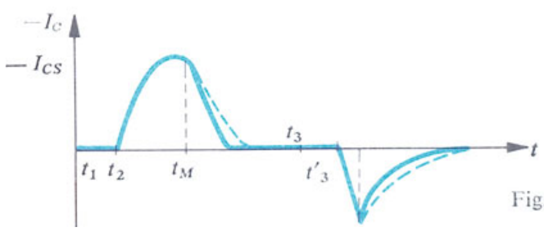


Fig. 290

(Fig. 290) and in the collector of the transistor (Fig. 291).

At instant t_2 a very large voltage variation tends to switch the transistor

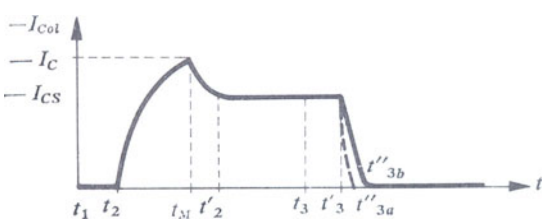


Fig. 291

from the cut-off to the saturation state (Fig. 289). No current flows in the load resistance from instant t_2 to the instant t_2', which is the time required for the reduction in the width of the base–emitter barrier. It is, however, possible that there will be a small negative current pulse due to the arrival of the charges required for the reduction in the width of the barrier (Fig. 289).

On the other hand, a current pulse appears in the capacitive section; it approaches a maximum value $-I_{CS}$, reached well after instant t_2' (Fig. 290).

As the collector current is defined as the sum of the two previous currents, it will consequently follow the variations of the current in the capacitive section and will increase in accordance with the curve shown in Fig. 291.

From instant t_2' a current appears in resistance R_1. It rises almost linearly up to its saturation value $-I_{CS}$.

An increase in the capacitance causes a lengthening of the time required for the establishment of this current (with C_2 larger than C_1 and t_{2b}'' longer than t_{2a}'', Fig. 289). The current in the capacitive section continues to increase (Fig. 290). The collector current then increases to a value greatly in excess of the saturation value $-I_{CS}$.

This current reaches its maximum value ($-I_C$) at instant t_M shown in Fig. 291. This time t_M is given by the equation:

$$t_M = t_d \log_e (\tau_c / t_d).$$

Due to the big decrease in the current in the capacitive section the collector current decreases and rapidly returns to the saturation value (instant t_2'). Any increase in the capacitance causes a lengthening of the fall time of the current in the capacitive section (dotted lines in Fig. 290).

From instant t_2' until instant t_3 the current remains at the saturation value in the load resistance and in the collector, but it is zero in the capacitive section.

At instant t_3 a large variation in the voltage in the other direction (Fig. 288) produces no variation in the current in the collector and in the load resistance and capacitive circuits. The desaturation time ($t_3' - t_3$) is dependent on the removal of the surplus charges stored in the base of the transistor.

From instant t_3' the current decreases in the resistance and disappears at instant t_3' (Fig. 289). An increase in the capacitance causes an increase in the fall time of the current.

A reverse current pulse appears at instant t_4' in the capacitive section and then varies linearly and decreases fairly quickly.

The collector current decreases as the value of the capacitance increases (Fig. 291).

3.3. Inductive and resistive load

In transistorized switching circuits we often meet with the type of load shown in Fig. 292 which consists of an inductance and a resistance R in series. In order to determine the available energy at the output of the stage we must first determine the value of the collector current, and then the variations of the voltage appearing at the terminal of the load (voltage variations represented by V_L in the figure).

An inductance, just as a capacitance, stores energy and restores it. For this reason with inductive loads we must differentiate between the following three possibilities:

(a) under-damping; (b) critical condition; (c) over-damping.

These three cases are associated with quantitative expressions which we will now discuss. They have a very important effect on the shape of the signals received at the output circuit.

Under-damping. By under-damping we understand the condition expressed by:

$$(1/\omega + R_L C_{b'c})^2 - 4LC_{b'c}/h_{FE} < 0.$$

Critical condition. The critical condition is defined by the relation:

$$(1/\omega + R_L C_{b'c})^2 - 4LC_{b'c}/h_{FE} = 0.$$

Over-damping. Finally, this condition is expressed by:

$$(1/\omega + R_L C_{b'c})^2 - 4LC_{b'c}/h_{FE} > 0.$$

It is therefore very important, if we wish to gain a good knowledge of the shape of the output signal, to have a precise idea of the operating condition of a transistor loaded by an inductance.

With under-damping, when the input voltage changes from zero to a value $-V_{g1}$ (Fig. 293a), the variations of the output current give rise to ringing, which are superimposed on the final mean value of the current $-I_C$ (Fig. 293b). The same applies to variations of the output voltage (Fig. 293c).

Since the critical condition is very rarely met with we will not consider it here, but we will deal later in more detail with the over-damping condition.

When the voltage changes from zero to a certain value at instant t_2 (Fig. 294a), the current rises very rapidly and approaches its final value (Fig. 294b). There is no ringing; the variations of the voltage with time at the terminals of the inductance and of the resistance are then as shown in Fig. 294c).

At instant t_2' the voltage rises until it reaches its final value; it remains at this value so long as a signal is applied at the input terminals.

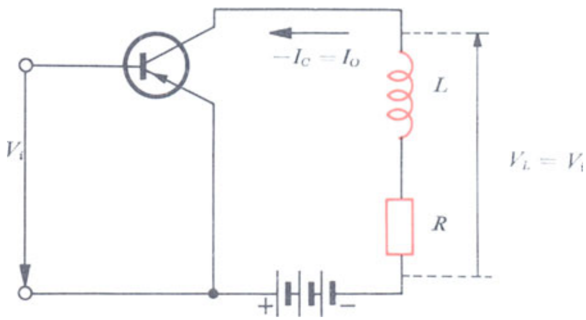


Fig. 292

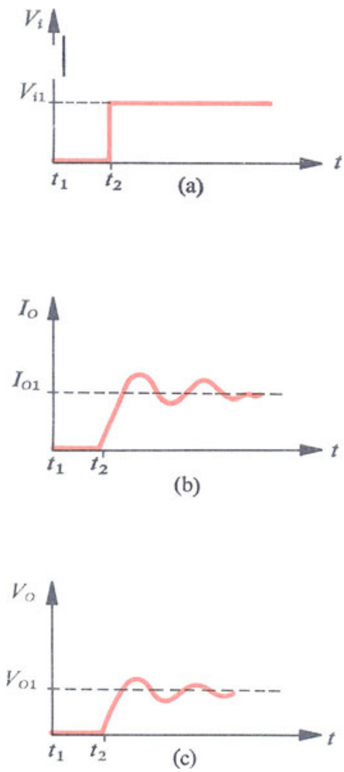


Fig. 293

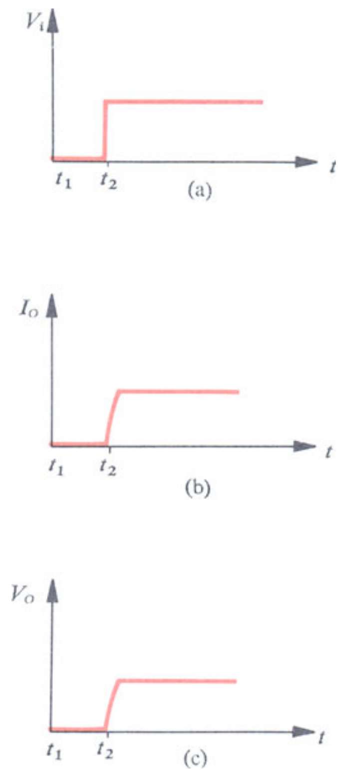


Fig. 294

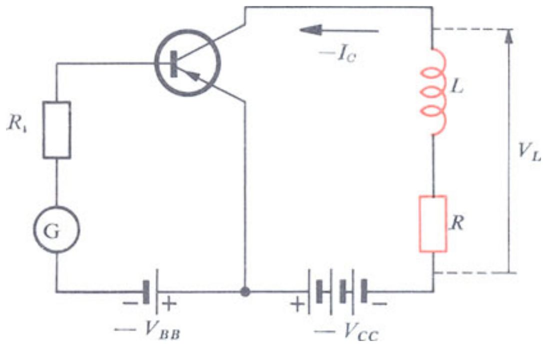


Fig. 295

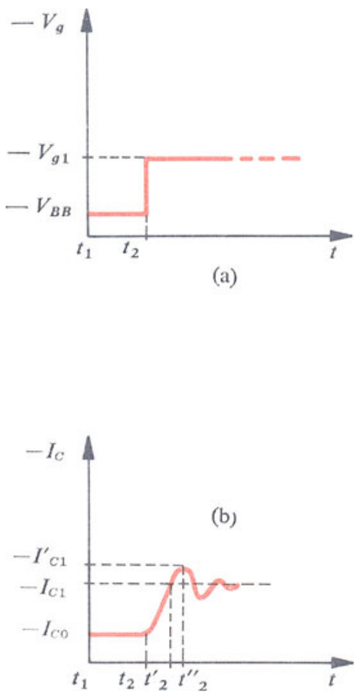


Fig. 296

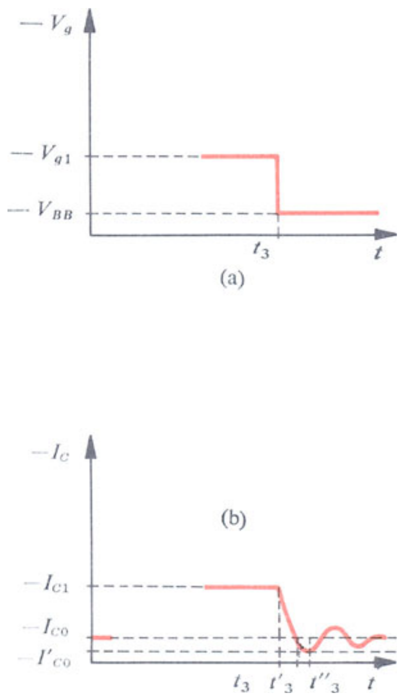


Fig. 297

With the two previous types of load we studied the behaviour of the system in the three following conditions:

(a) small-signal operation with the transistor switched from one state to another, where the two states are within the normal state;

(b) transient from the cut-off to the normal state;

(c) large-signal operation with the transistor switched from the cut-off to the saturation state.

Normal state. We will first consider the operation of the transistor loaded by an inductance and used only in the normal manner.

In the circuit in Fig. 295, the transistor is driven by a group of generators and loaded by an inductance in series with a resistance. The bias battery $-V_{BB}$ applies a steady negative voltage to the base of the transistor.

General study. From instant t_1 to the instant preceding t_2 the generator supplies no signal; the voltage applied by the battery causes the transistor to be in its normal state.

In the case of under-damping (Fig. 296a) in the absence of the signal from the generator, the collector current is equal to $-I_{CO}$ (Fig. 296b).

From instant t_2 the input voltage changes from $-V_{BB}$ to $-V_{g1}$ and the collector current increases. In view of the conditions of under-damping, the ringing effects appear and the collector current which reaches a value $-I_{C1}$ at instant t_2' continues to increase to the value $-I_{C1}'$ (value obtained at instant t_2'' in Fig. 296b).

The current then falls, passes through the value $-I_{C1}$ and falls further; it then rises and so on.... Damped oscillations are superimposed on the mean collector current.

When at instant t_3, the voltage supplied by the group of generators changes from $-V_{g1}$ to $-V_{BB}$ (Fig. 297a) the current falls and very soon reaches the value $-I_{CO}$ corresponding to this new bias condition (instant t_3' in Fig. 297b); it then continues to fall to the value $-I_{CO}$. The previous effect recurs again and the current swings from one side to the other of the value $-I_{CO}$. Damped oscillations then appear in the circuit.

With rectangular signal operation the inductive load produces ringing in the output circuit of the transistor for every increase or decrease in the collector current. This effect, however, occurs only in the case of under-damping.

To recapitulate, when a transistor operates permanently in the normal operating state and when perfectly rectangular voltage variations are applied at its input terminals, there are two cases to be considered:

(a) if the condition of under-damping is respected, the variations of the collector current and of the output voltage are as shown in Fig. 298;

(b) in the case of over-damping the corresponding variations are as shown in Fig. 299.

Under-damping. From instant t_1 to the instant preceding t_2 the signal applied at the input terminals is very small, the collector current is a minimum $-I_{CO}$ (Fig. 298b) and the output voltage is equal to V_{OO} (Fig. 298c).

At instant t_2 the variation of the input voltage (Fig. 298a) causes a rise in the collector current which reaches the final value $-I_{C1}$ at instant t_2' (Fig. 298b). It continues to increase and passes through the maximum value $-I_{C1}$ at instant t_2''; the variations occur in either direction until the current disappears.

The output voltage is affected by the same influences; from instant t_2 the voltage increases, assumes the value V_{O1}, rises above this value to the maximum V_{O1}' and then decreases to rise again to the steady value V_{O1} (Fig. 298c).

At instant t_3, the voltage supplied by the generator decreases (Fig. 298a). The collector current decreases and reaches the minimum value $-I_{CO}$ at instant t_3'; it continues to decrease to the value $-I_{CO}$ (instant t_3'' in Fig. 298b); it then increases and decreases again. As the oscillations are damped the current becomes stabilized at the value $-I_{CO}$ after a certain interval, and the output voltage varies in a similar manner (Fig. 298c).

Over-damping. With over-damping the variations of the current and voltage are much more simply explained. From instant t_2 when the voltage changes from $-V_{BB}$ to $-V_{g1}$ (Fig. 299a), the collector current rises from $-I_{CO}$ to I_{C1}; the time required for the rise is given by $t_2'-t_2$ (Fig. 299b); the same applies to the variations of the output voltage (Fig. 299c).

When the signal disappears at instant t_3, the current decreases; $t_3'-t_3$ gives the fall time of the current in the circuit (Fig. 299b). The output voltage decreases steadily to reach the final value $(-V_{OO})$ a certain time later (Fig. 299c).

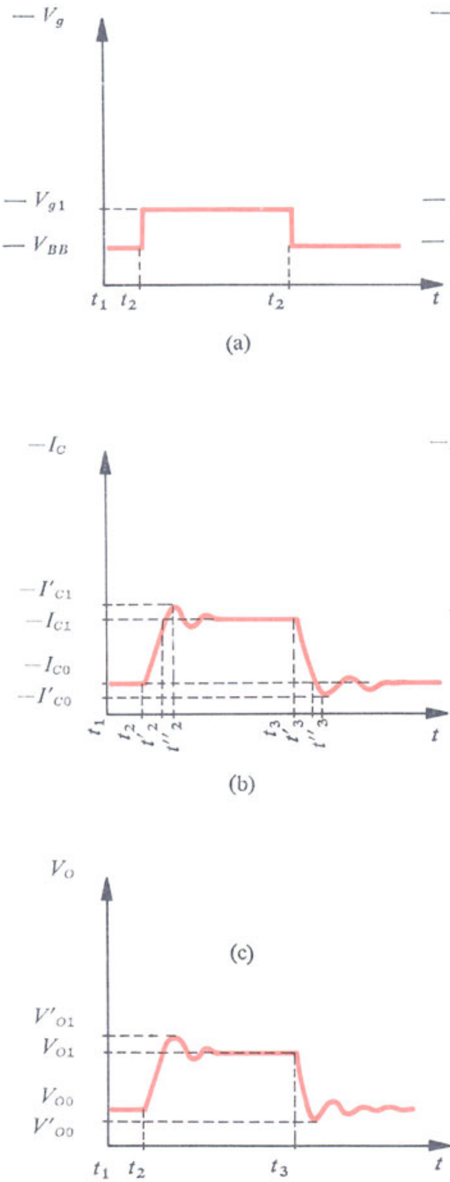


Fig. 298

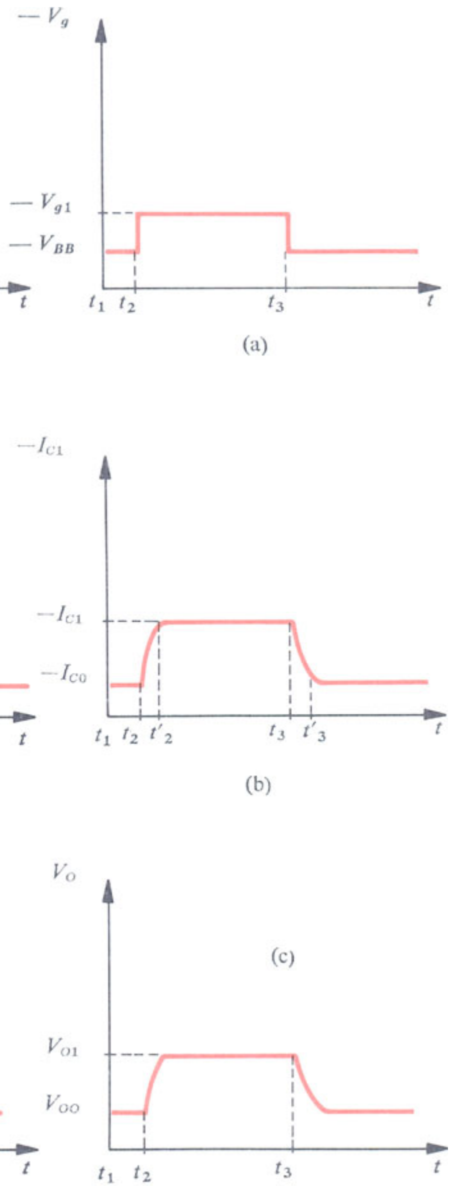


Fig. 299

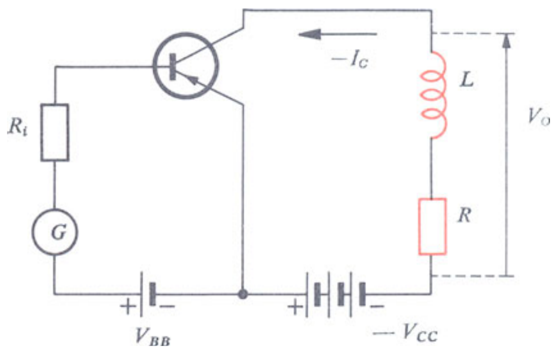


Fig. 300

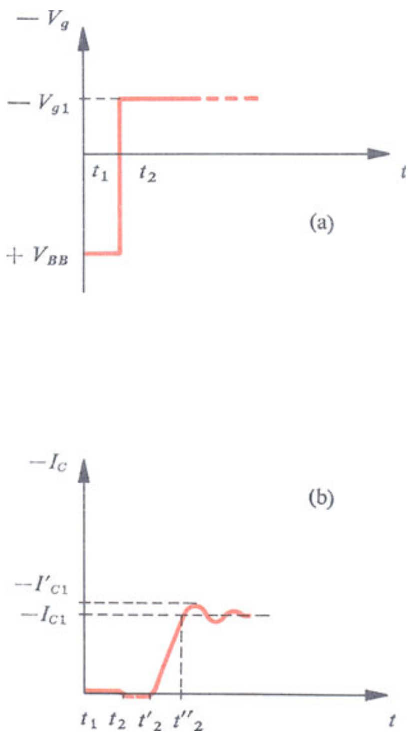


Fig. 301

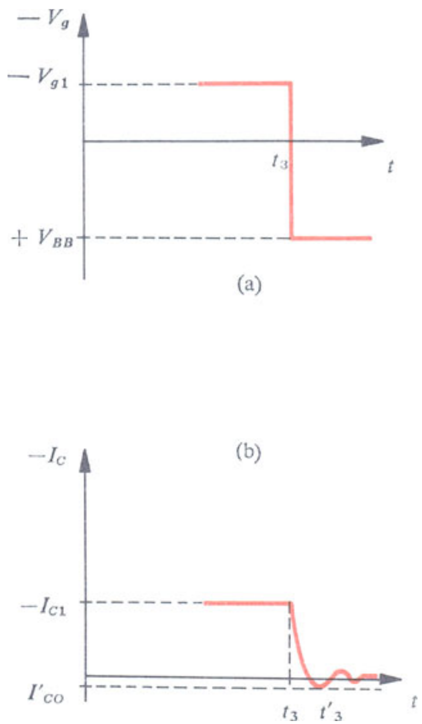


Fig. 302

3.4. Change from the cut-off to the normal state

There is a second possible method of operation with an inductive load, namely to switch the transistor from a cut-off to a normal state and vice versa. Referring to Fig. 300, it will be seen that the transistor is driven by a group of generators (generator G supplying a rectangular voltage in series with a battery whose positive pole is connected to the base); these two generators drive the transistor through a driving resistance R_i. The load still consists of an inductance L and a resistance R in series; the negative pole of the group of batteries $-V_{CC}$ supplies the collector of the transistor

The study of the dynamic behaviour of a circuit of this type requires consideration of two states:

(a) during the first transient period when the transistor is switched from the cut-off to the normal state;

(b) during the second transient period when the transistor is switched from the normal to the cut-off state.

In the first case from instant t_1 to the instant preceding t_2 the transistor is cut-off; the voltage applied between the base and emitter is practically equal to V_{BB} (Fig. 301a) and the collector current is zero (Fig. 301b).

At instant t_2, the voltage supplied by the group of generators changes suddenly from the value V_{BB} to the value $-V_{g1}$ The transistor should normally start to conduct and a current should appear in the output circuit. In fact, however, it is necessary before the holes can pass from the emitter into the base that the base–emitter barrier should narrow. The time essential for this reduction gives the delay time of the current in the collector circuit. It is given by $t_2'-t_2$ in Fig. 301.

At instant t_2 the holes pass from the emitter into the base and the current increases with time until it reaches the value $-I_{C1}$. With under-damping a ringing appears with its maximum at I_{C1}' in Fig. 301.

When the transistor passes from the normal to the cut-off state, that is at instant t_3 in Fig. 302a, the collector current, which before was equal to $-I_{C1}$, starts to decrease and disappears at instant t_3'. The current then passes through a value I_{CO}; as before a train of damped oscillations appears in the collector circuit.

The switching of the transistor loaded with an inductance from the cut-off to the normal state causes only slight modifications in the operation compared with that in the normal state. It is only the delay time of the current that differentiates between the two possible methods of use of a transistor.

When studying the behaviour of a pulse amplifier circuit we mentioned particularly the two possible methods of use:

(a) with under-damping;

(b) with over-damping.

We must therefore consider the same methods in the present study.

Under-damping. With under-damping (Fig. 303), from instant t_1 to the instant preceding t_2, the transistor is at cut-off. The current in the collector circuit is zero (Fig. 303b). In these conditions the voltage at the terminals of the load (L in series with R) is also equal to 0 (Fig. 303c).

At instant t_2, when the voltage changes suddenly from the value V_{BB} to the value $-V_{g1}$, no current appears in the circuit because of the necessity for the width of the base–emitter barrier to be first reduced. From instant t_2 to instant t_2', the absence of a current in the collector circuit (Fig. 303b) results in a zero value of the voltage at the terminals of the load (Fig. 303c).

At instant t_2 the holes start to diffuse into the base; the collector current increases rapidly to the value $-I_{C1}$ (instant t_2'' in Fig. 303b). At the same time the voltage at the terminal of the load increases to the value V_{O1} (Fig. 303c).

As we are dealing with the case of under-damping, the appearance of oscillations causes a new increase in the collector current (that is the maximum value $-I_{C1}'$ in Fig. 303b). The voltage at the terminals of the load is affected by forced oscillations, giving the maximum value of the voltage V_{O1}' (Fig. 303c). The train of damped oscillations then disappears and the current assumes its final value $-I_{C1}$.

At instant t_3, when the transistor is again cut-off, the shapes of the collector current and the voltage at the terminals of the load are as shown in Figs 303b and 303c respectively.

Over-damping. With over-damping, the variations of the current in the collector circuit and of the voltage at the terminals of the load are no longer influenced by the presence of trains of oscillations (in the final stage of the rise of the current or at the end of the stage of interruption of the current in the circuit).

In Figs 304b and 304c, which represent on the one hand the variations of the collector current with time, and on the other hand the variations of the output voltage with time, the shapes of the signals correspond almost exactly, except for the fall times, with those we have indicated for small signals.

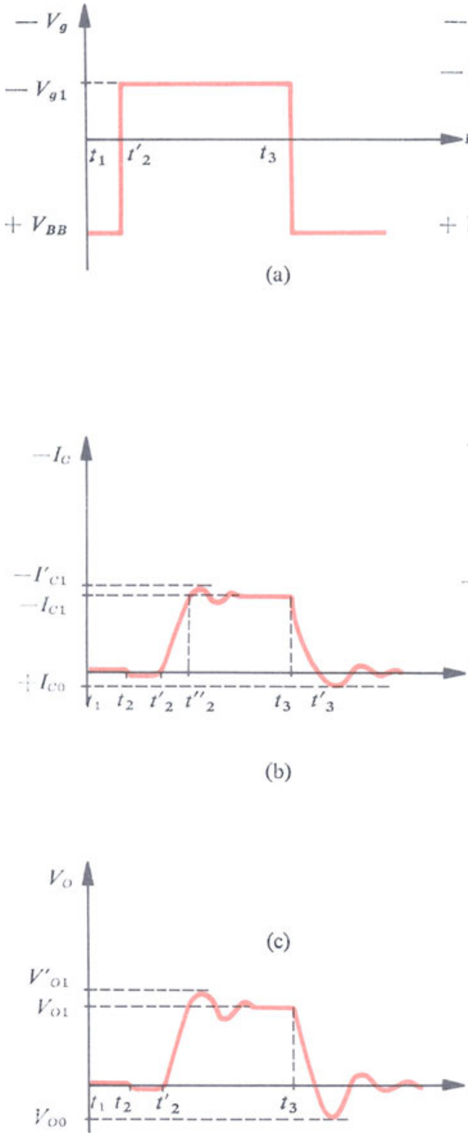


Fig. 303

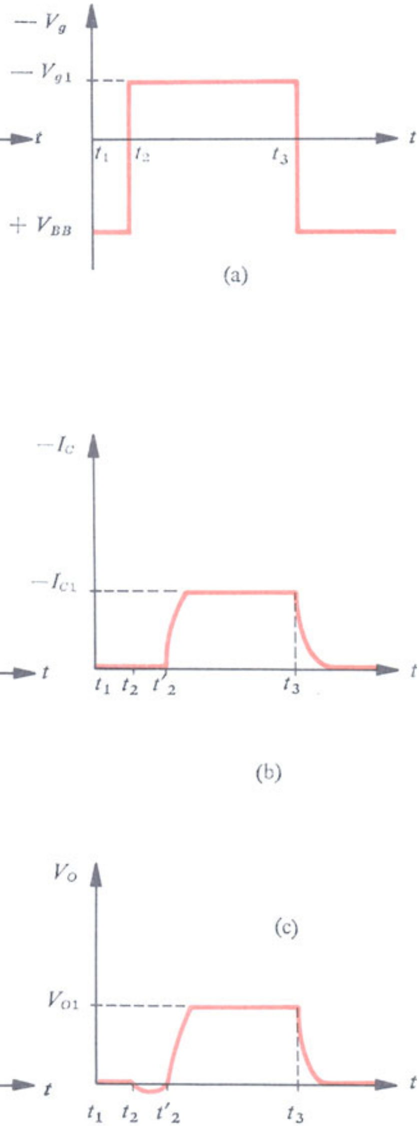


Fig. 304

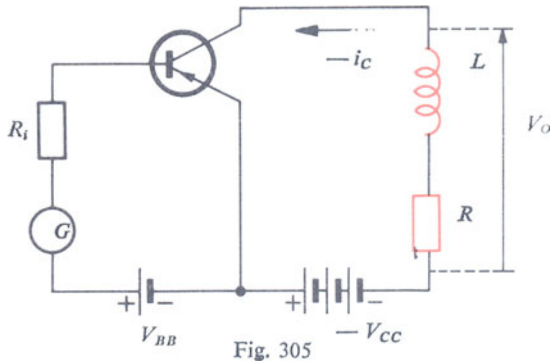


Fig. 305

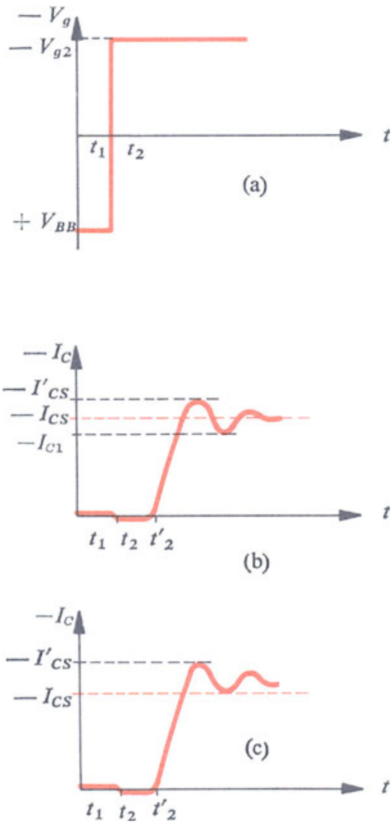


Fig. 306

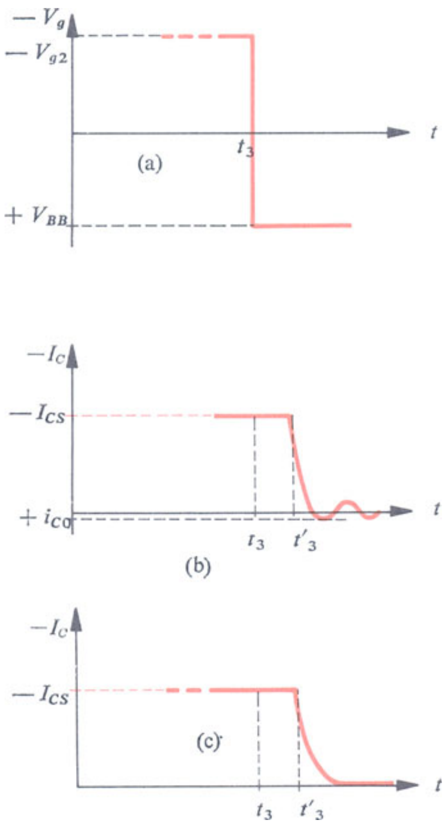


Fig. 307

3.5. Change from the cut-off state to saturation

The third possible use of a transistor loaded with an inductance is to switch it from the cut-off to the saturation state and vice versa. As will be seen in Fig. 305 the base of the transistor is driven by a group of generators through the driving resistance R_i; the load is made up as before of an inductance L and a resistance R in series; the collector is fed from the negative pole of a group of batteries $-V_{CC}$.

Because of the importance of transient effects, we must study the shape of the collector current when the pulse is applied (transition from the cut-off to the saturation state) and at the disappearance of the pulse (from saturation to the cut-off state).

Rise of the current. From instant t_1 to the instant preceding t_2, the transistor is at cut-off (Fig. 306a). The collector current is zero (Figs 306b and c).

At instant t_2 a strong voltage pulse supplied by the generator G tends to bring the transistor to saturation. In fact a certain time ($t_2'-t_2$ in Figs 306b and c) is required for the narrowing of the emitter–base barrier, corresponding to the absence of current in the collector circuit. At instant t_2' the current increases very rapidly and approaches the saturation value $-I_{CS}$ (Fig. 306b). It rises above that value to $-I_{CS}'$. It then decreases and falls below the saturation value, namely $-I_{C1}$ in Fig. 306b. This current approaches the final value $-I_{CS}$.

In a circuit of this type we note that the transistor may, because of the ringing appearing in the collector circuit, leave the saturation state. To prevent this we must choose values of the collector current such that the minimum value corresponding to the lowest peak of the train of damped oscillations lies above the value of the saturation collector current $-I_{CS}$ (Fig. 306c).

Disappearance of the current. At instant t_3 in Fig. 307a the collector current does not decrease at once because of the time required for the removal of the surplus charges stored in the base during the time the transistor is saturated.

From instant t_3 to instant t_3' the current remains at its saturation value (Figs 307b and c).

With under-damping, after instant t_3' the collector current decreases and a train of damped oscillations appears, causing this current to lie alternately on one side or the other of the abscissa (Fig. 307b).

With over-damping the current decreases and at once reaches its final value (Fig. 307c).

Switching of a transistor with an inductive load, from cut-off to saturation may cause:

(a) departure from the saturation state during the rise time of the current ($-I_{C1}$ in Fig. 306b);

(b) transistion from cut-off to the normal state during the period of interruption.

It will be seen from the two above studies that we must differentiate between under-damping and over-damping.

Under-damping. If the condition of under-damping is observed the shapes of the current and voltage in the output circuit are as shown in Fig. 308.

From instant t_1 to the instant preceding t_2 the transistor is non-conducting (Fig. 308a), no current flows through the collector circuit (Fig. 308b) and no voltage appears at the terminals of the load (Fig. 308c).

At instant t_2 when the signal is applied, the current and voltage are still zero in the output circuit ($t_2'-t_2$ in Figs 308b and c).

From instant t_2' the collector current starts to increase and rapidly reaches the saturation value $-I_{CS}$ and the collector voltage increases to the value V_{O1} (Fig. 308c). Since we are dealing with under-damping, the collector current exceeds the saturation value and passes through a maximum value equal to $-I_{CS}'$ (Fig. 308b). Simultaneously the output voltage reaches the level V_{O1}' in Fig. 308c. During the first negative half cycle the train of damped oscillations produces a collector current lower than the saturation value $-I_{CS}$ and at the same time an output voltage V_{O1}'' smaller than the final value V_{O1} (Figs 308b and c).

The collector current and the output voltage remain at their respective values until instant t_3 which corresponds to the next triggering of the system.

At this instant these two values still do not vary decause of the time required for the removal of the surplus charges stored in the base of the transistor ($t_3'-t_3$ in Figs 308b and c). From instant t_3' the collector current falls, disappears, assumes a small positive value and returns again to zero. The output voltage is affected by the same fluctuations with respect to the equilibrium value.

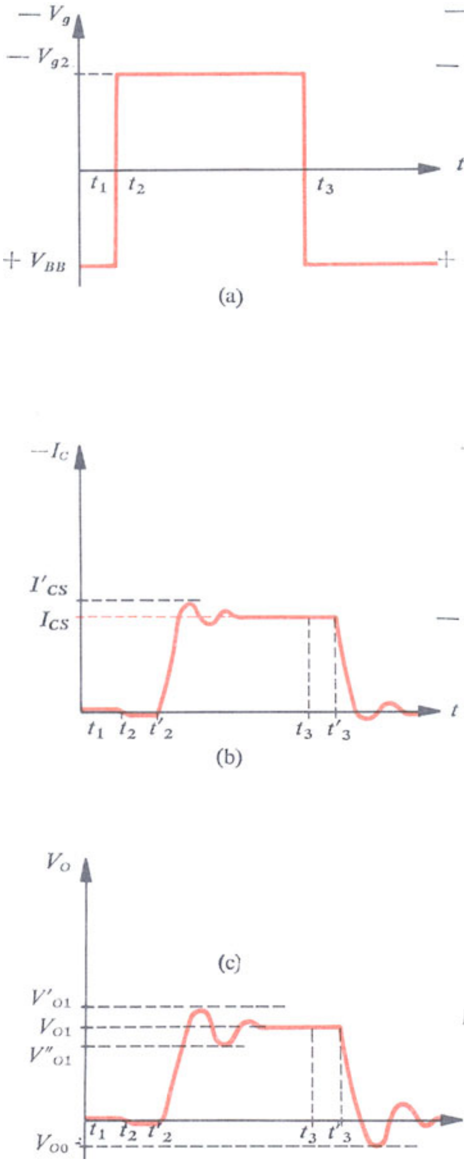


Fig. 308

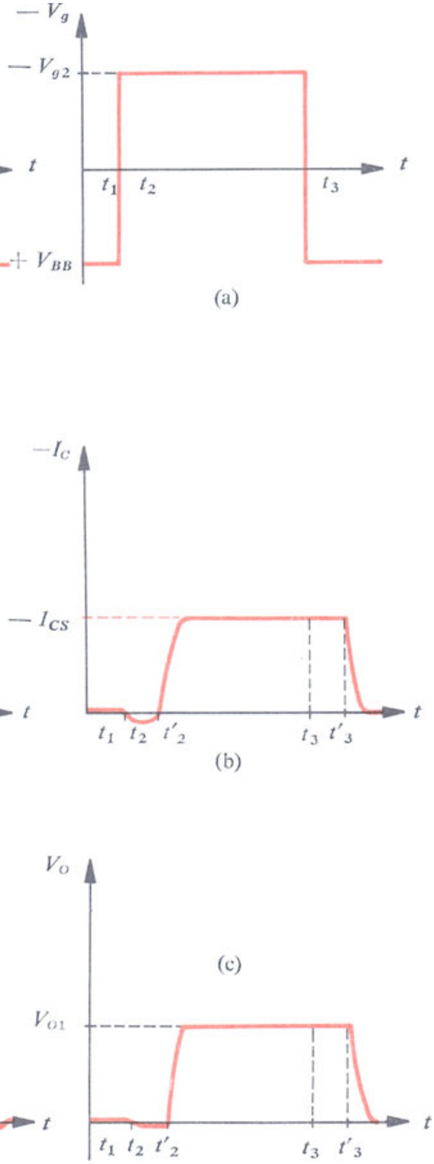


Fig. 309

Over-damping. Figure 309 shows the variations of the current and of the output voltage when the transistor is switched from cut-off to saturation and vice versa.

The total absence of ringing in the circuit causes practically rectangular variations in the output current, the collector current (Fig. 309b) and the output voltage (voltage at the terminals of the load Fig. 309c).

4. Thermal Behaviour of Switching Transistors—Limiting Operating Characteristics

Switching can be defined as the use of 'all or nothing' semiconductor devices. The production or reproduction of a rectangular signal is equivalent to causing the transistor to change from one level to another. This mode of operation is therefore associated with the production or amplification of perfectly rectangular signals such as that shown in full line in Fig. 310. The transit times, that is the triggering time to instants t_2 and t_3 are then considered as negligible. With slow switching we can assume that the signals supplied by the stage considered resemble closely the theoretical signals shown in the figure.

We have shown earlier that the rise and fall times occur during the periods when the collector current is varying. A true rectangular signal is therefore very like that shown in dashed line in Fig. 310; this signal is associated with the current rise time ($t_r = t_2' - t_2$) and the current fall time ($t_f = t_3' - t_3$).

4.1. Power considerations

In the power stages of audio or radio frequency circuits, the most important criterion is the definition of the maximum power which can be supplied by a stage without the danger of destruction of one or more transistors. In the field of switching and especially in the case of signal generation, it is interesting, starting from the data supplied by the manufacturer, to calculate the maximum power which can be switched without exceeding the prescribed limits.

Before passing on to the study of rectangular signals, it may be useful to remind ourselves of certain basic principles concerning the various powers dissipated in a transistorized stage, and to associate with them the limiting current and voltage characteristics.

We will first consider a transistorized amplifier (Fig. 311), in which it will be seen:

(a) that the battery supplies a certain power equal to P_B;

(b) that a certain amount of energy P_O is available at the terminals of resistance R_L;

(c) that a certain amount of power is dissipated in the collector of the transistor, i.e. in the collector–base junction, P_C.

These different powers are connected by the relation:

$$P_B = P_O + P_C.$$

These powers can be represented on the network of characteristics in Fig. 312 by rectangles whose area can be calculated.

The power supplied by the battery (green rectangle) is given by the formula:

$$P_B = V_{CC} \cdot I_C.$$

The power available at the terminals of the load resistance (in black) is supplied by the relation:

$$P_O = V_{CC} \cdot I_C / 4.$$

The power dissipated in the collector–base junction of the transistor (red rectangle) is given by the equation:

$$P_C = V_{CC} \cdot I_C / 2.$$

4.2. Reminder of the thermal behaviour of transistors

When dealing with thermal behaviour in the first two books of this series it was shown that the basic thermal equation of a transistor is:

$$T_j = T_a + R_{th(j-a)} \cdot P_C.$$

In this relation T_j represents the temperature of the junction, T_a the ambient temperature, $R_{th(j-a)}$ the thermal resistance between the collector–

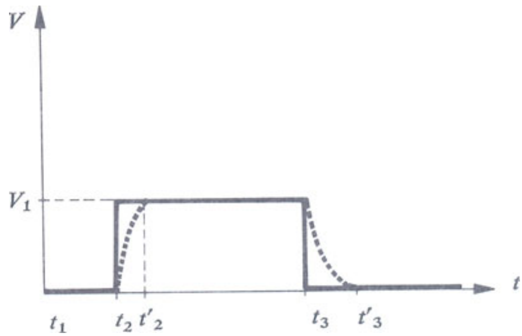


Fig. 310

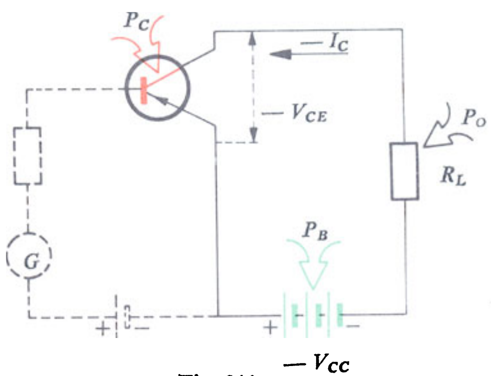


Fig. 311

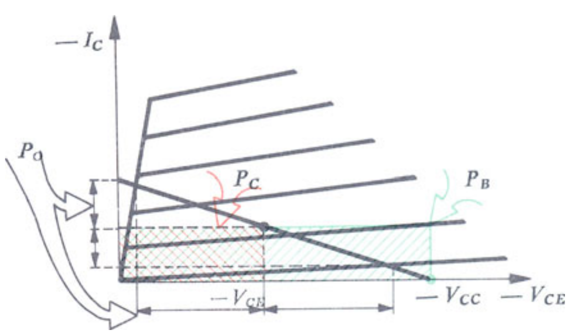


Fig. 312

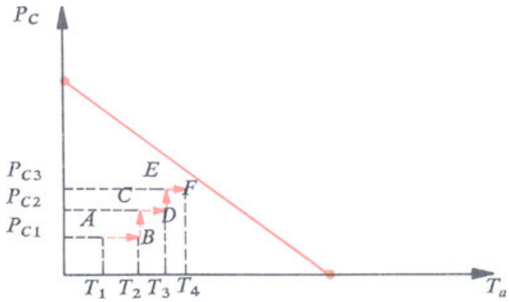
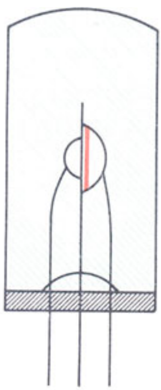
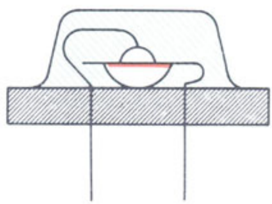


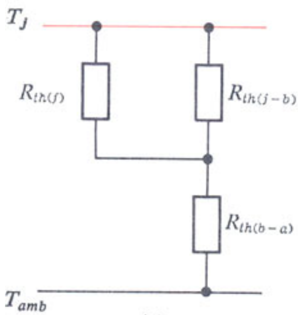
Fig. 313



(a)

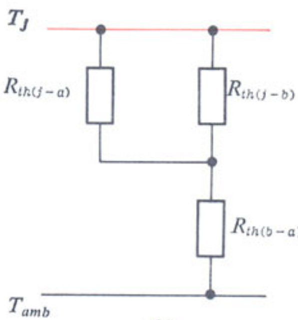


(a)



(b)

Fig. 314



(b)

Fig. 315

base junction and the surroundings, and P_C the power dissipated in the collector–base junction of the transistor.

The graphical presentation of this relation is shown in Fig. 313. If we assume that the power dissipated in the collector at instant t_1 is equal to P_{C1}, any increase in the junction temperature due to an increase in the ambient temperature will cause a displacement of the working point and at the same time an increase in the collector current. Consequently the power dissipated in the collector–base junction increases (P_{C2} in Fig. 313) This new increase in the collector power is associated, in accordance with the above expression, with a second rise of the junction temperature, and the process continues in accordance with the rising curve in the figure. If we wish to prevent the working point moving into the region above the characteristic shown in red, we must, starting from the above expression, try to obtain as low a thermal resistance as possible. What does the thermal resistance consist of? It is specified in data sheets in degrees per watt or in degrees per milliwatt, depending on the type of transistor used.

Low-power transistors

We will first consider 'small-signal' transistors shown in a sectional view in Fig. 315a. If we wish to establish the equivalent thermal circuit of a transistor of this type, all that is needed is to define the methods by which heat energy is transferred from the hottest to the coldest point.

This circuit is therefore inserted between two lines of thermal potential, one corresponding to the highest temperature (the red line T_j in Fig. 314b) and the other corresponding to the lowest temperature (the black line T_a in the same figure).

There are certain resistances between these thermal potential lines, namely:

(a) resistance $R_{th(j-b)}$ representing the thermal resistance between the junction and the case of the transistor;
(b) resistance $R_{th(f)}$ representing the possibility of the heat energy being removed through the connecting wires;
(c) resistance $R_{th(b-a)}$, the resistance between the case and the surroundings.

Power transistors

With this type of transistor we can consider a certain number of thermal resistances (Fig. 315b).

(a) resistance $R_{th(j-b)}$ (thermal resistance junction–case);
(b) resistance $R_{th(j-a)}$ (resistance junction–surroundings by direct diffusion);
(c) resistance $R_{th(b-a)}$ (thermal resistance case–surroundings).

Maximum output power

When studying the design of a transistorized circuit, the first step with large signals is almost always to define the maximum power which can be dissipated at the output. The data supplied by manufacturers include the power dissipated in the collector; we now have to associate this concept with the output power.

With sine-wave signals (Fig. 316a), the output power is a function of the peak-to-peak variation of the current in the collector circuit and of the peak-to-peak variation of the output power; these two variations are shown on the ordinate and abscissa respectively in the figure. The output power can be calculated by means of the expression:

$$P_O = \frac{V_O{}^\wedge}{\sqrt{2}} \cdot \frac{I_O{}^\wedge}{\sqrt{2}} = \frac{V_O{}^\wedge \cdot I_O{}^\wedge}{2}.$$

With rectangular signals, the output power is a function of the peak variation of the current and the peak variation of the voltage, but it is also dependent on the duration of the pulse with respect to the total cycle of the signal. The definition of this power will be discussed later.

In the two types of application, the output power increases as the variations of the current and voltage increase. Consequently if we wish to have available a greater power at the output of a stage, limiting values of the collector current and the collector emitter voltage should be chosen as near as possible to the maximum values specified by the manufacturer.

Maximum collector current

For all transistors the maximum permissible value of the collector current is given by the expression $-I_{CM}$ (Fig. 317a). If we locate the load line so as to obtain this maximum variation of the collector current (shown in red in the figure), the output power, for the same supply voltage, is much larger than the value shown in green which corresponds to an intermediate value of the load resistance.

This applies even more to pulse operation. It will be seen in Fig. 317b that a load line whose location depends on the maximum value of the collector current is accompanied by a peak value of the current in the output circuit equal to I_O', which is much higher than the value I_O shown in green.

Finally, in order to obtain the highest possible output power, we must always try to locate the working point near the highest values of the collector current, that is at the level $-I_{CM}$.

Maximum collector–emitter voltage

Since the output power is given by the product of the variations of the output voltage and the output current, its maximum value is obtained when

MAXIMUM OUTPUT POWER

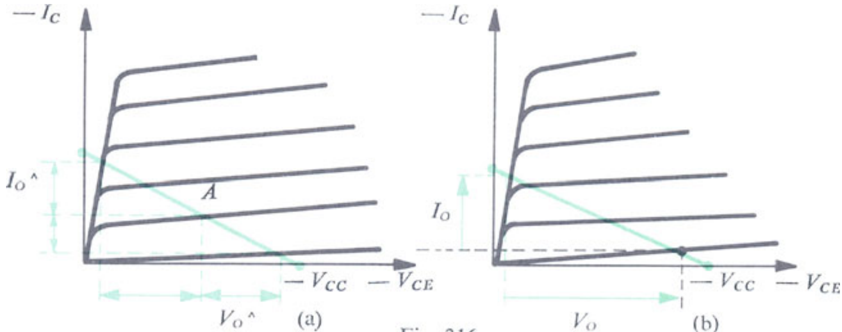


Fig. 316

MAXIMUM COLLECTOR CURRENT

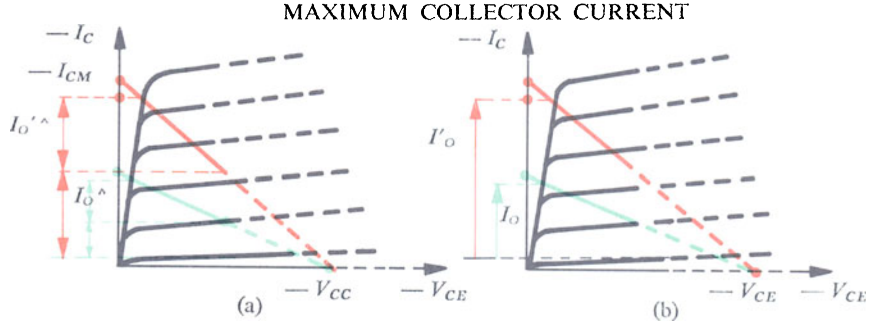


Fig. 317

MAXIMUM COLLECTOR–EMITTER VOLTAGE

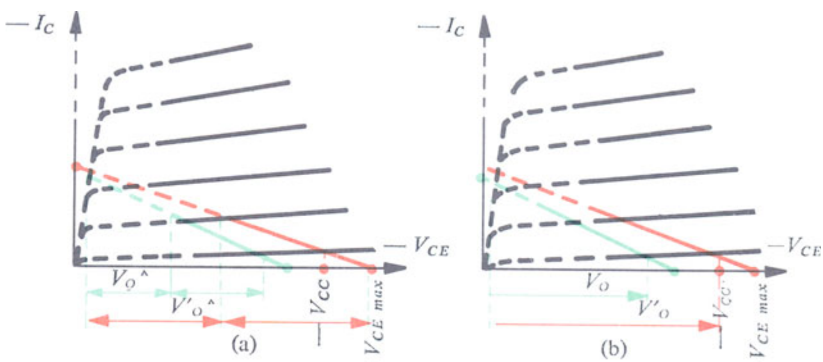


Fig. 318

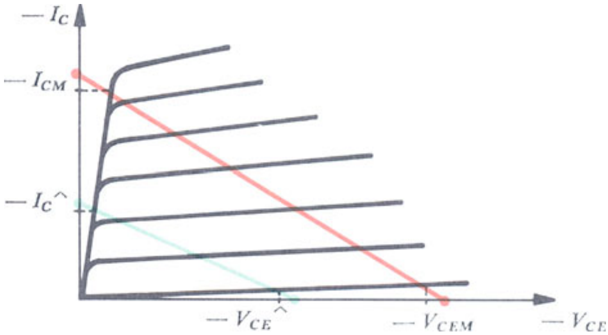


Fig. 319

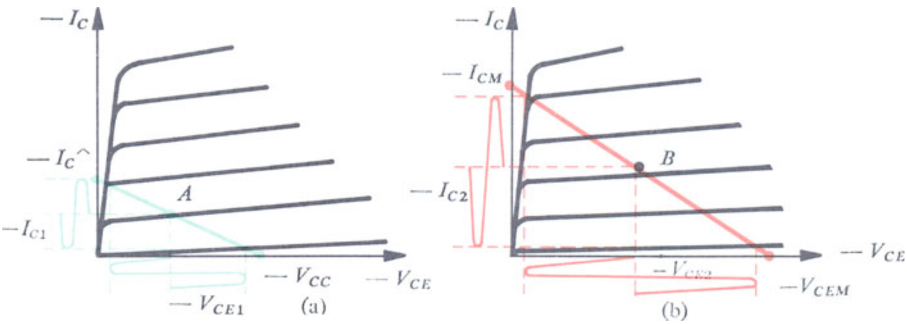


Fig. 320

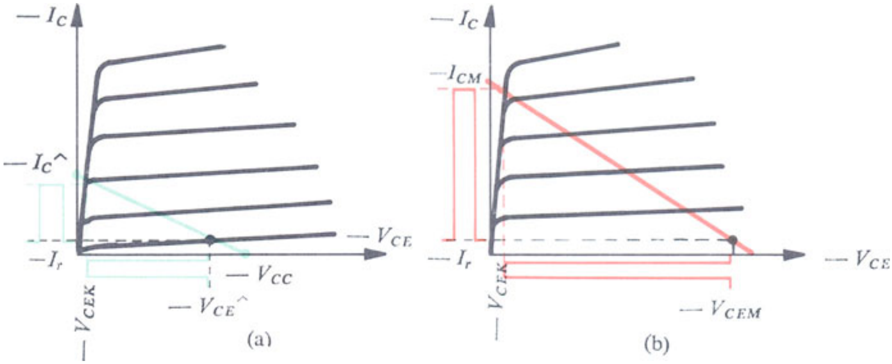


Fig. 321

the extreme position of the working point is near the maximum collector–emitter voltage ($-V_{CEM}$), specified by the manufacturer.

It will be seen in Figs 318a and b how important it is with the two types of application to locate the load line with one of its ends near the maximum voltage.

Optimum position of the load line

Figure 319 shows two possible positions of the load line on the network of output characteristics of a transistor:

(a) one (in green) associated with any given value of the load resistance with a supply voltage much lower than $-V_{CEM}$;

(b) the other (in red) obtained by means of a straight line passing through the two extreme values of the collector current and collector–emitter voltage.

Sine-wave signals. In the intermediate case (in green in Fig. 320) the extreme positions of the working point, when projected on to the collector current axis, give the limiting values of the peak-to-peak variations of the output current, and when projected on to the collector–emitter voltage axis, they give the limiting values of the peak-to-peak output voltage. These variations of the collector current and the collector–emitter voltage are equal to $-I_{C1}$ and near to $-V_{CC}$ respectively. In the red example (Fig. 320) the peak-to-peak variation of the collector current is given by the maximum value of this current $-I_{CM}$; and the peak-to-peak variation of the collector–emitter voltage is near to the maximum permissible voltage $-V_{CEM}$. In this last case the two variations of the current and the output voltage are much greater than the variations shown in green in the previous figure.

Rectangular signals. With rectangular signals the output energy is also a function of the variations of the current and output voltage.

If we examine the two cases corresponding exactly to the load lines and to the values of supply voltage previously chosen, we see that:

(a) for a load line of any given value associated with a supply voltage $-V_{CC}$ which is lower than $-V_{CEM}$, the variations of the collector current in the circuit are given by $-I_C$ and the variations of the output voltage by a peak value $-V_{CE}$ which is very little different from $-V_{CC}$;

(b) when the transistor is being used near its upper limit of performance, where the maximum values of the current and voltage (Fig. 321b) are concerned, the variation of the output current is now given by the maximum value of the current specified by the manufacturer, namely $-I_{CM}$; the variation of the output voltage is dependent on the maximum value of the collector–emitter voltage permissible for this type of transistor.

Just as for sine-wave signals, the output signal is much larger in the red example than in the intermediate case shown in green in this group of figures.

At this stage the reader's attention should be drawn to the fact that these conclusions are purely academic; other parameters occur effectively in the expression for the maximum amounts of power that can be dissipated at the output of a transistor whatever type of signals are to be amplified or generated (sine-wave or rectangular signals).

4.3. Power dissipated in a transistor

Before entering into a detailed calculation of the different amounts of power involved with rectangular signals, we will make a brief reminder of the amounts of power dissipated in a transistorized circuit, with special emphasis on the limitations of such a circuit.

We will consider the intermediate case in Fig. 322. The load line (brown) in this figure is associated with a value of supply voltage equal to $-V_{CC1}$. The transistor is used in class A operation; the quiescent point is located at the centre of the load line for the quiescent values of current and collector–emitter voltage, which are equal to $-I_{C1}$ and $-V_{CE1}$ respectively.

The power available at the output of this circuit with 100% modulation, is given by the relation:

$$P_O = V_{CE1} \cdot I_{C1}/2.$$

If we now operate the transistor at its limiting range, and dealing only with the limiting values of current and voltage, the power available at the output of the transistor is given by the expression:

$$P_O = V_{CE2} \cdot I_{C2}/2.$$

Comparing the two expressions, we see that in the case of Fig. 323 the output power is much larger than the power available in the example in Fig. 322.

The power dissipated continuously in the collector circuit of the transistor is given, in the two cases, by the area of the rectangle obtained by the projection of the quiescent point on the ordinate and abscissa:

(a) brown rectangle (Fig. 322);

(b) red rectangle (Fig. 323).

Is it possible, if we wish to have this amount of energy available at the output, to dissipate in the collector such an important amount of energy?

The data supplied by the manufacturer always include a maximum collector power as a function of the ambient temperature. If we locate this power axis on the bisector of the angle made by the ordinate and abscissa of the output characteristics network, the maximum permissible power is represented by a hyperbola.

In this case, the load line allowing the maximum output power without risking destruction of the transistor, is at a lower level (Fig. 324) than that in the red example (Fig. 323).

When a transistor operates in the optimum conditions shown in Fig. 324 the power dissipated in its collector is a maximum in the absence of any modulation, and is given by the brown rectangle in Fig. 325; this power is

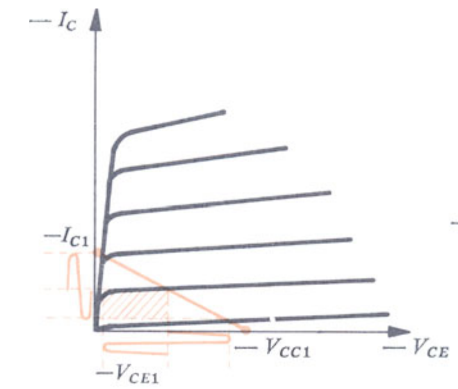


Fig. 322

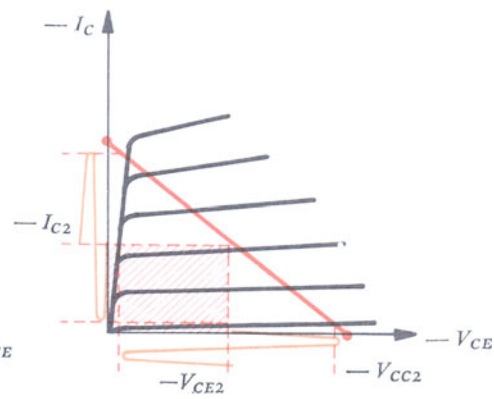


Fig. 323

given by the expression:

$$P_C = V_{CE3} \cdot I_{C3}.$$

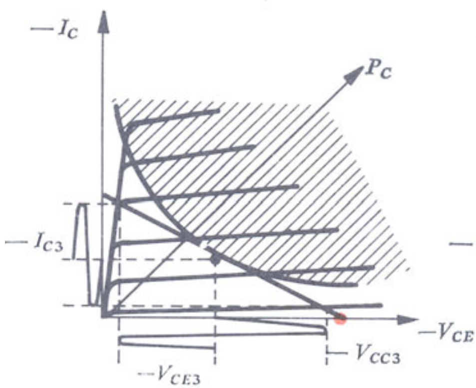


Fig. 324

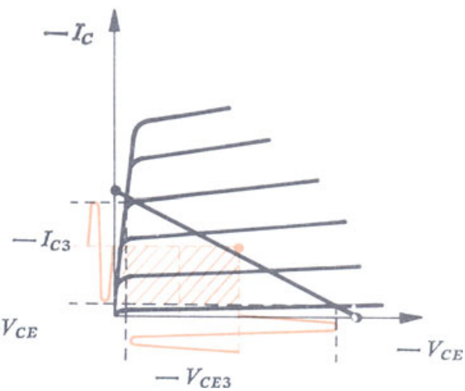


Fig. 325

With 100% modulation, the output power is practically equal to one half the collector power; the yellow rectangle defines then the corresponding power dissipated in the collector of the transistor.

In addition to audio- and radio-frequency applications, associated with sine-wave signals, we must determine the amounts of power dissipated in the collector of the transistor in other applications, especially:

(a) for d.c. amplifiers;
(b) for d.c. regulators;
(c) for rectangular signals.

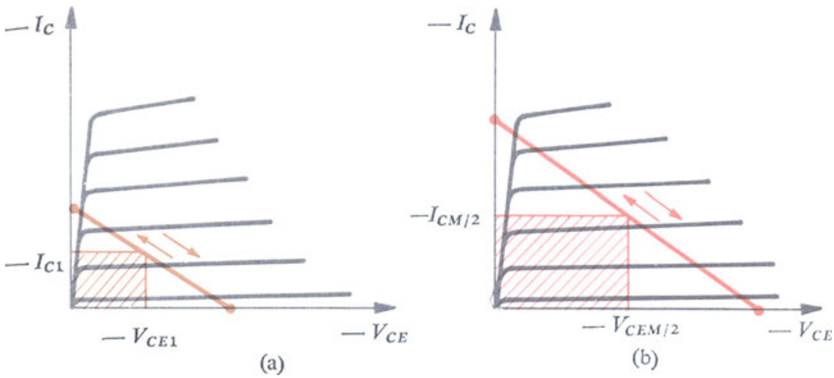


Fig. 326

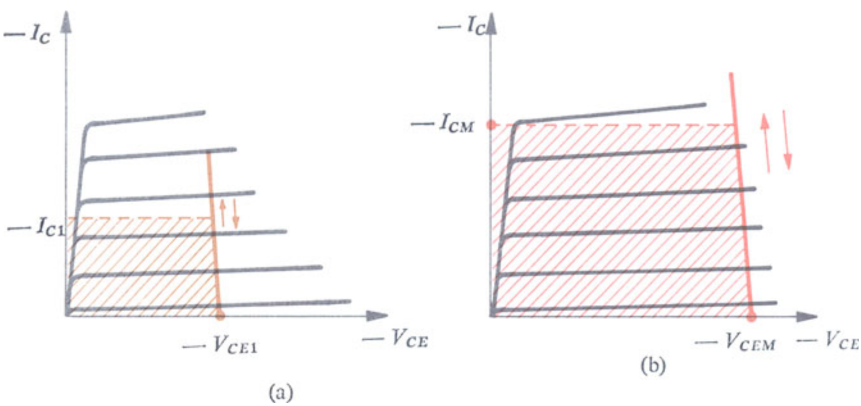


Fig. 327

Direct-current amplifier

When a transistor is used as a d.c. amplifier, the working point shifts from one side to the other of the quiescent position, as shown in Fig. 326a. The quiescent collector–emitter voltage is equal to $-V_{CE1}$ and the collector current to $-I_{C1}$.

The quiescent point is located at exactly the centre of the load line. In these conditions, the power dissipated in the collector of the transistor is directly connected to the area of the brown rectangle.

If we use the transistor at the limit of its possibilities, we see that the power associated with the red rectangle in Fig. 326b is greater.

Direct-current regulator

With regulation the problem is a little different. The load lines are still nearly vertical, because of the very small values of the load resistances used. If we choose a supply voltage lower than the maximum value of the collector–emitter voltage (Fig. 327a), the power dissipated in the collector of the transistor is a function of the product:

$$P_C \approx V_{CE1} \cdot I_{C1}.$$

The maximum value of the collector current $-I_{C1}$ marks the upper limit of a rectangle whose area is directly dependent on the above expression.

If we wish to obtain the maximum performance, the supply voltage is made equal to the maximum permissible collector–emitter voltage ($-V_{CEM}$) and the collector current equal to the maximum value $-I_{CM}$ specified by the manufacturer. The power dissipated in the collector of the transistor is then represented by the 'red shaded' rectangle in Fig. 327b, namely:

$$P_C \approx V_{CEM} \cdot I_{CM}.$$

In regulation the amounts of power dissipated in the collector are very large and much higher than the powers involved earlier. The maximum collector power characteristics, as functions of the ambient temperature, occur also in the maximum value of the supply voltage and collector current.

Powers of this magnitude cannot be satisfactorily dissipated in the collector–base junction of a transistor except for very short periods. We will see later that, in switching, this expression enables us to define a new power known as the 'maximum controlled power'.

Rectangular signals

In a switching circuit the transistor is used as an interrupter. It passes from one state to another, that is from saturation to cut-off and vice versa. Very precise values of the collector current and voltage are associated with each of these states. In order to define the energy dissipated in the collector of the transistor in any given type of operation, we must specify the corresponding values of the power dissipated.

When the transistor is saturated (position 1 in Fig. 328), a collector current equal to the saturation current ($-I_{C1}$) flows through the output circuit, and because of the magnitude of this current the voltage between the collector and the emitter is very low, namely $-V_{CO}$.

During the cut-off period (position 2 in Fig. 328), the collector current is a minimum at $-I_{CO}$; conversely, since the voltage drop in the load resistance is very small, the collector–emitter voltage is a maximum and near to the value of the supply voltage ($-V_{CC}$).

Before we pass on to the definition of the mean power dissipated in the collector of the transistor, it would be interesting to determine the amounts of power dissipated, on the one hand during the period of saturation, and on the other during the period of cut-off.

Power dissipated in the transistor during the period of saturation. When the transistor is saturated (Fig. 329), the collector current is a maximum $-I_{C1}$ and the collector–emitter voltage is a minimum $-V_{CO}$; in order to define the power dissipated in the collector–base junction of the transistor, all we need do is to calculate the area of the red shaded rectangle in this figure. Consequently the collector power is given by the expression:

$$P_{CS} = V_{CO} \cdot I_{C1}.$$

If the transistor remains permanently saturated, this power corresponds to a power continuously dissipated in the collector–base junction of the transistor. During saturation, and in spite of the magnitude of the current, the power dissipated is extremely small because of the low value of the voltage across the collector–base junction.

Power dissipated in the transistor during the period of cut-off. When the transistor is cut-off (Fig. 330), the power dissipated in the collector of the transistor is given by the black shaded area in the figure. This area is dependent on a very low value of the collector current $-I_{CO}$ and a high value of the collector–emitter voltage $-V_{CE}$; the power dissipated is then given by the relation:

$$P_{CO} = V_{CE} \cdot I_{CO} \approx V_{CC} \cdot I_C.$$

Just as in the saturation state, the power dissipated in the collector of the transistor is small in spite of the high value of the collector–emitter voltage; this is explained chiefly by the negligible value of the collector current ($-I_{CO}$). The saturation and cut-off states do not correspond with the normal states, but they are associated with very precise durations. It is therefore necessary to know the saturation and cut-off periods in order to calculate as accurately as possible the mean power dissipated in the transistor.

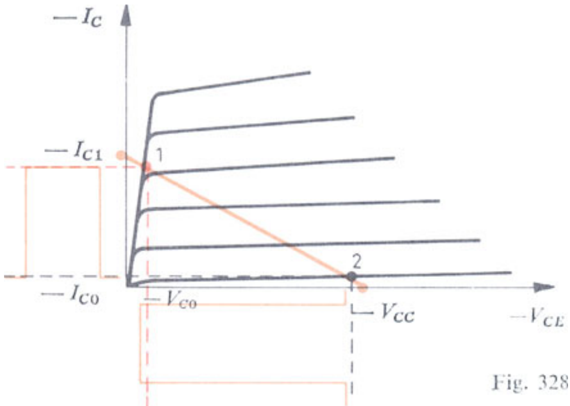


Fig. 328

Power dissipated in the transistor during saturation period

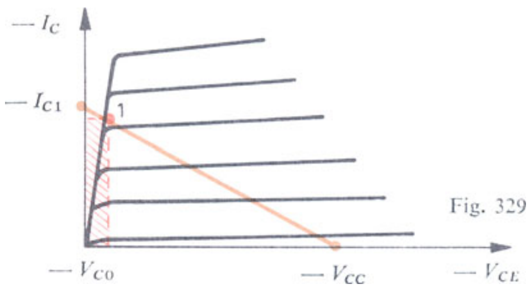


Fig. 329

Power dissipated in the transistor during cut-off time

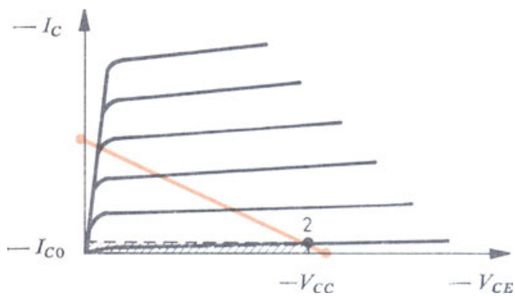


Fig. 330

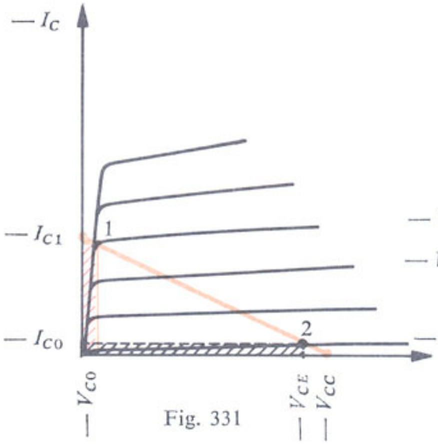


Fig. 331

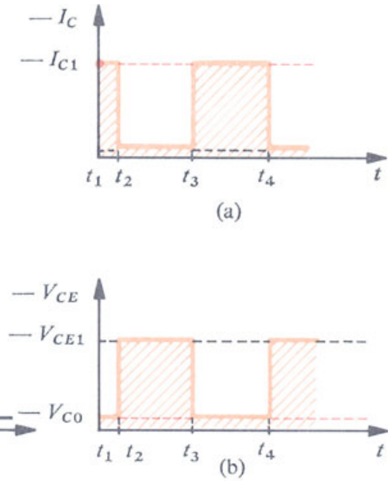
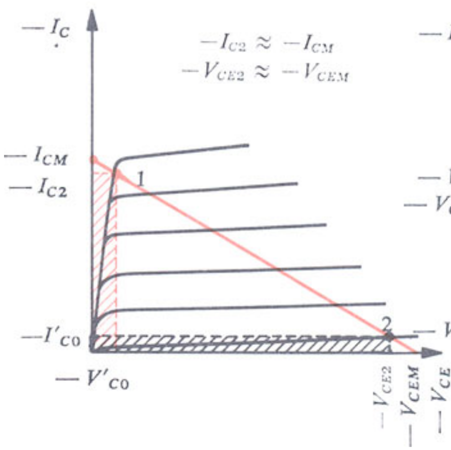
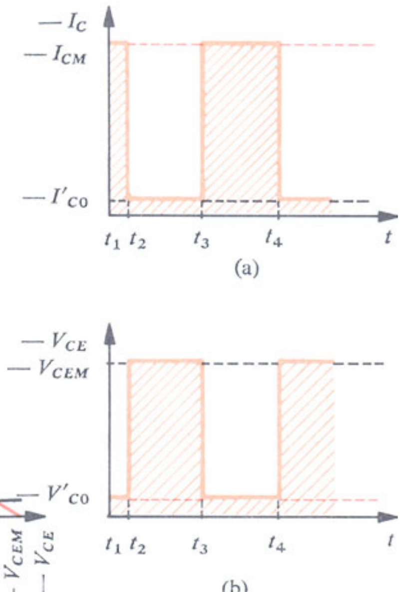


Fig. 332





4.4. Mean power dissipated in the transistor

Referring back to the example on the previous page, we see that at instant t_1 the transistor is operating in the saturation state.

On the network of output characteristics in Fig. 331, the operating condition is associated with position 1 of the working point and with the red shaded rectangle, whilst the cut-off state corresponds to position 2 on the load line and to the black shaded rectangle.

If we wish to know the variations of the power dissipated in the collector of the transistor, we must represent simultaneously the variations with time of the collector current and collector–emitter voltage; (see Figs 332a and b). At instant t_1 the transistor is saturated, the collector current is a maximum and equal to $-I_{C1}$, and this state is maintained until the instant preceding t_2. The collector–emitter voltage is a minimum equal to $-V_{CO}$.

At instant t_2, and without taking the inertial effects into account, the collector current changes immediately from $-I_{C1}$ to $-I_{CO}$; it remains at this value until the instant preceding t_3. During this period, the collector–emitter voltage becomes very large and approaches the value of the supply voltage (namely $-V_{CE1}$ in Fig. 332b).

At instant t_3 another change in the method of operation occurs; the transistor becomes saturated again. The collector current becomes very high and near to $-I_{C1}$. This condition is maintained until the instant preceding t_4 (Fig. 332a). The collector voltage then becomes very low again, i.e. $-V_{CO}$ in Fig. 332b. The variations of the current and voltage are perfectly rectangular; they are shown in the brown shaded areas in Figs 332a and b.

Before expressing the power dissipated in the collector of the transistor, we will consider the possibility of using this transistor at its maximum output, that is at the limiting values of current and voltage. Position 1 in Fig. 333 corresponds to a saturation state of the transistor and is associated with a saturation value of the current $-I_{C2}$ near to the value $-I_{CM}$. At position 2 the transistor is in the cut-off state and the collector–emitter voltage has the value $-V_{CE2}$ almost the same as $-V_{CEM}$.

If we wish to represent the variations of the current and of the output voltage with time, all that is necessary is to examine the limiting values of the collector current ($-I_{CM}$ and I_{CO}' in Fig. 334a) and of the collector–emitter voltage ($-V_{CO}'$ and $-V_{CEM}$ in Fig. 334b).

The shape of the received signal in the second case is exactly the same, although the amplitude is higher than the theoretical value already shown in the previous figure.

The variations of the voltage and current are in opposition; any increase in the current causes a decrease in the collector–emitter voltage and vice versa. The same applies when operating with sine-wave signals.

4.5. Calculation of the power dissipated in the transistor

The transistor in the circuit shown in Fig. 335 is driven through a group of generators (generator G and battery V_{BB} and the driving resistance R_i.

The load on the transistor consists of the resistance R_L; and the supply is from the negative pole of a group of batteries $-V_{CC}$.

On the network of output characteristics of the transistor shown in Fig. 336 the two operating states are represented respectively by position 1 of the working point for the saturation state, and by position 2 for the non-conducting state.

The signal supplied by the group of generators is shown in Fig. 337. In practice, the sum of the two voltages supplied by the generator G and the group of batteries $+V_{BB}$ causes the transistor to switch from cut-off (V_{BB}) to saturation ($-V_{g1}$) and vice versa. The duration of the saturation state is dependent on the time of application of the voltage $-V_{g1}$ across the input of the transistor.

If T represents the total duration of the signal, then t_1 represents the duration of the conduction of the transistor, and t_2 its cut-off period.

We will now represent the variations with time of the collector–emitter voltage and of the collector current in order to enable us to calculate the amounts of power effectively dissipated in the collector–base junction of the transistor during these two different operating states.

In Fig. 338a the transistor is first in the cut-off state and the collector–emitter voltage is very high ($-V_{CE1}$).

At the instant when the transistor becomes saturated the voltage disappears immediately ($-V_{CO} = 0$); it remains at this low value until the instant when the applied signal causes the transistor to become cut-off again.

If t_1 represents the period of saturation of the transistor and corresponds to a very low value of the voltage, this period is associated with a very high value of the collector current, namely $-I_{C1}$ in Fig. 338b. On the other hand, during the cut-off period the current is very small ($-I_{CO}$).

In order to calculate the power dissipated in the collector of the transistor all we need for each state is to determine the area of the rectangles in Fig. 336.

When the transistor is in the saturated state (that is the period t_1 in Fig. 338c) the power dissipated in the collector is given by the red shaded area in Fig. 336, that is by the relation: $P_{CS} = V_{CO} \cdot I_{C1}(t_1) \approx V_{CEK} I_{C1}(t_1)$.

During the cut-off period (period t_2 in Fig. 338c) the power dissipated in the collector is given by the black shaded area in Fig. 336. It is expressed by the equation: $P_{CO} = V_{CE1} \cdot I_{CO}(t_2) \approx V_{CC} \cdot I_{CO}(t_1)$.

The definition of the instantaneous values of the power dissipated in the collector is equivalent to the determination of the variations of the rectangular powers; P_{CO} determines the power dissipated in the transistor

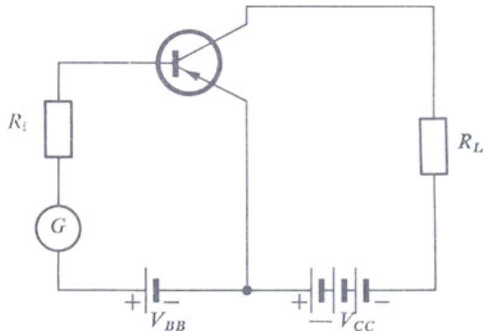


Fig. 335

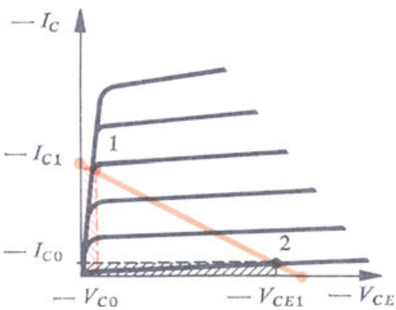
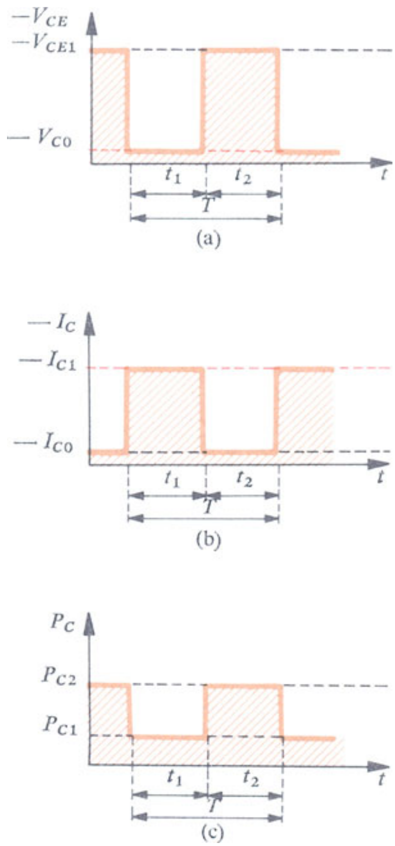


Fig. 336



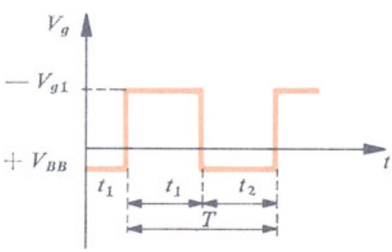


Fig. 337

Fig. 338

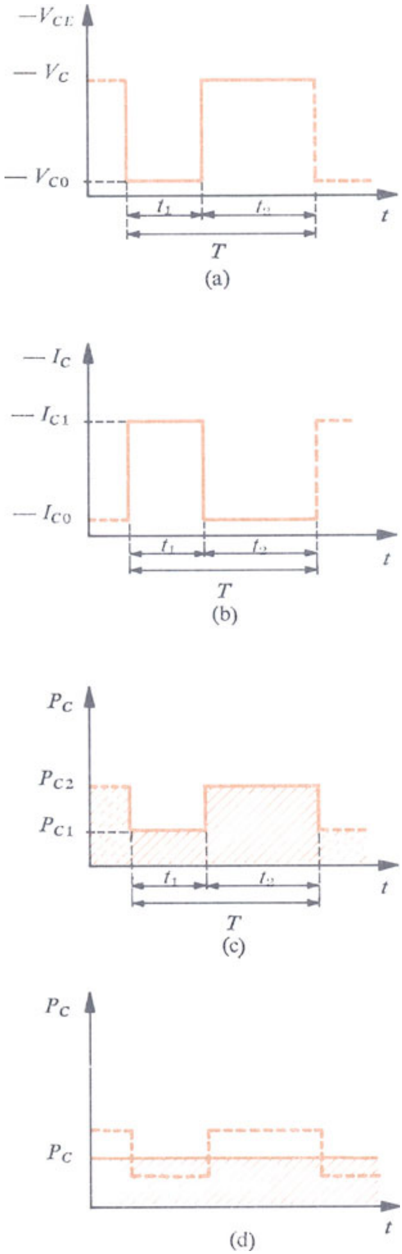


Fig. 339

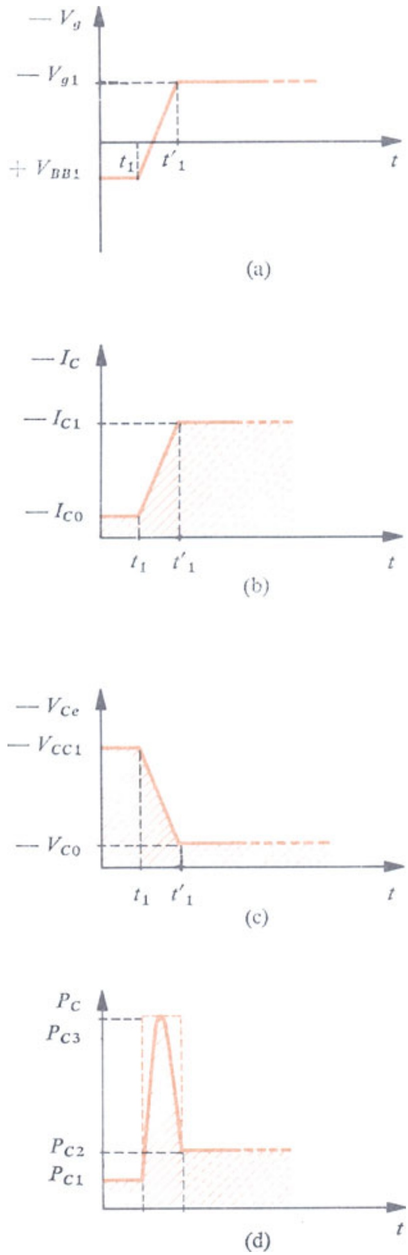


Fig. 340

during the cut-off period and P_{CS} is associated with the power dissipated in the collector of the same transistor during the period of saturation.

We have just defined the variations with time of the collector–emitter voltage and collector current. The corresponding signals are shown in Figs. 339a and b respectively. The period of saturation, that is the time of conduction of the transistor, is represented by t_1 on the axes of the two figures; the cut-off period is associated with the time t_2.

Mean power dissipated in the transistor

The real values of the power dissipated in the collector of the transistor during these two periods are shown in Fig. 339c.

During the cut-off stage the power dissipated in the transistor is equal to P_{C1} and during the saturation period it is given by P_{C2}. The brown shaded area then represents the mean power dissipated in the collector of the transistor.

$$P_{C1} = P_{CO} \qquad P_{C2} = P_{CS}.$$

It is now very important to determine the influence of these variations of power on the power effectively dissipated in the collector of the transistor during continuous operation. The data supplied by the manufacturer always give the power dissipated in continuous operation; i is therefore necessary to pass from instantaneous values of power to a value of the power dissipated in continuous operation.

If t_1 represents the time of the period of conduction and t_2 the period of cut-off, then knowing that P_{C1} is associated with the saturation state and P_{C2} with the cut-off state, the mean power dissipated in the collector of the transistor is given by the relation:

$$P_C = \frac{P_{C2}t_1 + P_{C1}t_2}{T} \qquad \text{or} \qquad P_C = \frac{V_{CEK}I_{CS}t_1 + V_{CC}I_{CO}t_2}{T}.$$

This mean power lies between the two values P_{C1} and P_{C2} (brown shaded area in Fig. 339d).

We have so far assumed instantaneous triggering, that is the complete absence of transient effect in the variations of voltage, current and power.

In rapid switching it is difficult to obtain such signals since the time of transition from one state to another can no longer be ignored.

Influence of transient effects

Let us therefore assume that the variation of input voltage is not instantaneous but progressive, that is a voltage changing from $+V_{BB1}$ to $-V_{g1}$ from instant t_1 to instant t_1'.

At instant t_1 the collector current is low (cut-off state) and it then rises steadily to the value $-I_{C1}$, which it reaches at instant t_1' (Fig. 340b). The collector–emitter voltage is a maximum at instant t_1 (Fig. 340c).

Before instant t_1 the transistor was cut-off and the power dissipated in the collector was equal to P_{C1}. After instant t_1' the transistor is saturated and the power dissipated in it is equal to P_{C2} (Fig. 340d).

During the transient period, that is from instant t_1 to instant t_1', the instantaneous values of power are given by the product of the corresponding values of the current and voltage. The power varies as shown in the curve in Fig. 340d.

Let us now consider a complete signal. The collector–emitter voltage changes from a maximum value ($-V_{C1}$ in Fig. 341) to a minimum value at the instant when the signal is actually applied at the input of the transistor.

If t_1 represents the period of conduction and t' the time taken for the change from one state to another, the collector current changes during the transient period from the minimum value $-I_{CO}$ associated with the cut-off state (Fig. 342) to the maximum value $-I_{CS}$.

The transistor then returns to the cut-off state and the voltage increases from $-V_{CO}$ to $-V_{C1}$ during the time t'', and simultaneously the current falls from $-I_{CS}$ to $-I_{CO}$. In order to define the different amounts of power dissipated in the collector of the transistor as a function of time, all we have to do is to determine for each stage the product of variations of the voltage and current defined earlier.

Calculation of the amounts of power dissipated during the various stages

The transistor is in the cut-off state. The power dissipated in the collector is then equal to P_{C1} (Fig. 343).

At the instant when the pulse is applied the current rises and the voltage decreases. During the transient period the power dissipated in the collector increases to the value P_{C3} and then decreases to the value P_{C2} when the transistor is saturated. This power remains at the value P_{C2} until the transistor is next switched to the cut-off state. The collector current decreases again and the collector–emitter voltage rises. Another power peak appears, rising to the value P_{C3} in Fig. 343. The power then returns to the value P_{C1}.

In order to define the mean power dissipated in the collector of a transistor it is necessary to consider the power peaks represented by the brown shaded areas in Fig. 343. If we know the values of P_{C1}, P_{C2} and P_{C3} all that is then necessary is to use the relation:

$$P_C = \frac{P_{C2}t_1 + P_{C1}t_2 + 2(P_{C3}t')}{T}.$$

By means of this expression we can obtain the power dissipated in continuous operation of the transistor; this power is now at a level higher than P_{C1} and P_{C2} (powers associated with the cut-off and saturation states).

The presence of power pulses of appreciable duration causes an increase in the power dissipated in the collector of the transistor. The greater the

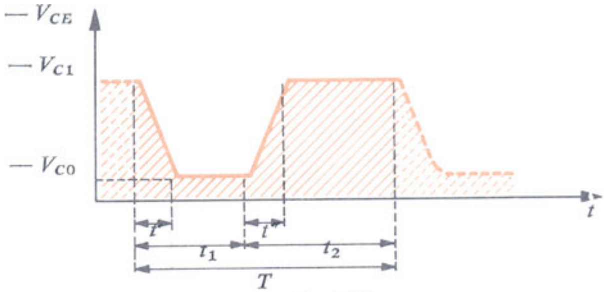


Fig. 341

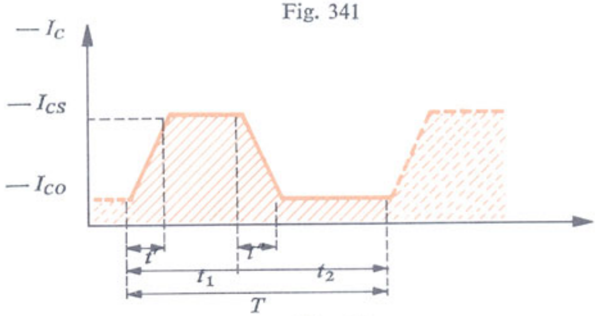


Fig. 342

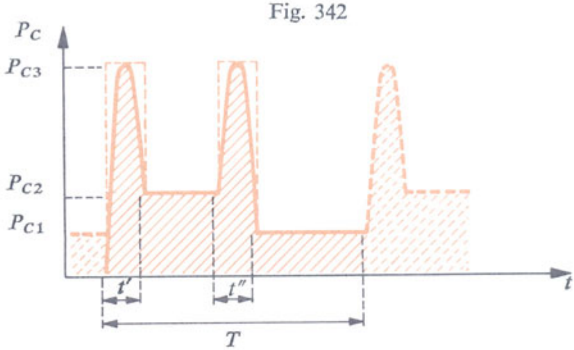


Fig. 343

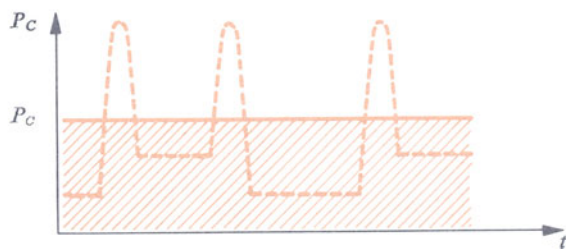


Fig. 344

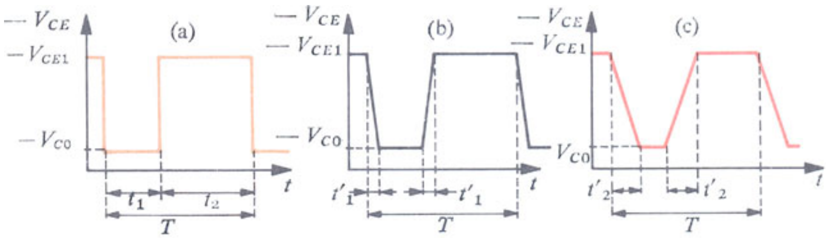


Fig. 345

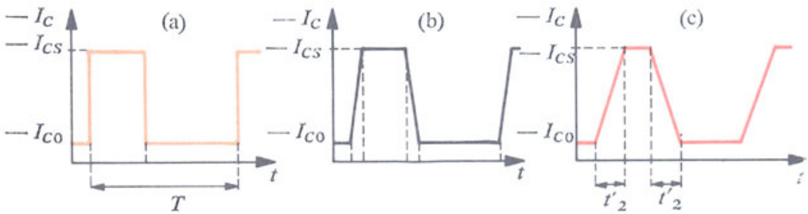


Fig. 346

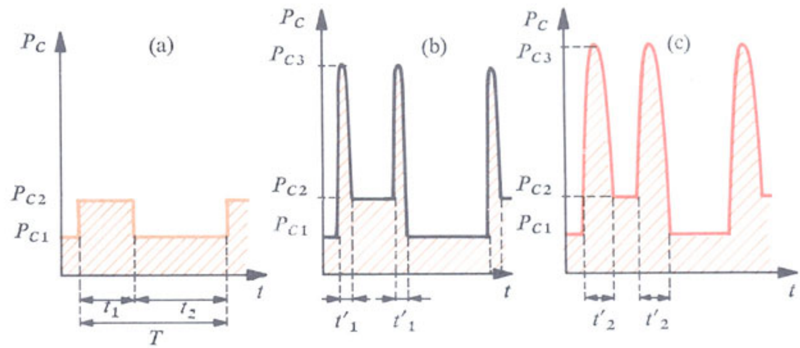


Fig. 347

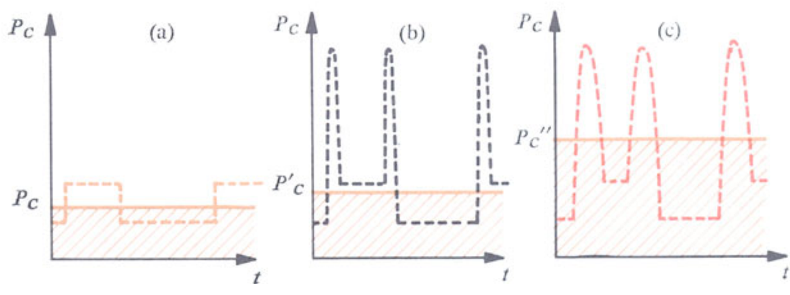


Fig. 348

transient period with respect to the pulse duration the more rapidly will this increase take place.

The power dissipated during the two transient periods may be represented by a rectangle the area of which occurs in the quantitative expression for this power, namely:

$$P_{C3} = \frac{V_{CC} \cdot I_{CM}}{4}.$$

Finally, the switching time has a considerable influence on the power dissipated in the collector of the transistor.

Figure 345a shows the variations with time of the collector–emitter voltage, the switching time being considered negligible. The saturation period is given by the duration t_1 and the cut-off period by t_2. In Fig. 345b the variations of the collector–emitter voltage are no longer instantaneous but require a certain time, namely t_1' on the abscissa.

In Fig. 345c the switching transients become very long, i.e. t_2' much larger than t_1'.

The variations with time of the current correspond exactly to the variations of voltage defined earlier.

Influence of the lengthening of the switching time

In the absence of transition time (Fig. 346a) the collector current changes instantaneously from $-I_{CO}$ to $-I_{CS}$ and vice versa.

For a transition time equal to t_1', the variations of the collector current take only place gradually from $-I_{CO}$ to $-I_{CS}$ and vice versa (Fig. 346b).

For an even longer time t_2', the variations shown in red in Fig. 346c become very large.

For each of these cases it is possible to represent the values of the instantaneous power dissipated in the collector of the transistor.

In the absence of transient effects the power changes from P_{C1} to P_{C2} and vice versa (see the brown shaded area in Fig. 347a).

For a relatively short transition time, two power pulses of value P_{C3} are added to the amounts of power dissipated in continuous operation (Fig. 347b).

For a longer transition time, the power pulses considerably greater than P_{C2} are represented by a shaded area which is very large in relation to the powers dissipated during the cut-off period (P_{C1}) and during the period of conduction (P_{C2}).

In order to calculate the power effectively dissipated in continuous operation in the collector–base junction of the transistor, all that is necessary is to make use of the formula mentioned earlier, for each of these three cases, namely:

$$P_C = \frac{P_{C2}t_1 + P_{C2}t_2 + P_{C3}t'}{T} \qquad T = t_1 + t_2 + 2t'.$$

In the absence of transients, this power is shown in Fig. 348 between the power dissipated in cut-off and the power associated with the state of saturation.

In the case of a relatively rapid transition with respect to the pulse duration (that is $t_1' \ll t_1$) the mean power dissipated in the collector of the transistor is located at a higher level (P_C' in Fig. 348b).

When the transition time increases ($t_2' < t_1$), the power dissipated in the collector in continous operation (P_C'') becomes much larger than the corresponding powers in the cut-off and saturation states (Fig. 348c).

4.6. Resistive load

Our discussion so far has been concerned with the use of a transistor with a resistive load 'R_L' (Fig. 349a). On the network of output characteristics of the transistor associated with the load line, the two operating states are located in positions 1 and 2. Position 1 corresponds to the cut-off of the transistor and position 2 to saturation. When the transistor changes from one operating state to another, the working point shifts:

(a) from position 1 to position 2 (along the load line) 'from cut-off to saturation';

(b) from position 2 to position 1 (still on the load line) 'from saturation to cut-off'.

The powers are calculated from the product of the currents and voltages obtained by the projection of the working points on the ordinates and abscissae of the network of characteristics.

4.7. Inductive load

Another very common method of use in switching is to load the transistor with an inductance L. We have already determined the influence of this inductance on the shape of the current flowing through it (collector current of the transistor). We will now attempt to define, by means of the network of output characteristics (Fig. 350b), the effective path of the working point when the transistor is switched from cut-off to saturation and vice versa.

In stage 1 the transistor is cut-off; it then changes its state and approaches saturation; from point B it becomes saturated and the working point reaches position C.

The transistor now receives the command to return to the cut-off state. A certain amount of energy is stored in the inductance and the working point shifts from position C to position D (stage 3 shown in yellow in the figure). The inductance then discharges and the collector current decreases as a function of time (stage A shown in black in the figure).

The displacement of the working point when the load is inductive takes place along a path very different from the simple straight line associated with the purely resistive load.

It is very important that the extreme positions *A*, *B*, *C* and *D* of the working point should be determined accurately in order to avoid the

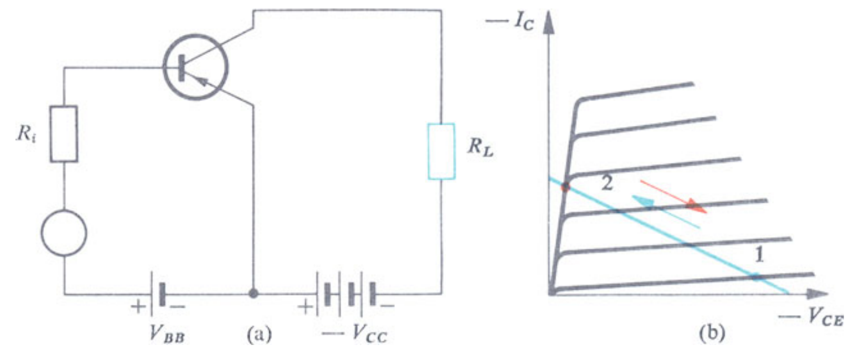


Fig. 349

breakdown of the transistor. The projection of point *C* marks the maximum value of the collector current and point *A* the the limiting value of the collector–emitter voltage. In addition to these two limiting positions, point *D* certainly constitutes the most hazardous operating condition because of the high values of voltage and current associated with it. The risk of breakdown is however limited, the time of transition to the zone *D*

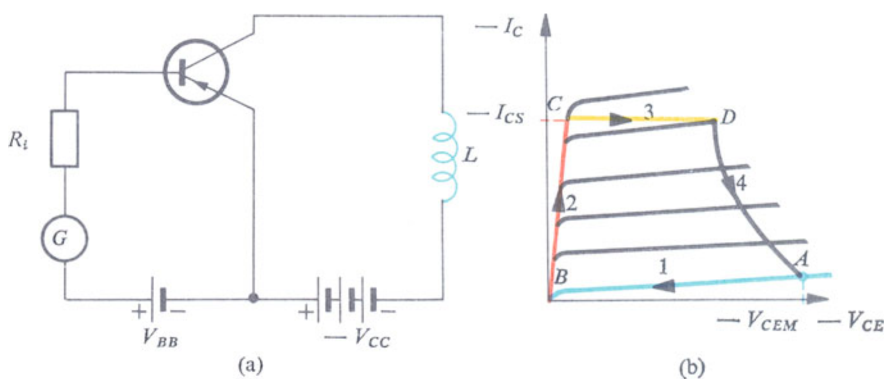


Fig. 350

is always very short, and the positions *C* and *A* are by far the most important. They should take into account the maximum values of current and voltage specified by the manufacturer.

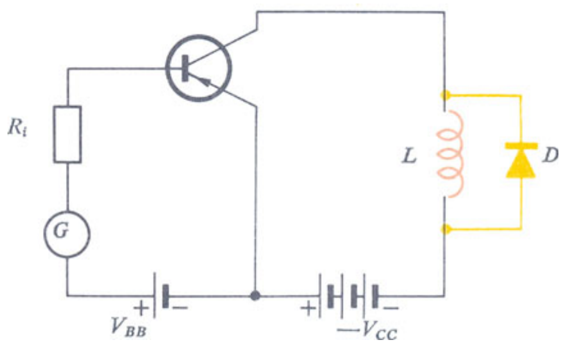


Fig. 351

It should be remembered that these limiting values of collector–emitter voltage should never be exceeded even for very short periods.

The circuit in Fig. 351 shows a transistor switched from cut-off to saturation and vice versa, loaded with an inductance.

On the network of output characteristics of this transistor (Fig. 352) the brown curve represents the effective path of the working point during the different stages mentioned on the previous page. During the transition

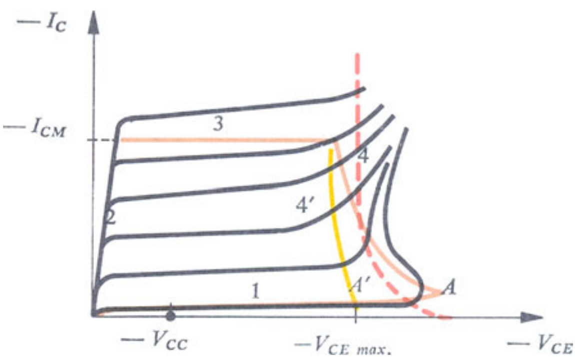


Fig. 352

from cut-off to saturation, that is from stage 1 to stage 2, the danger of breakdown or destruction is relatively small. On the other hand, during the final phase of stage 2 the collector current reaches the saturation value which should always be less than $-I_{CM}$ (maximum permissible value of the collector current).

During stage 3 the power dissipated in the collector of the transistor becomes much larger, and it is therefore necessary to ensure that the corresponding duration is shorter than the integration time.

In stage 4 the collector–emitter voltage is very high. It may then exceed the maximum value specified by the manufacturer. The projection of point A on the abscissa corresponds to a collector–emitter voltage much higher than $-V_{CEM}$.

4.8. Influence of a diode in parallel with the inductance

Is there any danger in locating the working point in conditions of this sort? If we examine the network of output characteristics of the transistor, we see that for point A there is collector–emitter voltage such that the output of the transistor behaves as negative-resistance component.

This voltage is much higher than the maximum permissible voltage and there is therefore a danger of punch-through of the transistor.

It is therefore always necessary to avoid this part of the path, and to connect a diode in parallel with the inductance.

The role of this diode is to prevent voltage surges across the load. It affects the path of the working point on the network of output characteristics; the curve corresponding to stage 4 moves to position 4′ (in yellow in Fig. 352).

The working point does not move beyond position A', which is the limit associated with the maximum value of the collector–emitter voltage on the abscissa of the output characteristics.

This diode has the advantage that it suppresses all effects of surge voltages across the load, and prevents the danger of breakdown of the transistor due especially to punch-through.

The danger of punch-through is explained by the fact that the collector–emitter voltage is applied in full to the terminals of the collector–base junction of the transistor; this reverse bias involves a very large widening of the corresponding barrier, resulting in the danger of punch-through of the base of the transistor.

4.9. Instantaneous values of the power dissipated in the transistor

In addition to the maximum values of the voltage and current, it is necessary to calculate the instantaneous powers dissipated in the collector of the transistor.

If we examine Fig. 353 it will be seen that this power varies very widely. If we take as the limiting value for this transistor a maximum permissible dissipated power equal to P_{C3} we can plot a hyperbola of equal power. This hyperbola represents the geometric locus of all the points corresponding to this power P_{C3}.

We will now consider the path of the working point with an inductive load. The amounts of power dissipated may be much greater than the power P_{C3} specified by the manufacturer. Its maximum value is P_{C1} (in red in Fig. 353) and its intermediate values are P_{C2} (in brown in the same figure).

When the working point is located above the hyperbola of equal power, that is between the two yellow points shown on the red dotted line, the power dissipated in the collector of the transistor is greater than the maximum permissible power. The limiting power is defined for continuous operation and it is therefore possible, in a transient period, to dissipate considerably higher powers. It is not out of the question for an instantaneous power equal to P_{C1} to be dissipated, provided its duration is limited; it is also possible to represent the variations of the power with time and to calculate the effective value of the mean power dissipated in the collector of the transistor.

With a purely resistive load the problem would seem much simpler. If we define, as before, a hyperbola of equal power associated with a maximum collector power equal to P_{C3} (Fig. 354), the load line corresponding to an optimum output power should be a tangent to the hyperbola of equal power.

Where efficiency is concerned, with extreme positions of the working point, this point will always be located for values of the supply voltage very little different from the maximum permissible collector–emitter voltage.

It should however be possible to locate the load line with respect to the maximum value of current; this solution is less interesting because it involves a stronger quiescent current, resulting in a lower efficiency.

4.10. Controlled power

The power dissipated in the collector–base junction of the transistor certainly represents one of the most important items of information,

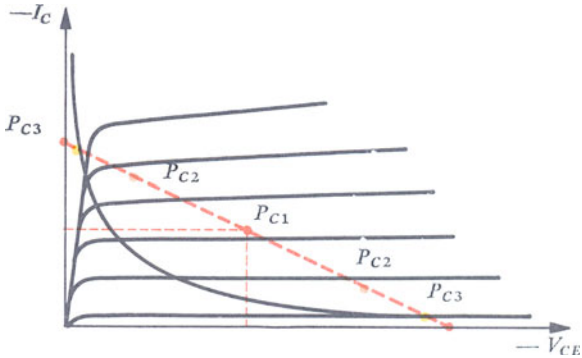


Fig. 353

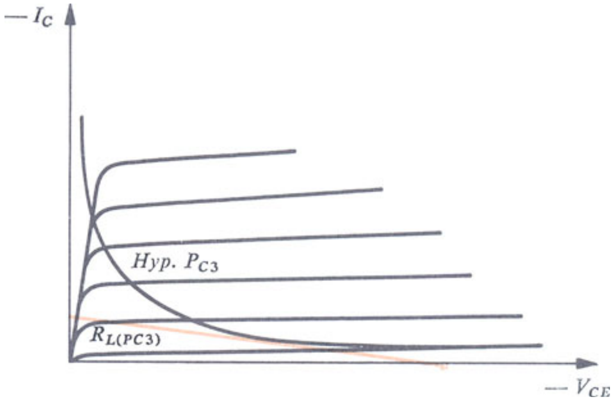


Fig. 354

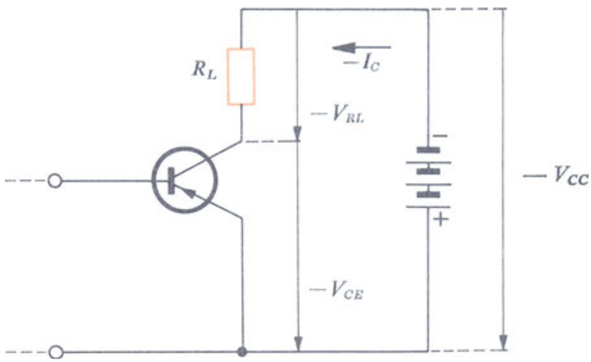


Fig. 355

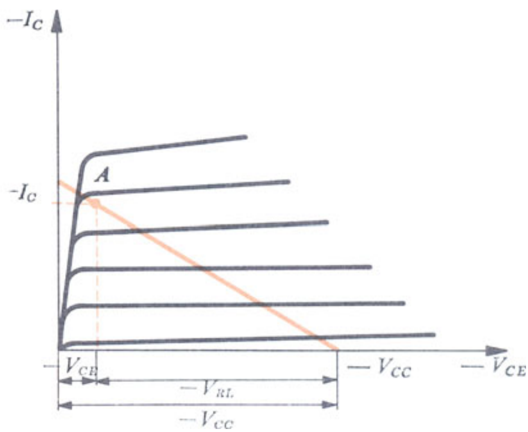


Fig. 356

especially in power switching. In audio-frequency operation we mentioned earlier the possibility of associating with this collector power the maximum output power. Is it possible in these same conditions to define a factor for the calculation of the maximum permissible output power in switching with relation to the maximum power that the collector can dissipate?

With pulse operation we speak rather of controlled power than output power. What is meant by controlled power?

Figure 355 illustrates a transistor in common-emitter configuration loaded by a resistance R_L and supplied by a group of batteries $-V_{CC}$.

By definition the power dissipated in the collector of the transistor is given by the product of the collector-emitter voltage and the current flowing through the collector of the transistor, namely:

$$P_C = -V_{CE} \cdot -I_C.$$

The power supplied by the battery is defined as the product of the supply voltage and the current flowing through the collector of the transistor, namely:

$$P_B = -V_{CC} \cdot -I_C.$$

By 'controlled power' we mean that power which can be controlled effectively at the terminals of the load resistance, and it is given by the relation:

$$P_{cont} = -V_{RL} \cdot -I_{RL} = -V_{RL} \cdot -I_C.$$

The controlled power is dependent on the product of the collector current and the voltage across the load resistance.

Controlled voltage

In these conditions the value of the controlled voltage is defined by the expression:

$$V_{cont} = -V_{RL} = -V_{CC} - (-V_{CE}).$$

We will now consider the network of output characteristics of the transistor in Fig. 356. The working point is in position A on the load line. Is it now possible to know both the collector-emitter voltage and the voltage across the load resistance?

All we need do for this is to project point A on to the abscissa. The distance separating the origin from the intersection of this vertical line, shown in dotted broken lines in the figure, gives the value of the collector-emitter voltage. The voltage across the load resistance is then equal to the voltage of the battery less the collector-emitter voltage of the transistor, that is V_{RL} in the figure.

In this same way we can define the value of the collector current by the projection of point A on to the axis $-I_C$.

Maximum controlled power

The maximum controlled power is a function of the product of the maximum voltage appearing across the load resistance and the maximum current that can flow through it, namely:

$$P_{cont\,max} = -V_{RL\,max} \cdot -I_{C\,max}.$$

By means of the network of output characteristics of the transistor we can express these two concepts and obtain from them the maximum value of the controlled power.

With a resistive load the maximum power that can be applied between the collector and emitter of the transistor is given by $-V_{CEM}$ (Fig. 357). The 'knee' voltage is designated in data sheets by $-V_{CEK}$. Consequently the value of the voltage at the terminals of the load resistance when the transistor is saturated, is equal to the difference between the maximum value of the collector–emitter voltage and its knee value, namely:

$$V_{RL\,max} = -V_{CEM} - (-V_{CEK}).$$

The value of the voltage across the load resistance increases as the maximum collector–emitter voltage increases for a given transistor.

The collector current plays an important part in the expression for the maximum controlled power in the load of a transistor. For this reason it is advisable to mark on the family of output characteristics of a transistor the limiting conditions for the choice of a collector current equal to its maximum value $-I_{CM}$. If we assume a load line corresponding to a resistance of the same value as before, the maximum value of the current locates the working point in the position shown in red in Fig. 358. The transistor is then saturated and the collector–emitter voltage is a minimum.

If we know the maximum values of the collector–emitter voltage and of the voltage across the load resistance, as well as the limiting value of the current in the collector circuit, we can easily calculate the maximum value of the power available across the load resistance, namely:

$$P_{cont\,(max)} = (-V_{CEM} - (-V_{CEK})) \cdot -I_{CM}.$$

It is nevertheless possible to use diagrams for the calculation of the ratio between the maximum controlled power and the maximum power dissipated in the collector of the transistor.

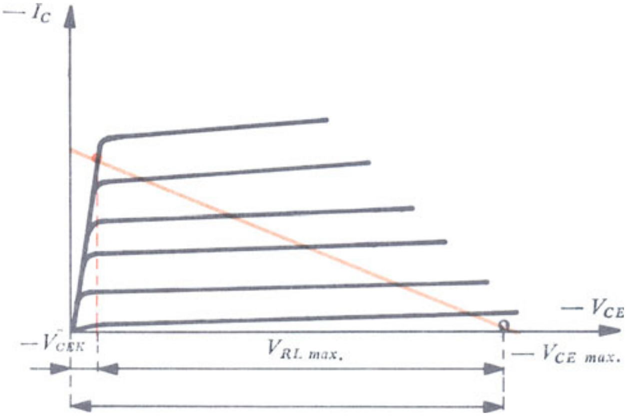


Fig. 357

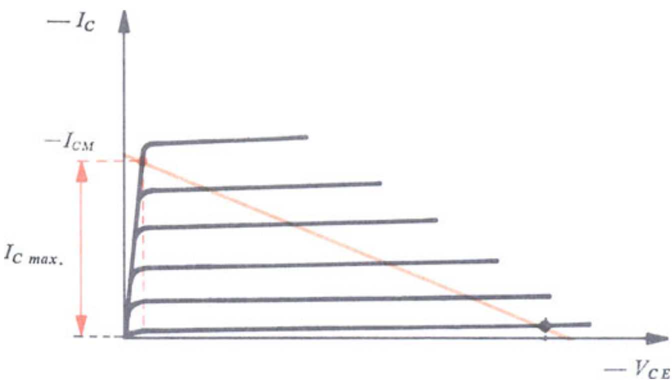


Fig. 358

If we refer to the various expressions for the calculation of the powers associated with a switching transistor, we see that it is possible, by examinin the family of output characteristics of the transistor in Fig. 359, to determine them again.

The power dissipated in the collector of the transistor when it is operating

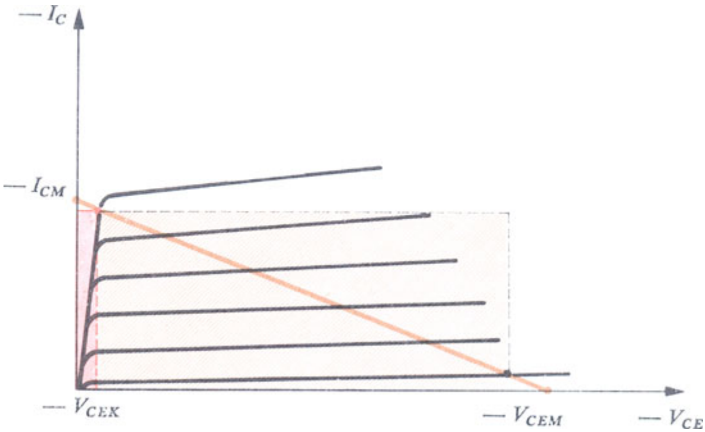


Fig. 359

in the saturated state is given by the red shaded area in Fig. 359:

$$P_{CS} = -V_{CEK} \cdot -I_{CM}.$$

The maximum power dissipated in the load resistance is then defined by

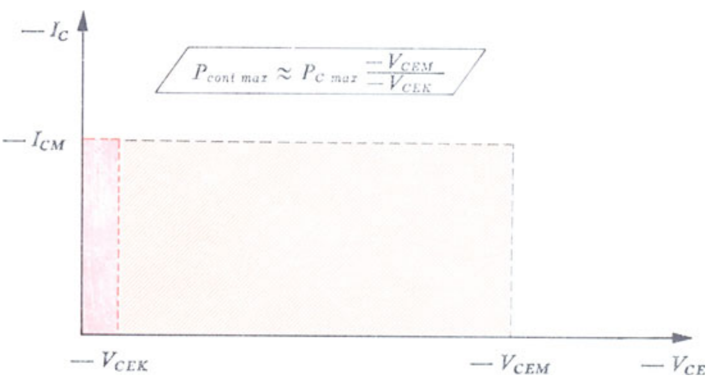


Fig. 360

the brown shaded rectangle in the same figure:

$$P_{cont\ (max)} = (-V_{CEM} - (-V_{CEK}) \cdot -I_{CM}.$$

The power supplied by the battery is then equal to the sum of the two above powers and it is dependent on the group of two rectangles (red and brown in the figure):

$$P_B = P_{CS} + P_{cont\ (max)}.$$

In order to determine the proportionality factor between the maximum controlled power and the maximum power dissipated in the collector of the transistor, all we need do is to establish the relationship connecting the two areas of these rectangles. The reader is referred to Fig. 360 which represents:

(a) on the one hand the power dissipated in the transistor (red shaded rectangle); this power is defined by the expression P_{CS} corresponding to the maximum value of the power dissipated in the collector of the transistor;

(b) on the other hand the maximum controlled power (brown shaded rectangle); it is associated with the power dissipated in the collector.

These two rectangles have a common dimension, namely the width of the brown rectangle, which is approximately equivalent to $-I_{CM}$. The factor required, therefore, depends exclusively on the ratio between the width of the red rectangle and the length of the brown rectangles.

These two lengths are shown on the voltage axis and correspond to the limiting values of the collector–emitter voltage: namely $-V_{CEK}$ for the low voltages and $-V_{CEM}$ for the high voltages. The ratio of the two rectangles is therefore directly dependent on the quotient of the two voltages:

$$-V_{CEM}/-V_{CEK}.$$

The area of the red rectangle gives the maximum power dissipated in the collector of the transistor, and that of the brown rectangle the maximum controlled power in the load resistance; the ratio between these two powers is given directly by the expression:

$$P_{cont\ (max)}/P_C = -V_{CEM}/-V_{CEK}.$$

This relation is only approximate, as the residual characteristics of the transistor have been ignored. However, it is usually adequate in practice for the very rapid calculation of the power that can be controlled in the load from the maximum power dissipated in the collector of the transistor, which is the only item of information supplied by the manufacturer, namely:

$$P_{con\ (max)} = P_{CS} \cdot -V_{CEM}/-V_{CEK}.$$

4.11. Equivalent thermal circuits

From the thermal point of view transistors must always be divided into two main categories:

(a) small-signal transistors;

(b) large-signal transistors.

In order to define the transient thermal behaviour of each type of transistor more satisfactorily we must remind ourselves of the construction of the static equivalent thermal circuits.

Small-signal transistors

A small-signal transistor almost always resembles the drawing shown in Fig. 361a. In this transistor most of the power is dissipated in the collector–base junction. In order to construct the equivalent thermal circuit of a transistor of this type it is first necessary to define the two extreme values of temperature. The top line in Fig. 361b represents the temperature equal to T_j. On the other hand, the minimum value of the temperature T_a is shown by a black line at a lower level. A certain number of resistive elements are connected between these two lines of thermal potential.

The simplified thermal equivalent circuit contains at least two resistances:

(a) resistance $R_{th(j-c)}$ or the resistance between the junction and the case of the transistor;

(b) resistance $R_{th(c-a)}$ representing the case-to-ambient resistance.

The manufacturer always supplies a simplified equivalent circuit (Fig. 361c) consisting only of a resistance ($R_{th(j-a)}$), which when expressed in degrees per milliwatt represents the thermal resistance between the junction and the surroundings.

Large-signal transistors

The thermal equivalent circuit for large-signal transistors (Fig. 362a) is more complicated. There are several resistances between the highest level of thermal potential T_j and the coldest point T_a:

(a) $R_{th(j-b)}$ the thermal resistance junction-to-mounting base;

(b) $R_{th(b-h)}$ resistance of the insulating strip;

(c) $R_{th(h-a)}$ thermal resistance of the heatsink.

For the sake of simplicity this equivalent circuit may perhaps be represented as shown in Fig. 362, where the two thermal resistances are:

(a) the resistance between the junction and the heatsink $R_{th(j-h)}$;

(b) the thermal resistance of the heatsink associated with the power transistor ($R_{th(h-a)}$).

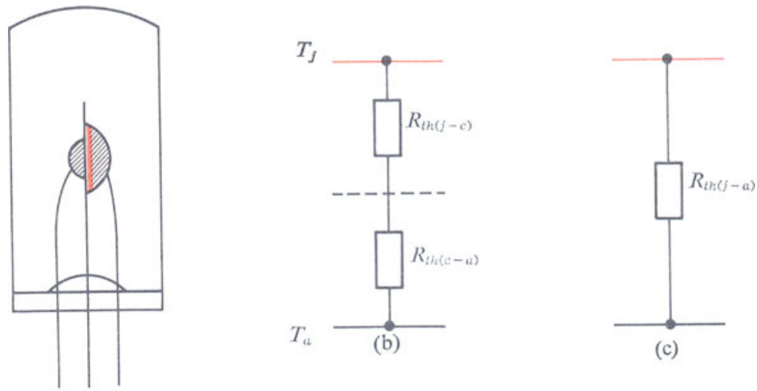


Fig. 361

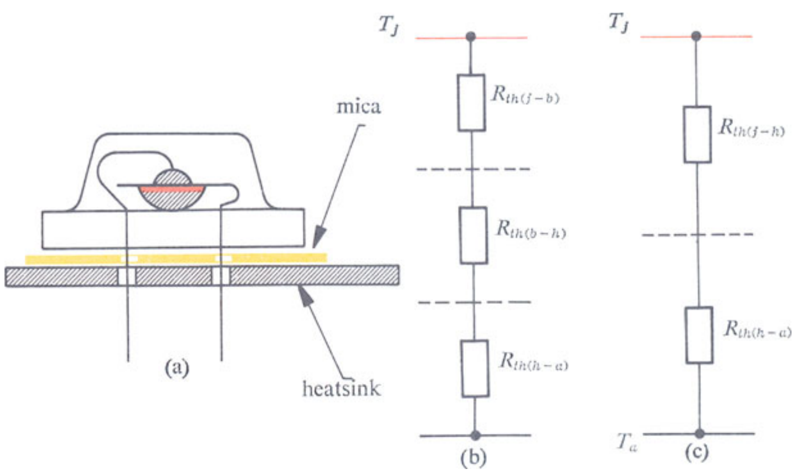


Fig. 362

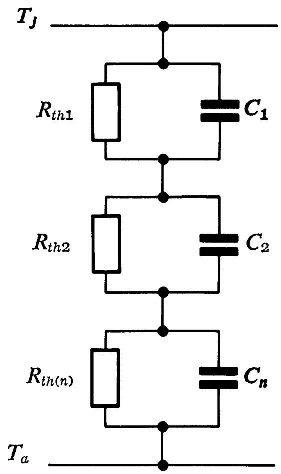


Fig. 363

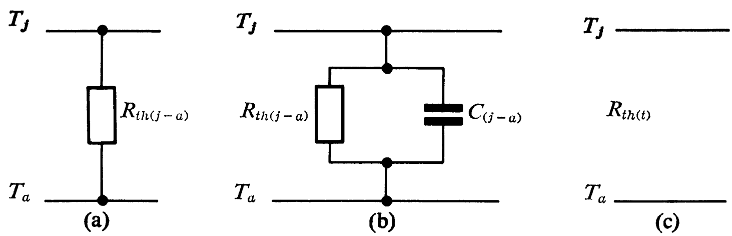


Fig. 364

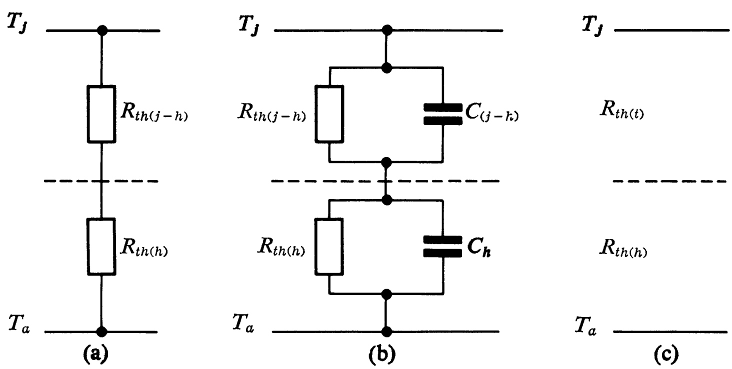


Fig. 365

4.12. Dynamic equivalent thermal circuits. Transient thermal impedance

If in practice we wish to construct the dynamic equivalent circuit of a transistor between the two lines of thermal potential T_j and T_a, account must be taken of the fact that certain of the elements are charged and discharged with the same cyclic variations as the variations of the heat energy.

Referring to Fig. 363, it will be seen that the thermal equivalent circuit is represented by the connection in series of a certain number of *RC* elements. In a circuit of this sort the elements R_{th1}, R_{th2} and $R_{th(n)}$ represent the different thermal resistances associated with this equivalent circuit, and the elements C_1, C_2 and C_n the capacitive elements. Finally, the dynamic thermal equivalent circuit of a transistor can be represented by resistances and capacitors in parallel. If we refer back to the simplified thermal equivalent circuit of the 'small-signal' transistor (Fig. 364a) we see the necessity of improving this circuit by connecting a capacitance $C_{(j-a)}$ in parallel with the resistance $R_{th(j-a)}$. Once we know these two elements we can calculate the value of the transient thermal impedance $R_{th(t)}$ of Fig. 364c. Its value is dependent on the pulse duration and on the repetition frequency of the signal; the curves representing the variations of the transient thermal impedance as a function of these parameters will be discussed later.

With large-signal transistors, the simplified equivalent circuit of Fig. 365a, which consists of two thermal resistances $R_{th(j-h)}$ and $R_{th(h)}$ in series, should be modified; it then takes the form shown in Fig. 365b. Associated with the thermal resistance $R_{th(j-h)}$ there is a thermal capacitance $C_{(j-h)}$ intrinsic to the transistor and a thermal capacitance C_h intrinsic to the thermal resistance $R_{th(h)}$ of the heatsink. It is now easy, since we know the effective values of these two impedances, to construct a new and simpler equivalent circuit, as shown in Fig. 365c.

In this new figure $R_{th(t)}$ represents the transient thermal impedance intrinsic to the transistor and $R_{th(h)}$ the transient thermal resistance of the heatsink.

With small-signal transistors, the two transient thermal impedances are connected on the one hand with the pulse duration, and on the other hand with the duty factor. It is therefore very important, especially in switching, to define the operating conditions of the transistor exactly so as to associate with it precise values of the transient thermal impedance. We will now pass to the demonstration of the effect of this parameter on the maximum values of power dissipated in the collector of the transistor.

4.13. Influence of the transient thermal impedance

Depending on the shapes of the signals received at the output of the transistor, the equations for the calculation of the maximum power dissipated in the collector may be very different.

'Small-signal' transistors

With small-signal operation, Fig. 366 illustrates the case in which the transient thermal resistance $R_{th(t)}$ is used. This transient value can be used only for pulses of duration less than 300 seconds. In these conditions the expression for the maximum collector power is given by the relation:

$$P_{CM} = \frac{T_j - T_a - R_{th(j-a)} \times P_{C1}}{R_{th(t)}}.$$

On the other hand if the pulse duration is longer than the 300 seconds mentioned above, for example the pulse shown in Fig. 367, the transient thermal impedance is no longer usable, and we must introduce the thermal resistance 'junction–ambient' ($R_{th(j-a)}$ in Fig. 367a). The maximum collector power is given by the relation:

$$P_{CM} = \frac{T_j - T_a - R_{th(j-a)} \times P_{C1}}{R_{th(j-a)}}.$$

'Large-signal' transistors

With large-signal operation the limiting duration of the pulse corresponding to the use of transient thermal impedances is much shorter. In practice the thermal impedance is used for pulse durations of less than one second. The thermal equivalent circuit of the transistor corresponds to that in Fig. 368a, and the maximum collector power can be calculated by means of the expression:

$$P_{CM} = \frac{T_j - T_a - (R_{th(j-h)} + R_{th(h-a)}) \times P_{C1}}{R_{th(t)} + dR_{th(h-a)}}.$$

If the pulse duration is longer than one second (Fig. 369), the thermal equivalent circuit of the transistor is limited to the representation of the two thermal resistances $R_{th(h-a)}$ (junction–heatsink resistance) and the resistance of the heatsink. In these conditions the maximum power dissipated in the collector of the transistor is given by the equation:

$$P_{CM} = \frac{T_j - T_a - (R_{th(j-h)} + R_{th(h-a)})P_{C1}}{R_{th(j-h)} + R_{th(h-a)}}.$$

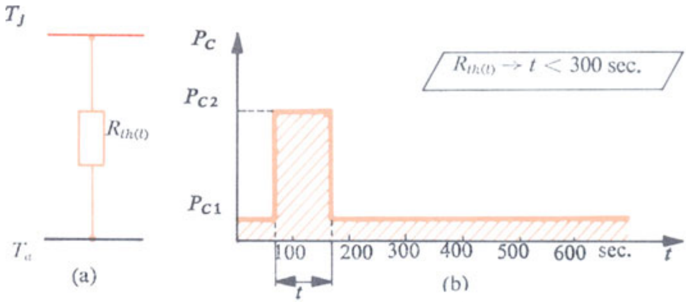


Fig. 366

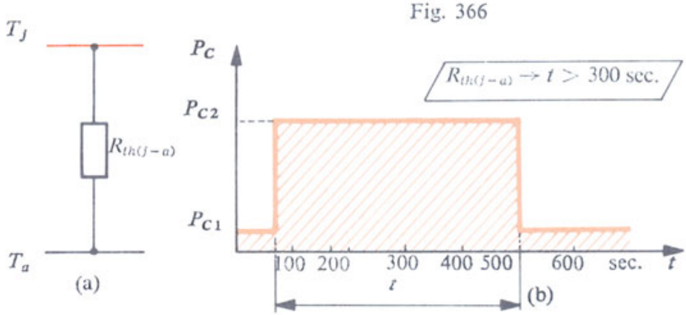


Fig. 367

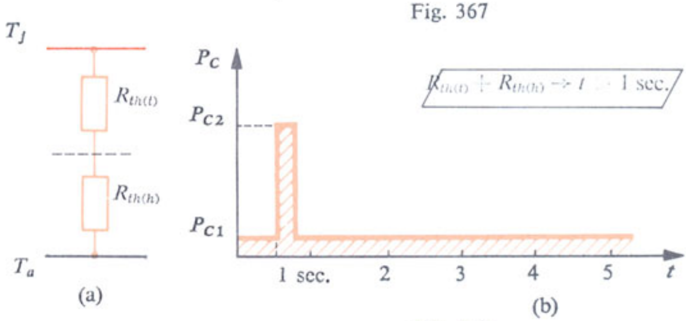


Fig. 368

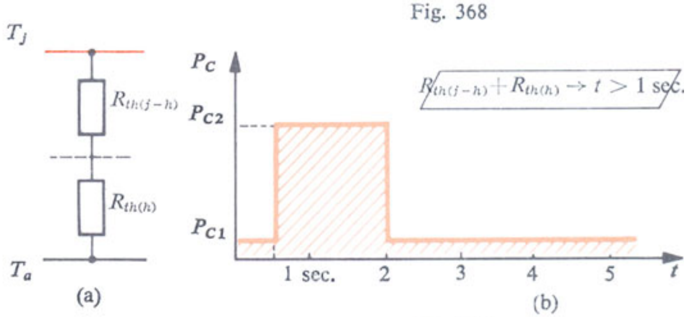


Fig. 369

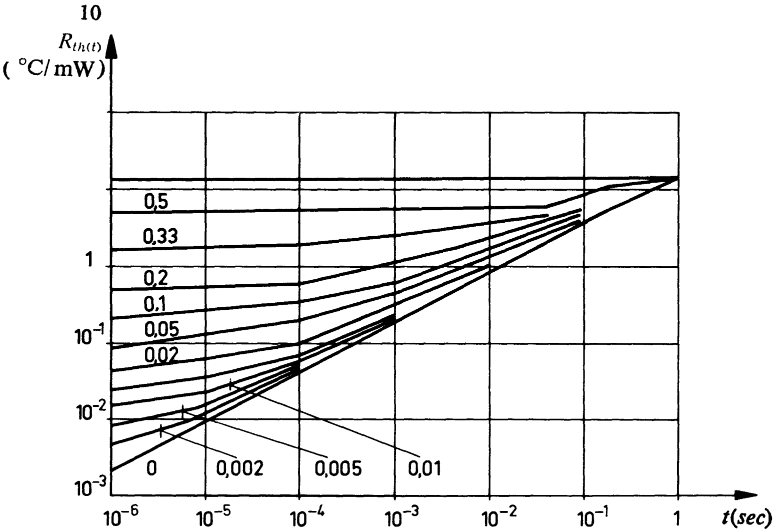


Fig. 370

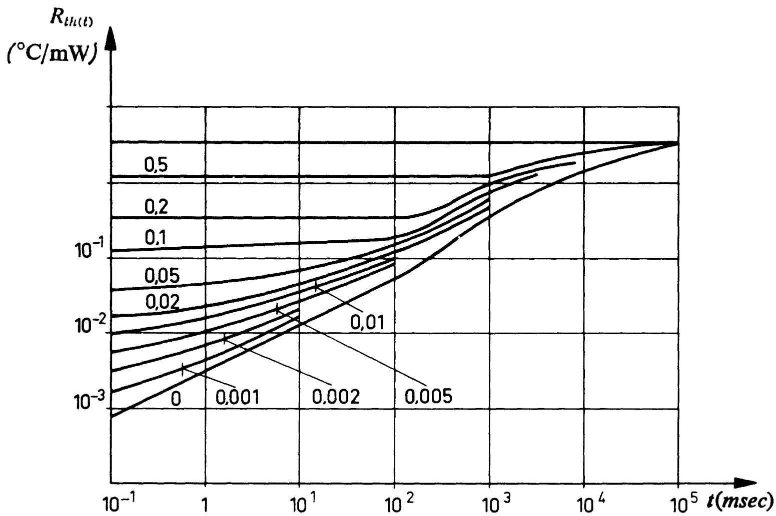


Fig. 371

The thermal impedance of a transistor is of special importance in rapid switching circuits. Given that this impedance is much lower than the nominal thermal resistance of a component, it is important to be able to define it accurately in order to enable us to determine the maximum permissible powers in switching.

We must therefore know not only the real value of the thermal impedance of the transistor but also the influence of the electrical characteristics of the signal on the value of this impedance.

Large-signal transistors. Figure 370 shows a series of curves by means of which it is possible to calculate, whatever the type of signal, the value of the transient thermal impedance. In this figure the thermal impedance is plotted on the ordinate, the thermal resistance being equal to 2 °/W, and the abscissa is graduated in time and defines the pulse duration; the duty factor *d* constitutes the third parameter. In continuous operation the thermal impedance is equal to the thermal resistance of the transistor.

In switching, with a duty factor of 0·5, the thermal impedance for pulse durations of less than one second become equal to one half the previous thermal resistance. A decrease in the duty factor causes a reduction in the thermal impedance.

Because of the lower value of the thermal impedance the maximum permissible power in transient operation is always higher than the limiting value for continuous operation.

Small-signal transistors. We will now investigate the variations of the transient thermal impedance of a 'small-signal' transistor. If we take as an example a transistor having a thermal resistance of 300 °C per watt, this resistance is represented by a straight line at the top of Fig. 371.

The ordinate is graduated in values of thermal impedance and the abscissa in time; the duty factor occurs as a parameter of the network of curves.

In this case the thermal impedance appears for signals of duration of less than 10 seconds. As for large-signal transistors it is very important to take account of the real calculated value of the transient thermal impedance with respect to the thermal resistance in order to calculate the value of the maximum thermal power that can be dissipated in the collector of the transistor without the danger of thermal runaway or breakdown.

4.14. Instantaneous junction temperature

In addition to the concept of transient thermal impedance it is important to be able to determine the real value of the temperature of the junction of a transistor.

In practice, whatever the type of transistor, manufacturers always provide information of the limiting values of temperature, namely:

(a) for germanium transistors, 75 °C without restriction as to duration, 90 °C with certain precautions and 100 °C for a limited life;

(b) for silicon transistors, 150 °C in normal operation, 175 °C with certain precautions, and 200 °C for a limited life.

The calculation of the maximum power dissipated in the collector moreover does not allow us to define accurately the real value of the temperature of the collector–base junction of the transistor, and especially to know whether this temperature exceeds the limits specified by the manufacturer.

Consequently, any serious study of the use of a switching transistor calls for the calculation of the instantaneous values of the temperature of its collector–base junction.

Let us take as an example a transistor used as a switch. From instant t_1 to the instant preceding t_2, the transistor is cut-off; no power is then dissipated in the collector (Fig. 372a). In these conditions the junction temperature is equal to the ambient temperature (Fig. 372b).

At instant t_2 the power changes suddenly and reaches the value P_{C1} and remains at this value until instant t_3; the junction temperature immediately rises to the value T_{j1}. At instant t_3 the collector power disappears and the temperature at once falls. Figure 372 shows the theoretical variations of the temperature of the junction as a function of a rectangular variation of the power dissipated in the collector. In practice, and for the same power variations (Fig. 373a), the temperature variations of the junction don't correspond any more with the rectangular signal in Fig. 372b. In fact, at instant t_2 the junction temperature starts to rise, but reaches its final value only after a certain time (t_2' in Fig. 373b).

4.15. Thermal rise time constant

The time required for the temperature to change from its minimum value T_a to its maximum value T_{j1} gives the constant of thermal rise time (θ_{th}).

At instant t_3 the junction temperature falls during the same times, namely:

$$_3{}' - t_3 = t_2{}' - t_2.$$

We will now consider a power transistor and assume that the variation

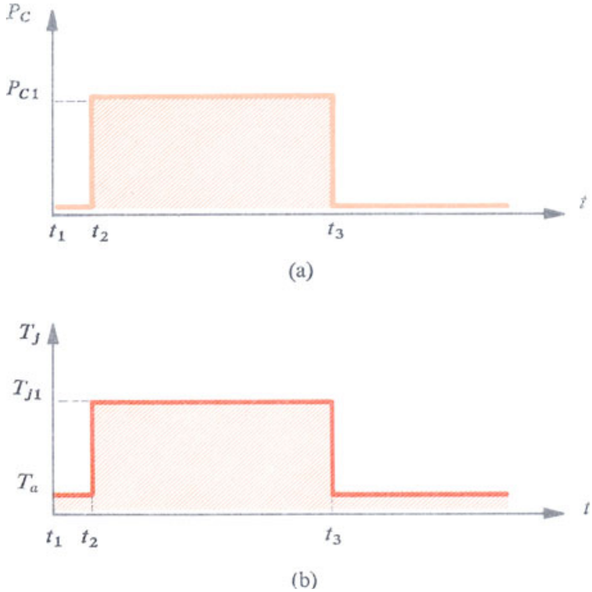


Fig. 372

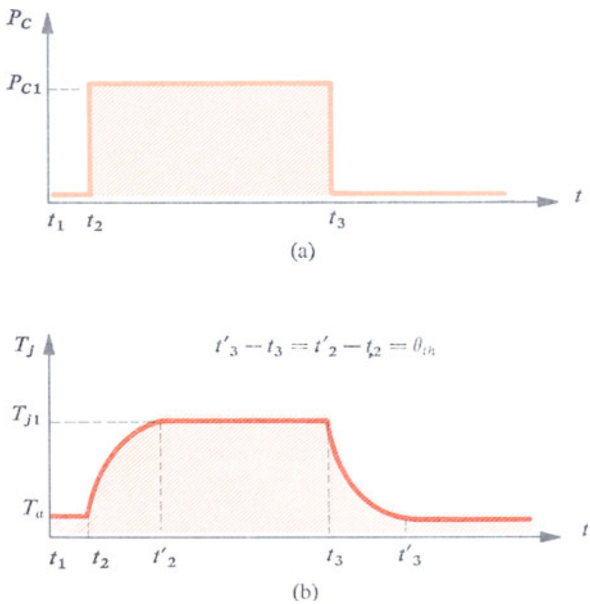


Fig. 373

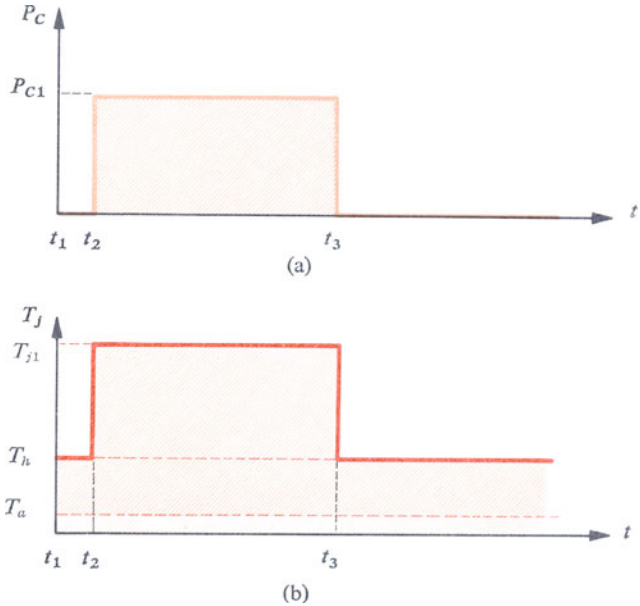


Fig. 374

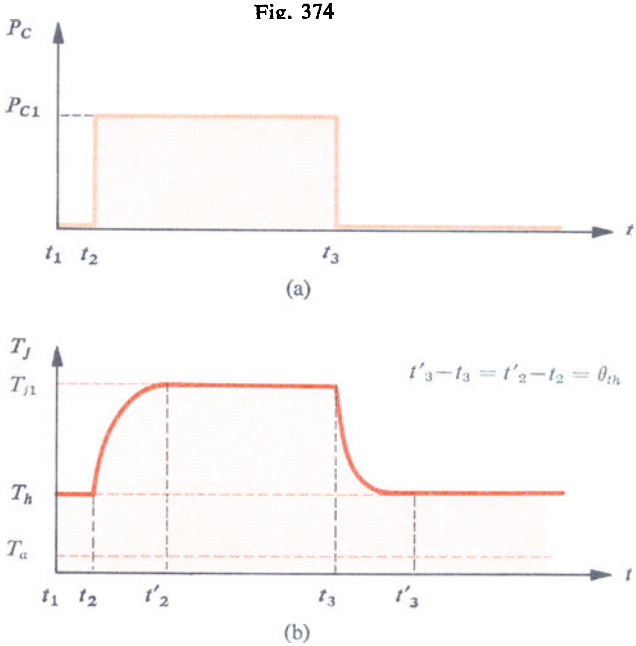


Fig. 375

of the power dissipated in the collector is such that from instant t_1 to the instant preceding t_2 the temperature of the junction is higher than the ambient temperature and equal to the temperature of the heatsink T_h.

When no power is dissipated in the collector of the transistor (Fig. 374a) the temperature of the junction remains at T_h as shown in Fig. 374b. From instant t_2 to instant t_3 a power equal to P_{C1} can be represented as a theoretical instantaneous variation of temperature; the same applies in the other direction at instant t_3 (Fig. 374b).

A more detailed study would produce equations for the variations of junction temperature very different from those in Fig. 374b. In fact, from instant t_1 to the instant preceding t_2 the absence of power dissipated in the collector causes the junction temperature to be near that of the heatsink T_h. It reaches its final value T_{j1} at instant t_2.

The thermal time constant of the transistor is given by $t_2'-t_2$.

The disappearance of all power dissipated in the collector is accompanied by a decrease in the junction temperature. This temperature reaches its minimum value at instant t_3'.

The thermal time constant of the transistor is given also by $t_3'-t_3$.

If we wish to calculate the various values of the temperature of the collector–base junction of the transistor as functions of the signal applied at the input, we must know:

(a) the thermal time constant of the transistor;

(b) the pulse duration;

(c) the duty cycle.

The next few pages will be devoted to a definition of the effect of these different characteristics on the effective temperature of the collector–base junction of the transistor.

If the thermal time constant is lower than the pulse duration and than the pulse spacing, the junction temperature will never exceed the value T_{j1} shown in Fig. 375b.

If the value of the heatsink temperature is higher than the ambient temperature, then in the example chosen the power dissipated in the collector will be different from zero. It can, however, always be ignored by comparison with P_{C1}.

The thermal time constants are short enough for the temperatures to reach their equilibrium values quickly, that is:

(a) a maximum value during the period of dissipation;

(b) a minimum value in the absence of a signal.

Thermal time constant short by comparison with the pulse duration

When the power dissipated in the collector is zero, that is from instant t_1 to instant t_2 in Fig. 376a, the temperature of the collector–base junction is equal to T_h (Fig. 376b).

At instant t_2 the temperature of the junction increases and reaches the value $T_{j\,max}$ some time later; it then falls from instant t_3 as the power dissipated in the collector decreases; the minimum value is reached at an instant between t_3 and t_4. This effect is repeated for each pulse. Since $T_{j\,max}$ defines the maximum value of the junction temperature it is possible by means of the expression:

$$P_{C\,max} = \frac{T_{j\,max} - T_h}{R_{th(t)}}$$

to calculate the maximum value of the power P_{C1} that can be dissipated in the collector–base junction of the transistor.

Shorter pulses and those with a higher recurrence frequency

From instant t_1 to the instant preceding t_2 the power dissipated in the collector is near to zero (Fig. 377a). The junction temperature, based on the first pulse, should have the value T_a. In practice, however, our study is concerned with continuous operation, and the temperature will, therefore, be $T_{j\,min}$, higher than T_h, which is itself higher than T_a. The appearance of a power pulse P_{C1} causes a rise in temperature (Fig. 377b).

At instant t_3 the pulse disappears and the temperature starts to fall, reaching the value $T_{j\,min}$ at instant t_4. With each pulse the temperature rises and falls successively, resulting in the following limiting temperatures:

(a) for low values, a level $T_{j\,min}$ higher than the temperature T_h of the heatsink;

(b) for high values the level $T_{j\,max}$.

Starting from this graphical representation it is easy, by associating with it the expression:

$$P_{C\,max} = \frac{T_{j\,max} - T_{j\,min}}{R_{th(t)}}$$

to calculate, if we know the maximum value of the junction temperature, the minimum value of this temperature and the maximum power that can be dissipated in the collector of the transistor.

We have based our discussion so far on the use of large-signal transistors, but the reasoning is the same for small-signal transistors; all we need do is to modify the two above expressions thus:

$$P_{C\,max} = \frac{T_{j\,max} - T_a}{R_{th(t)}} \quad \text{and} \quad P_{C\,max} = \frac{T_{j\,max} - T_{j\,min}}{R_{th(t)}}.$$

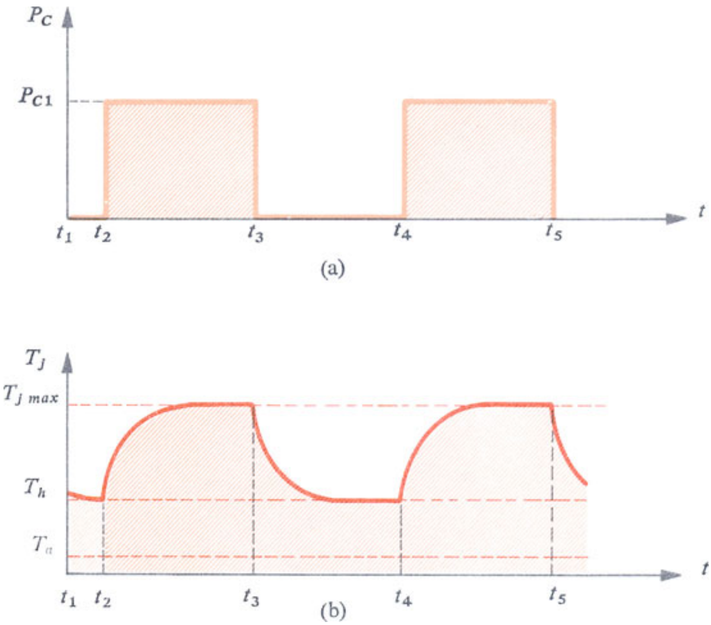


Fig. 376

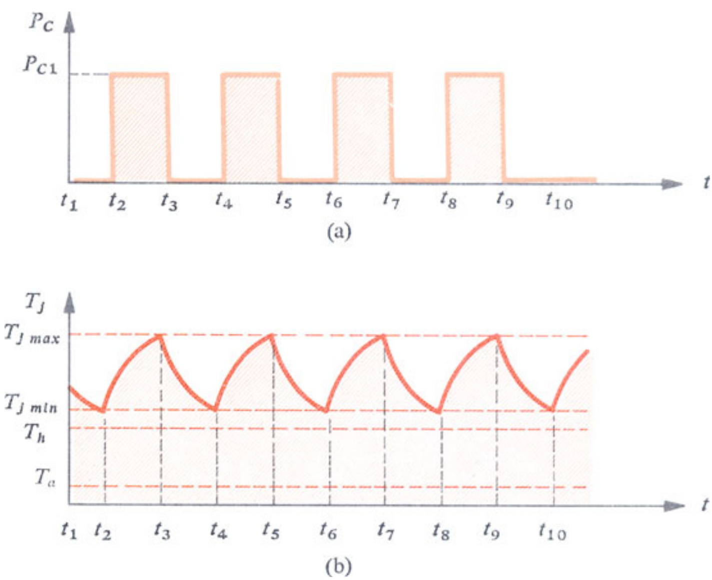


Fig. 377

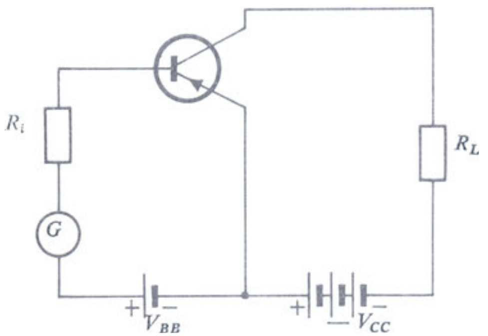


Fig. 378

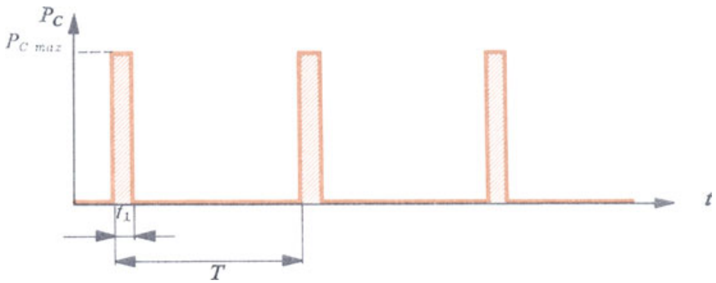


Fig. 379

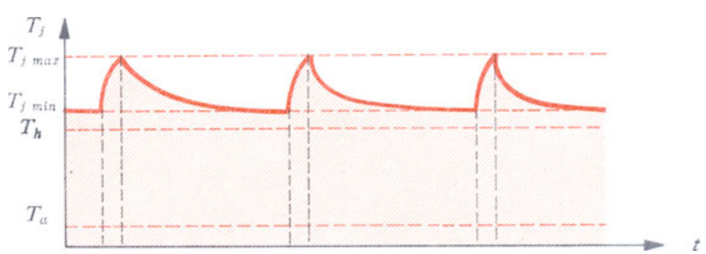


Fig. 380

Signals of short duration by comparison with their spacing

The circuit in Fig. 378 consists of a transistor driven by a generator of rectangular signals and loaded by a resistance R_L. The supply and bias are from the group of batteries $-V_{CC}$ and the battery V_{BB} respectively.

Instantaneous powers dissipated in the collector–base junction of a transistor. During the first stage the transistor is cut-off and the power dissipated is practically negligible. This absence of power is represented in Fig. 379 by a brown line on the abscissa.

When the transistor is saturated the power dissipated in the collector–base junction is a maximum and equal to $P_{C\,max}$. If t_1 represents the time during which this transistor remains saturated, the corresponding power remains constant.

The power disappears when the transistor becomes cut-off again.

Each triggering pulse is associated with a power pulse dissipated in the transistor.

We will now consider the case of a short pulse, e.g. of duration t_1 considerably shorter than that of the complete cycle of the signal. The instantaneous values of the collector power are then as shown in Fig. 379.

Instantaneous values of the temperature of the junction. The temperature of the collector–base junction of the transistor varies with the same periodicity as the variations of the power dissipated in the transistor.

If we refer back to the thermal equivalent circuit of the transistor, we see the reason for the marking of the heatsink temperature (T_h) on the ordinate of Fig. 380. This temperature is higher than the ambient temperature.

The collector–base junction of the transistor is separated from the heatsink by a certain number of components which offer a resistance to the passage of the heat energy. If the spacing of the pulses does not allow the removal of all the energy stored in the transistor, the temperature of the collector–base junction will be higher than the temperature of the heatsink ($T_{j\,min}$ in Fig. 380).

When the power pulse appears in the transistor the junction temperature increases to the maximum value $T_{j\,max}$. When this pulse disappears, the junction temperature falls to the value $T_{j\,min}$. The shapes of the rise and fall of the temperature are functions of the thermal time constants of the circuit. The thermal time constant is shorter during the period of the rise in temperature, due apparently to the very rapid increase in the temperature.

Pulses of long duration with respect to their separation

Figure 381 shows the variations of the power dissipated in the collector–base junction of the transistor, which is driven by a generator of rectangular signals $\frac{1}{2}T$. During the period of cut-off the power is zero, and during saturation it is a maximum and equal to $P_{C1\,max}$.

The variations of the junction temperature can be shown on a graph in which the ordinate is graduated in terms of the following temperatures:

(a) the ambient temperature T_a;

(b) the heatsink temperature T_h;

(c) the minimum junction temperature $T_{j\,min}$;

(d) the maximum junction temperature $T_{j\,max}$.

These different temperature levels are plotted on the ordinate of Fig. 382.

The temperature of the collector–base junction is always higher than that of the heatsink, but it is stabilized at a very definite value $T_{j\,min}$. Any increase in the power dissipated in the transistor causes a temperature rise to a level $T_{j\,max}$. In the absence of power dissipated in the transistor the junction temperature decreases.

It will be seen at once from Fig. 382 that the temperature of the junction varies periodically with the variations of the power dissipated in the collector of the transistor.

Maximum operating characteristics. A similar reasoning applies to the definition of the maximum values of the power and temperature of the junction.

In Fig. 383 the same variations of the collector power produce variations in the temperature of the junction of the same shape as those mentioned above. Figure 384 shows the temperature variations of the junction from one value to another, and finally becoming stable at:

(a) a minimum temperature ($T_{j\,min}$) in the absence of a pulse;

(b) maximum temperature ($T_{j\,max}$) in the presence of pulses.

Given the limiting characteristics taken as reference, a maximum value of the power dissipated in the transistor corresponds to a junction temperature which should be the maximum specified by the manufacturer.

For germanium transistors this temperature is 75 °C, and for silicon transistors 150 °C.

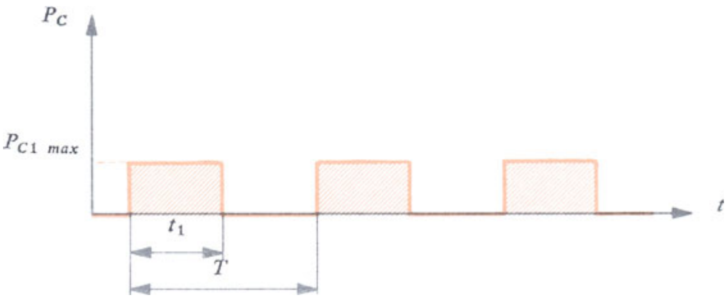


Fig. 381

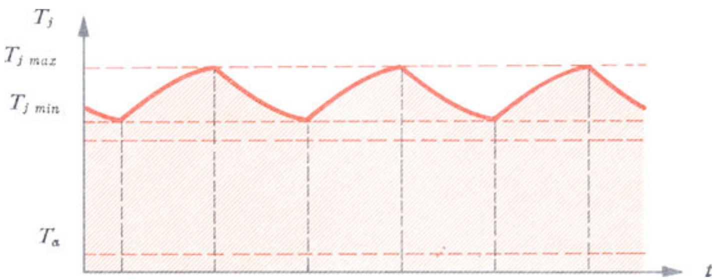


Fig. 382

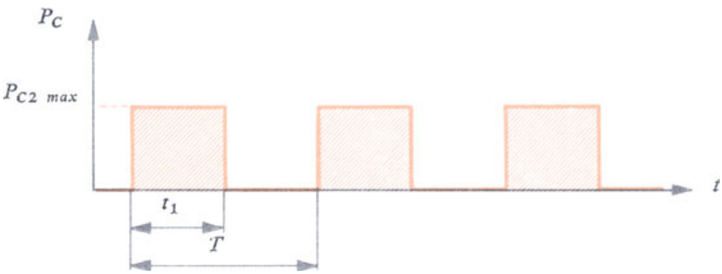


Fig. 383

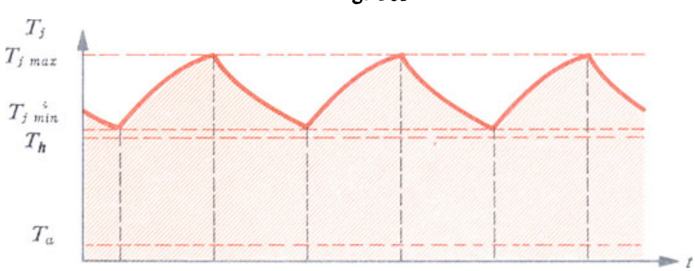


Fig. 384

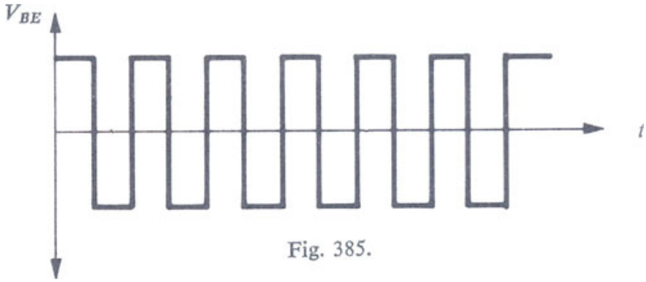


Fig. 385.

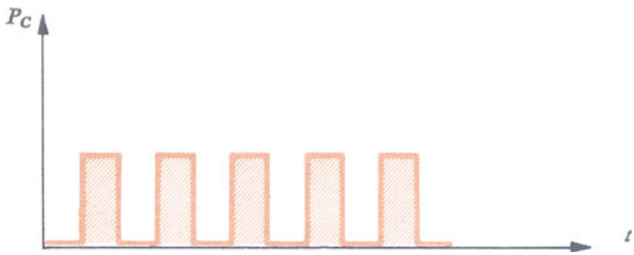


Fig. 386

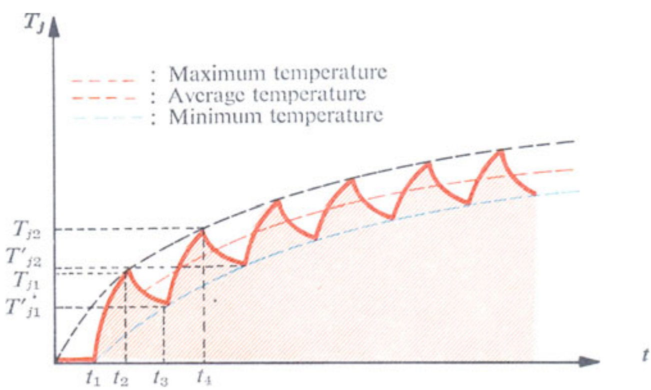


Fig. 387

Time for the establishment of the steady state

In the above study we have assumed a completely established state and values of junction temperature located between very precise limits. What happens when the signal is applied to the input terminals of the transistor? This transient period may be one in which the transistor may operate in the absence of thermal stability.

In Fig. 385 instant t_1 marks the application of a first voltage pulse between the base and emitter of the transistor, which becomes saturated, and the power dissipated reaches a maximum until the disappearance of the pulse.

At instant t_1 the power dissipated in the collector–base junction is practically zero, and the junction temperature is then near to the ambient temperature (Fig. 387).

The dissipation of a certain amount of power in the transistor causes a rapid rise in temperature, which reaches a level T_{j1} at instant t_2 in the figure.

At instant t_2 the pulse disappears and the junction temperature decreases (T_{j1}').

At instant t_3 a second pulse appears and the junction temperature starts to rise (T_{j2} in Fig. 387).

The pulse disappears at instant t_4 and the temperature falls to the value T_{j2}'.

As the pulses continue to appear both the maximum and minimum junction temperatures rise steadily:

$$P_{C\,max} = \frac{T_{j\,max} - T_{j\,min}}{R_{th(t)} + dR_{th(h-a)}}.$$

By means of this figure we can plot a curve representing the mean values of the junction temperature. The red curve in broken line in Fig. 387 shows this quite clearly. This curve is enclosed by the envelope representing the variations of the maximum junction temperature (in black in the figure) and the line connecting the various minimum values of this temperature (in blue in the figure).

This study of the time required for the establishment of the equilibrium value of the junction temperature is of great interest not only for the calculation of the initial periods, but also for the quantitative study of the thermal behaviour of a transistor in the absence of stabilization.

Diagrammatic method for the determination of the maximum temperature of the junction

The rectangular variations of the power dissipated in the collector–base junction of a transistor (Fig. 389) may be represented by successive variations in one direction or the other; these variations are superimposed and are in the positive direction for the appearance of each pulse, and in the negative direction when it disappears. They are dependent on the variations of the input voltage shown in Fig. 388.

At instant t_1 the presence of a power pulse is shown by a variation in the positive direction in Fig. 390.

The pulse disappears at instant t_2. This variation of the power dissipated in the other direction explains the representation of a pulse in the negative direction and of amplitude equal to that of the first pulse.

A second pulse at instant t_3 in Fig. 389 produces another jump upwards from the level corresponding to the first pulse.

The disappearance of this second pulse produces a variation in the other direction than that in the previous conditions, that is to the original level associated with the negative value of the disappearance of the first pulse.

We could easily continue this study by the superimposition in one direction or the other of the successive variations due to the arrival and disappearance of pulses constituting the information.

From the second figure we can obtain information for the plotting of the temperature variations associated with each pulse.

For this purpose the reader is referred to the previous figures. When the first pulse is applied (at instant t_1) the resulting variation of power which it produces in the collector–base junction of the transistor causes an increase in the temperature of this junction in accordance with the red curve in Fig. 391.

At instant t_2 the pulse disappears, and the corresponding sudden negative increase in Fig. 390 causes a decrease in the temperature of the junction. This variation takes place in the time shown in red dotted lines in the curve.

The arrival of the second pulse at instant t_3 causes a new rise in the temperature, as shown in the blue full-line curve in Fig. 391.

At instant t_4 the pulse disappears and the temperature of the junction decreases as shown in the blue dotted line curve.

The same applies to the succeeding pulses, and the figure showing the variations corresponding to the junction temperature consists of a number of more or less parallel curves.

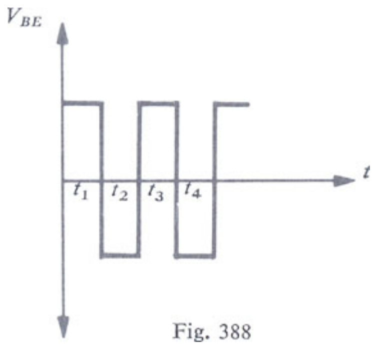


Fig. 388

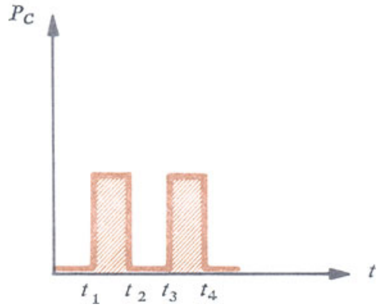


Fig. 389

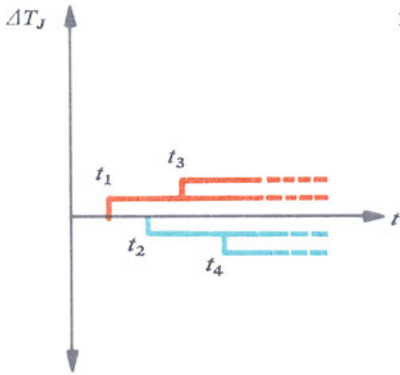


Fig. 390

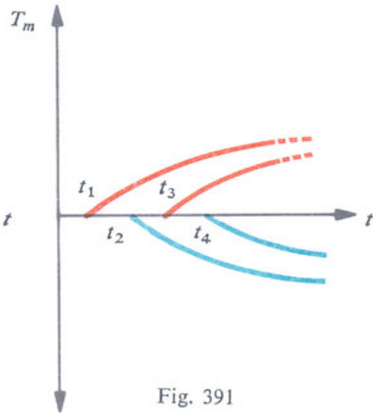


Fig. 391

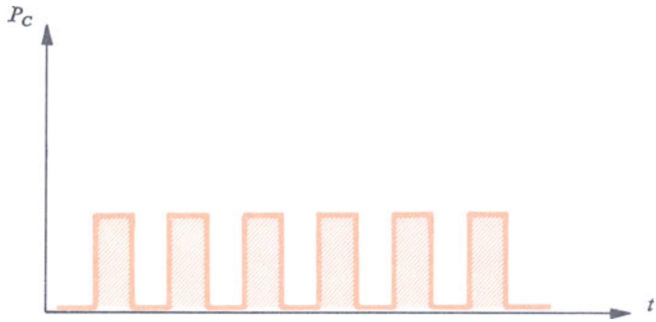


Fig. 392

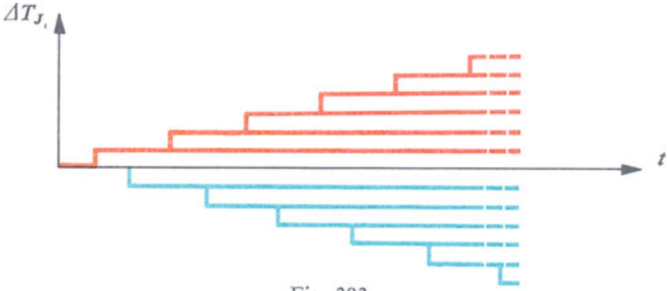


Fig. 393

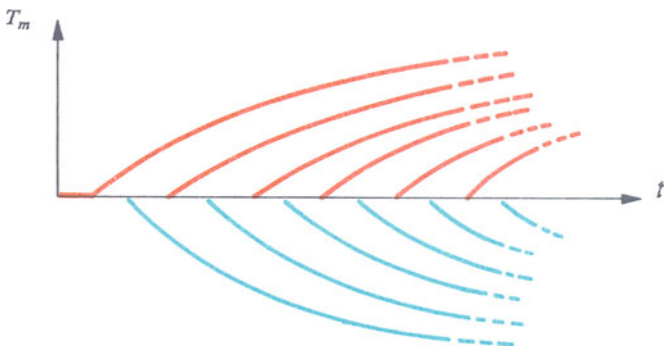


Fig. 394

Any increase in the power dissipated in the collector–base junction of a transistor may be represented in the figure by a tendency for an increase in the junction temperature, that is a curve in the positive direction. On the other hand a decrease in the collector power causes a variation in the other direction of the junction temperature, giving a curve in the negative direction.

The maximum junction temperature is a function of the number of positive variations of the power, as well as the negative variations.

In practice, the maximum temperature of the junction is dependent on the algebraic sum of the different junction temperatures associated with each pulse; this temperature is, therefore, dependent on the number of positive pulses and on the same number of negative pulses.

This diagrammatic method is very often used (Figs. 392, 393, 394); it has the advantage that, althought not being very accurate, it is easy to use. This theory is made use of especially for the preparation of certain curves representing the switching specifications for a transistor.

We emphasized earlier the importance of the concept of the transient thermal impedance. These parameters are still not often supplied by the transistor manufacturers, but they can be determined quantitatively:

(a) either with the aid of circuits specially adapted for this type of measurements;
(b) or by calculation from the theory of positive and negative variations just discussed.

5. Switching Circuits and Multivibrators

5.1. Principle of switching

The switching function constitutes one of the most important fields for the use of semiconductor devices. It consists of the transformation of d.c. energy (voltage or current) into energy of transient nature (rectangular, triangular signals, etc.)

Before defining the main types of circuits associated with the use of semiconductor devices for switching, we must remind ourselves of the principle of operation of a simple switch in order to be able to compare it with the behaviour of a diode or transistor.

Production of a rectangular signal

If we examine the circuit in Fig. 395 we see that it consists of a switch S as well as the battery V and resistor R. The switch may occupy successively the two positions: (a) O, open; (b) C, closed.

The representation of the variations with time of the current and voltage in this circuit involves the use of the graphs in Fig. 396. The top section illustrates the current-voltage characteristic of the circuit, and the bottom the variations with time of the variations of the voltage at the terminals of the switch.

Switch in open position 'O'. From instant t_1 to the instant preceding t_2 the switch S is open. No current flows through the circuit and the whole voltage supplied by the battery appears at the terminals of the switch (point V on the abscissa in Fig. 396).

Switch in closed position 'C'. From instant t_2 to instant t_3 the switch is closed. The current is a maximum; the whole voltage appears at the terminals of the resistance R, which corresponds to a zero voltage at the terminals of the switch (point O on the abscissa in Fig. 396).

The voltage at the terminals of the switch is therefore directly dependent on its position:

(a) in the open position it is a maximum at V (from instant t_1 to instant t_2);

(b) in the closed position it is zero (from instant t_2 to instant t_3).

The operation of the switch from one position to the other causes a rectangular voltage to appear at its terminals.

This illustrates the principle underlying the generation of rectangular signals, but it is still necessary to emphasize the restricted meaning of the term 'generation', which refers here much more to the transformation of energy than to the effective production of rectangular signals. This consideration will be of special importance when we are dealing with the calculation of the efficiency of a generator of rectangular signals.

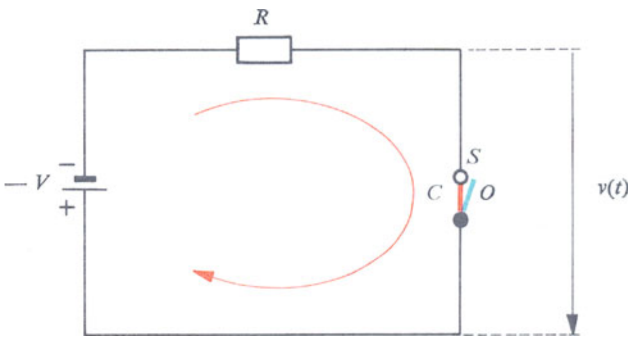


Fig. 395

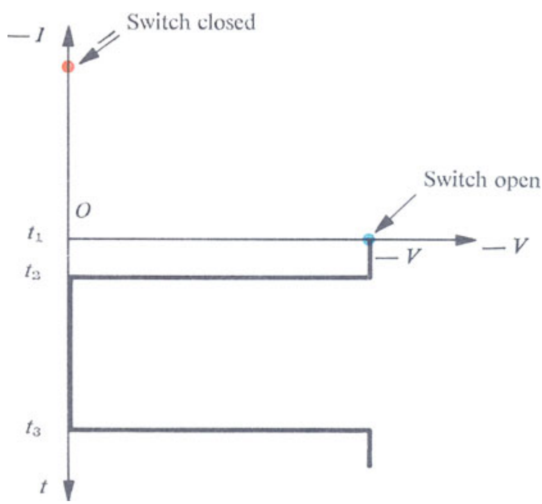


Fig. 396

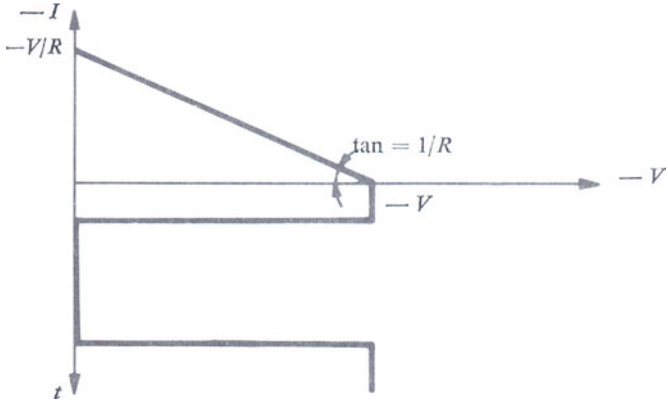
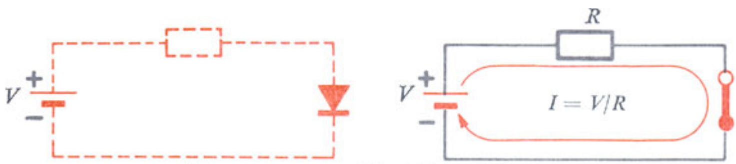
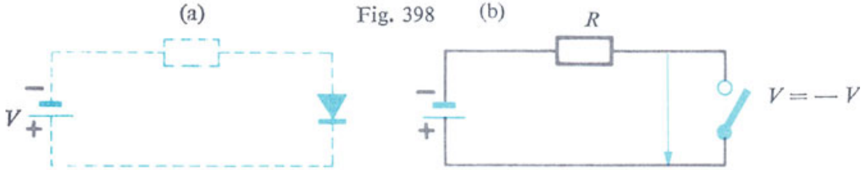


Fig. 397



(a) Fig. 398 (b)



(a) Fig. 399 (b)

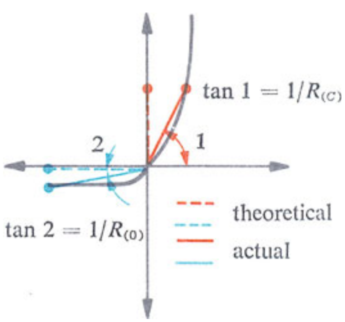


Fig. 400

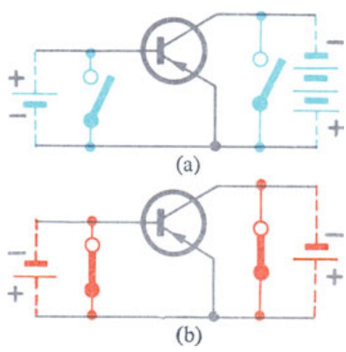


Fig. 401

Diagrammatic determination of the influence of the resistance R

In the top section of Fig. 397 we see that the two electrical parameters (voltage and current) may assume the following limiting values:

(a) 0 and $-V$ for the voltage;

(b) 0 and $-V/R$ for the current.

If we join the points defining the maximum value of the current and the limiting value of the voltage we obtain a straight line which represents diagrammatically the presence of the resistance R in the circuit. The tangent of the angle V is practically equal to the inverse of this resistance:

$$\tan V = \frac{V/R}{V} = \frac{1}{R}.$$

Resemblance of a diode to a switch

The classical mechanical switch may with advantage be replaced by a diode or a transistor.

We will confine our study at this stage to the purely static behaviour of the diode. When it is biased in the forward direction (Fig. 398a) its resistance is practically negligible, and it can therefore be represented electrically by a switch in the closed position (Fig. 398a). On the other hand, a reverse bias (Fig. 399a) corresponds to a very high value of the resistance of the diode, that is to a switch in the open position (Fig. 399b).

These two resistances may be easily calculated by means of the network of forward and reverse characteristics of the diode. The quantitative determination of the two resistances is performed by the method shown in Fig. 400.

Resemblance of a transistor to two switches

By definition the transistor consists of two junctions:

(a) the emitter–base junction;

(b) the collector–base junction.

In the chapter entitled 'transistors in switching' the necessity to switch the transistor from cut-off to saturation and from saturation to cut-off was pointed out. In the cut-off condition the two junctions are biased in the reverse direction; the two switches which represent them electrically are therefore in the closed position (Fig. 401a). In the saturated condition, the two junctions are biased in the forward direction, which is equivalent to the switches being closed (Fig. 401b).

Influence of a contact resistance

The switch shown in Fig. 402 is assumed to be an ideal one in which:

(a) the resistance is zero when it is closed;

(b) the resistance is infinite when it is open.

The variations of the voltage received at the terminals of the output of this circuit are shown in Fig. 403. The signal is perfectly rectangular and has the following limiting values;

(a) zero voltage in the short-circuited position;

(b) a voltage equal to $-V$ in the open position.

In electronic switching circuits it is difficult to compare the semiconductor devices used with switches with identical characteristics; in practice, in addition to the effect of inertia, they always offer some resistance to the flow of current. This resistance is low; it is represented in the circuit in Fig. 404 by a low resistance (r_s) in series with the switch.

In the open position the current in the circuit is zero, and the voltage across the switch is a maximum and equal to $-V$.

In the closed position the current in the circuit is a maximum, although it is lower than the value in the previous example, namely:

$$I = V/(R+r_s).$$

A diagrammatic solution will easily be found with the aid of the network of characteristics in Fig. 405.

The straight line connecting the points:

(a) abscissa 0, ordinate $I = V/R$ (point A);

(b) abscissa $-V$, ordinate 0 (point B); represents diagrammatically the influence of the resistance R.

We can now plot another line from the origin making an angle with the abscissa such that its tangent is equal to the inverse of the series resistance r_s, that is:

$$\tan = 1/r_s.$$

The two lines intersect at the point C, and the projection of this point on the ordinate marks the limiting value of the current when the switch is closed [$I = V/(R+r_s)$]; the projection of the same point on to the abscissa marks the value of the voltage when the switch is open (point C').

The voltage available across the output of a circuit of this nature has the shape shown in Fig. 405. The contact resistance r_s affects only the amplitude of the signal.

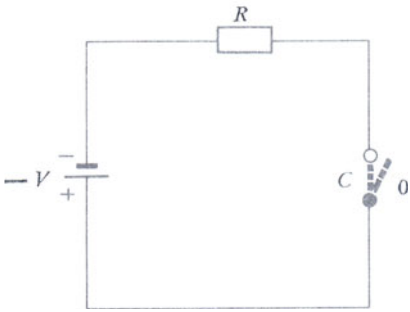


Fig. 402

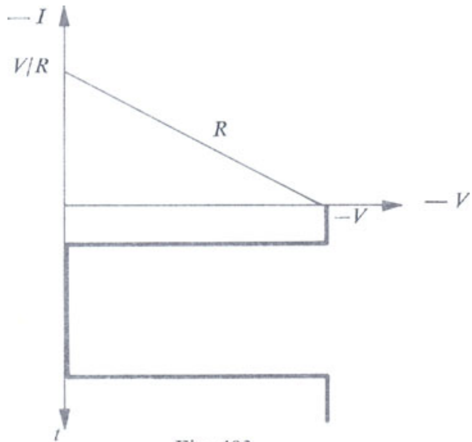


Fig. 403

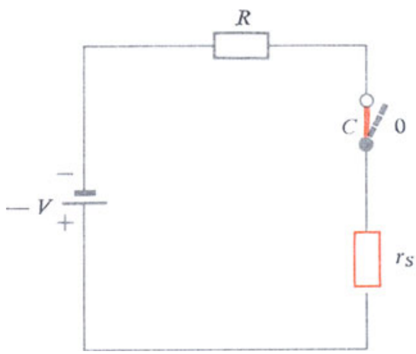


Fig. 404

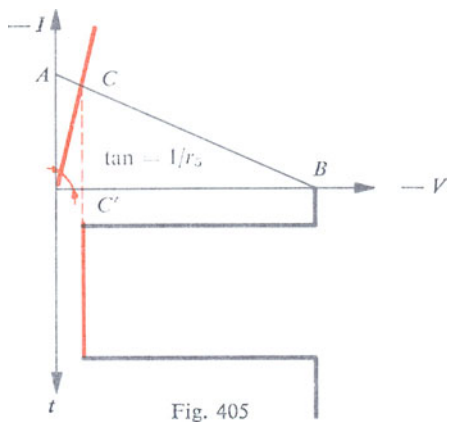


Fig. 405

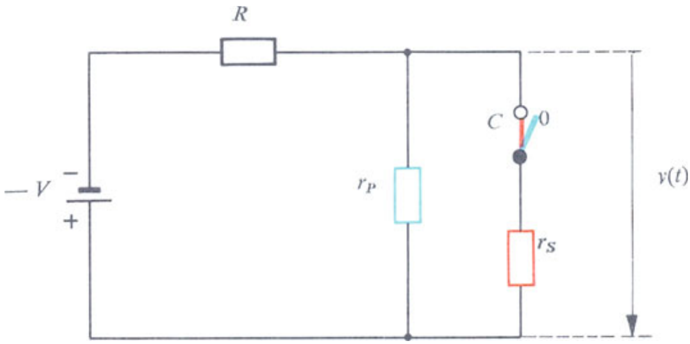


Fig. 406

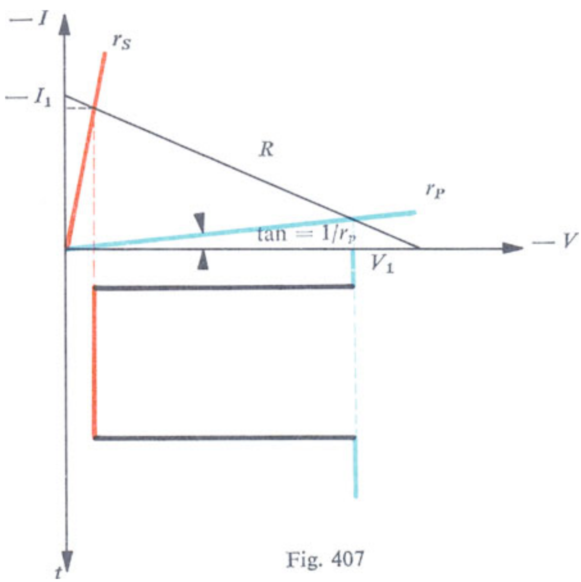


Fig. 407

Influence of a leakage resistance

Although the switch often presents a slight resistance to the flow of current when it is closed, it allows a small current to flow when it is open. This effect is represented electrically by a resistance r_p in parallel with the switch (Fig. 406).

The value of this resistance is high by comparison with the series resistance r_s.

When the switch is closed the voltage across the terminals of the switch is equal to:

$$V_C = V \cdot [r_s/(R+r_s)].$$

In the open position the voltage is no longer equal to V because of the voltage drop caused by the small current through resistance R, that is:

$$V_0 = V \cdot [r_p/(R+r_p)].$$

In this case we can immediately obtain the variations of the voltage at the terminals of the switch diagrammatically. In Fig. 407 the black line still represents the influence of the resistance R, the red line the series resistance r_s and the blue line the parallel resistance r_p.

This parallel resistance can be plotted in accordance with the method discussed in the previous chapter. All that is needed is to draw a line through the origin and making with the horizontal (the abscissa) an angle such that its tangent is equal to the inverse of the resistance:

$$\tan = 1/r_p.$$

The point of intersection of this line (in blue in the figure) with the black line gives by projection:

(a) on the ordinate, the value of the current through the circuit with the switch open;

(b) on the abscissa, the voltage at the terminals of the switch still open.

The voltage at the terminals of a circuit of this nature is as shown in blue in Fig. 407. The leakage resistance r_p affects the maximum value of the signal without however modifying its shape.

Effect of inertia

With slow switching it can be assumed that the transition from one state to the other occurs instantaneously; in these conditions the voltage variations are instantaneous and the signal is perfectly rectangular.

If the speed of switching is increased, the time necessary for the change from one position to the other becomes less and less negligible, thus producing a progressive variation of voltage during the transient periods.

This effect is explained electrically by the connection in parallel of a capacitor with the switch (Fig. 408).

When the switch is closed the voltage is a minimum and equal to V_A' in Fig. 409.

The transition from the closed to the open position produces at the terminals of the switch, or of the capacitor, a voltage variation equal to:

$$V_B' - V_A'.$$

The capacitor is charged through the resistance R. If the time constant (CR) is short, the voltage rapidly reaches its final value after the switch has been operated (in brown in the figure).

The change from the open to the closed position now causes a voltage variation in the other direction; the capacitor discharges through a combination of resistances in parallel (R and r_p). If we again assume a relatively short time constant (since the resistance r_p is much larger than the resistance R), the signal changes from the maximum to the minimum value in the time shown on the abscissa in Fig. 409.

A higher resistance R would cause a lengthening of the time constant of the circuit, and this would distort the signal, for example as shown in brown in Fig. 409.

5.2. Transistor in common emitter. Switching function

At the beginning of this chapter we likened the transistor in common base to two switches in the open position during cut-off, and closed during saturation.

Transistors in the common base configuration are rarely met with in switching circuits. Can we represent this transistor in common emitter simply by means of a group of switches?

It can be seen from Fig. 410 that the following two divisions can be made:

(a) the input circuit;

(b) the output circuit.

Representation of the input circuit. This circuit consists only of the emitter-base junction of the transistor. In the non-conducting state it offers a high resistance to the flow of current (switch open, in blue in Fig. 410).

In the saturation state the base–emitter junction is strongly biased in the forward direction; its resistance is then very low and it can therefore be represented electrically by a closed switch (in red in the figure).

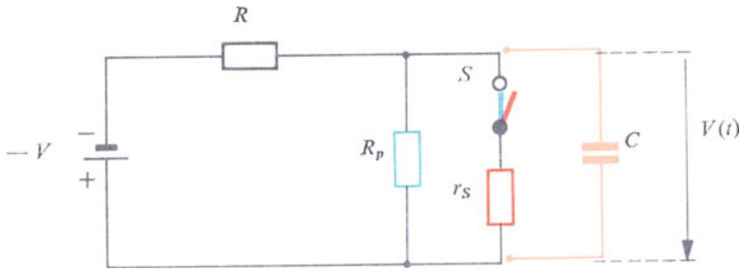


Fig. 408

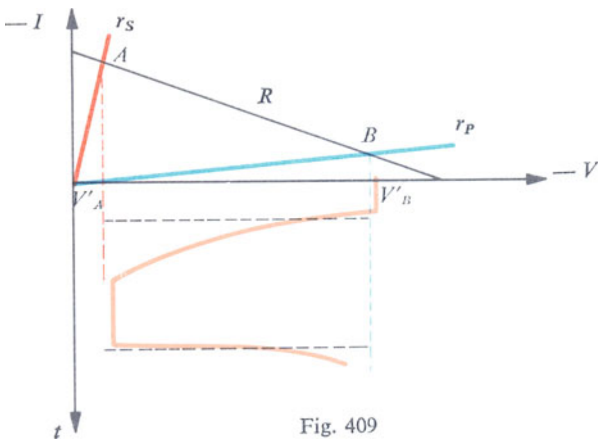


Fig. 409

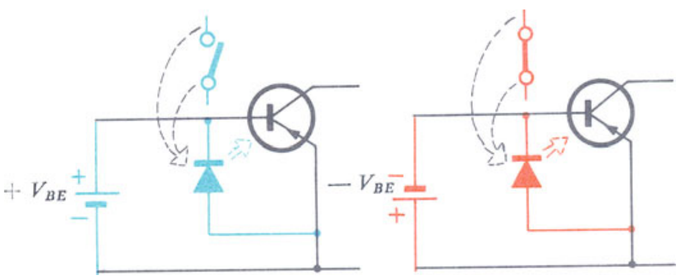


Fig. 410

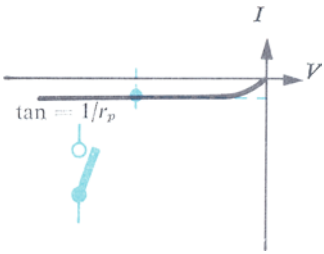


Fig. 411

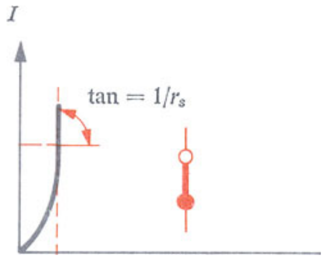
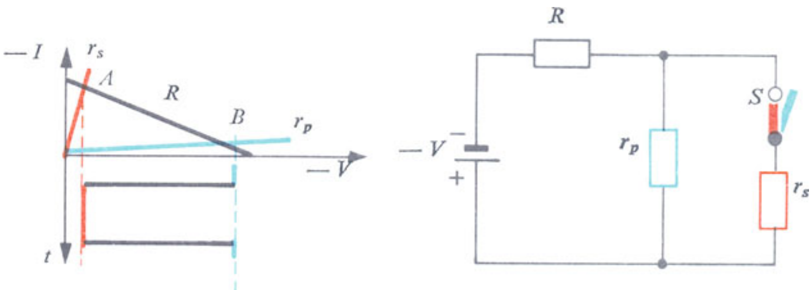


Fig. 412



(a) Fig. 413 (b)

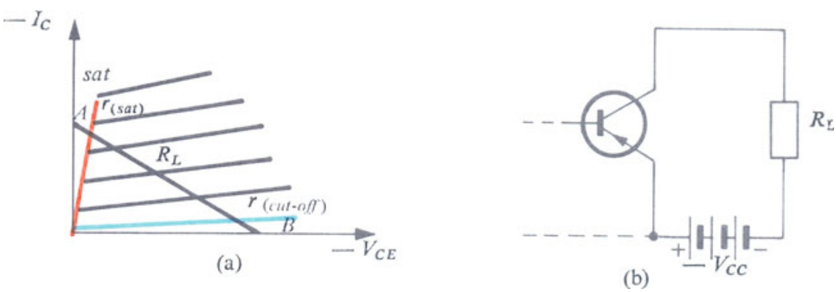


Fig. 414

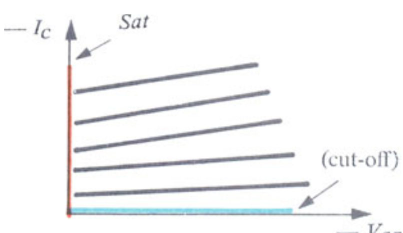


Fig. 415

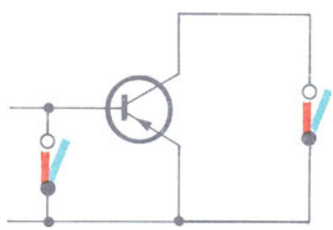


Fig. 416

We will omit the contact and leakage resistances which were included earlier. The leakage resistance r_p could be compared to the reverse resistance of the junction calculated in accordance with the method in Fig. 411. The contact resistance is associated with the forward resistance of this junction determined quantitatively by the methods given in Fig. 412.

Representation of the output circuit. It would seem more difficult to construct a simplified electrical equivalent circuit for the output circuit of the transistor, which consists in fact of the collector–base and base–emitter junctions.

Referring again to the drawing in Fig. 413a, we see that the various components shown enable the effects of the various resistance R, r_s and r_p to be determined (circuit in Fig. 413b).

It will be seen from the network of output characteristics of a transistor that the variations of the collector current are shown on the ordinate and those of the collector–emitter voltage on the abscissa (Fig. 414a). The transistor is loaded with a resistance R associated in the diagram with a load line shown in black in the figure.

In the cut-off state the two junctions are biased in the reverse direction and the working point is located at B (point of intersection of the load line with the residual characteristic of the cut–off).

During the period of saturation the working point is at A in the figure, that is at the point of intersection of the residual characteristics of the saturation and the load line. The corresponding circuit is shown in Fig. 414b.

An examination of Figs. 413a and 414a will show that the behaviour of the output circuit of a transistor, when it is switched from cut–off to saturation and then back to cut-off, can be likened directly to the operation of a simple switch changing from the open to the closed position and back to the initial open position.

In order to simplify the study of the principles underlying the operation of multivibrators, for example, no account will be taken of the residual characteristics; the output circuit will then be represented by a simple switch without contact resistance or leakage resistance. The corresponding network of output characteristics will then be an ideal one as shown in Fig. 415.

As for the circuit in common base, it is now easy to replace a transistor in common emitter by a system consisting of two switches (Fig. 416).

These two switches will be closed during the period of conduction (state of saturation) and open during the period of non-conduction (state of cut-off).

Rapid switching

The effect of inertia represented by the connecting of a capacitor in parallel with the switch (Fig. 417) makes itself felt when the transistor is switched rapidly from one state to the other.

In order to define the exact shape of the collector current by means of a group of electrical circuits, it is necessary to replace the switches of the previous page by the components of the overall equivalent circuit of the transistor acting as a switch (see Fig. 418). All these components have very variable values in terms of the real position of the working point; it then becomes very difficult in these conditions to calculate the real variation of the currents and voltages in the input and output circuits by a method of this nature.

The simplest method in this case is to retain for the transistor the simile of an association of two perfect switches, and, in order to obtain maximum accuracy in the calculation of the transient effects, to include the triggering and delay times, which are always supplied by the manufacturer.

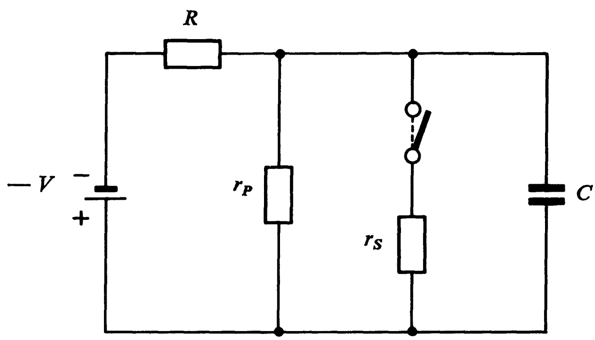


Fig. 417

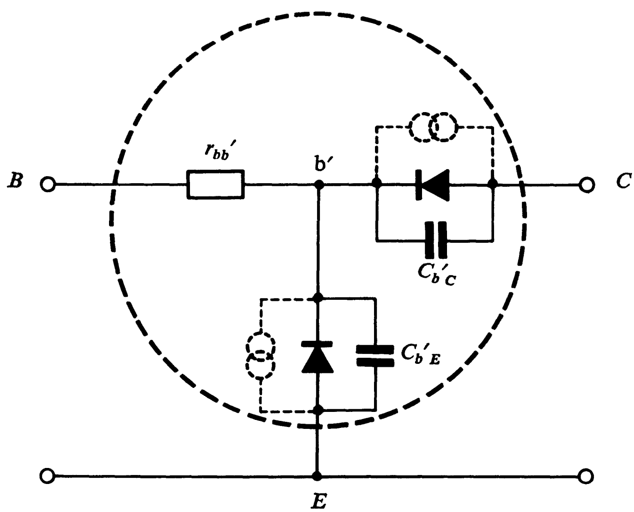


Fig. 418

6. Principal Switching Circuits

6.1. Chief functions of switching circuits

There are a number of circuits acting as 'switching' circuits, but we will of course not give the complete list here. It would be much more useful to study in detail the three basic applications of electronic switching circuits, which are:

(a) signal generation;
(b) signal shaping;
(c) counting.

6.2. Generation of signals

This function consists in the obtaining of rectangular signals, for example, by means of a system of one or more transistors. We referred earlier to the danger of the use of the term 'generation', since it refers much more to the conversion of energy than to the production of the actual signals.

This function is obtained electrically by means of free relaxation systems such as the astable or unstable multivibrator.

The function of 'signal generation' represented in Fig. 419 is performed by means of the circuit in Fig. 420 which represents the basic diagram of an astable multivibrator.

This type of multivibrator consists of two transistors (T_1 and T_2) supplied by a single group of batteries ($-V_{CC}$) and biased by two separate batteries ($-V_{BB1}$ for transistor T_1 and V_{BB2} for transistor T_2). The necessary dynamic connections for the transmission of energy from the output of one transistor to the input of the other consist of two capacitors C_1 and C_2.

A circuit of this type switches automatically from one state to the other, namely:

(a) for the first state, T_1 saturated and T_2 cut-off;
(b) for the second state T_1 cut-off and T_2 saturated.

We see therefore that the transistion from one state to the other involves very rapid current variations in the output circuits of the transistor, making it possible that the variations of the output will approach the rectangular shape mentioned earlier.

6.3. Pulse shaping

The basic information in any switching system (computers, industrial circuits, etc.) consists of pulses, which have to pass through a certain number of electrical circuits each of which introduces distortion which may even amount to the almost complete suppression of the information (Fig. 421).

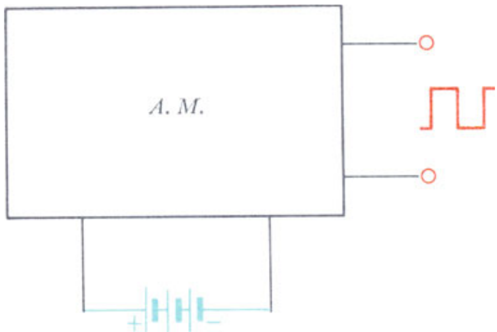


Fig. 419

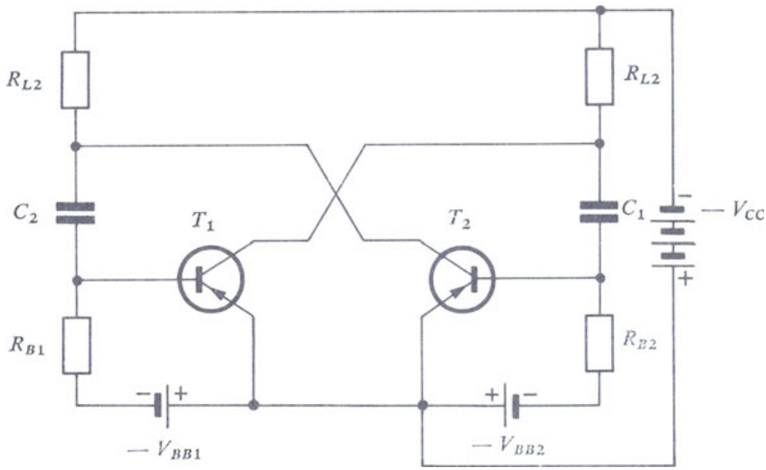


Fig. 420

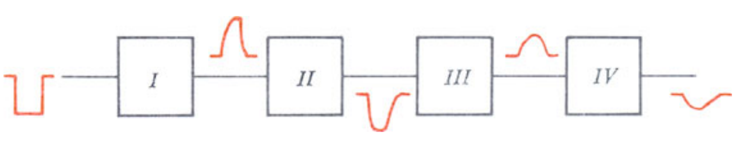


Fig. 421

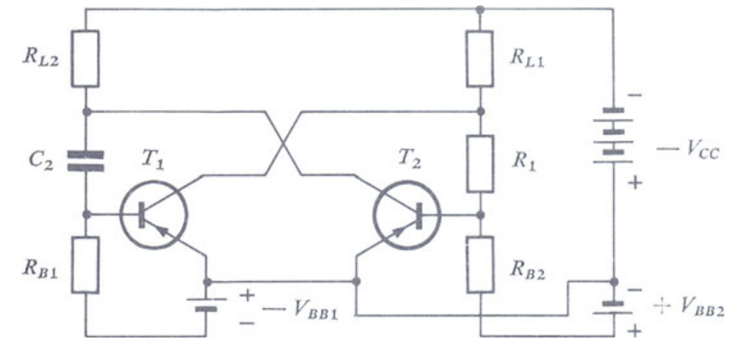


Fig. 422

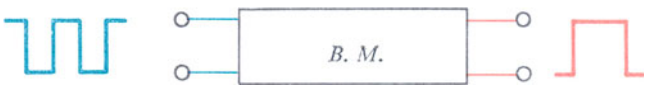


Fig. 423

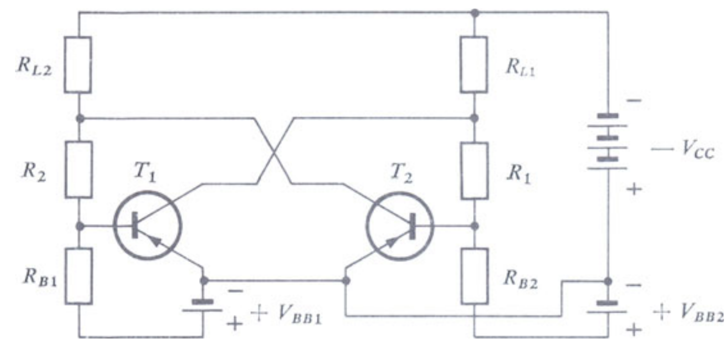


Fig. 424

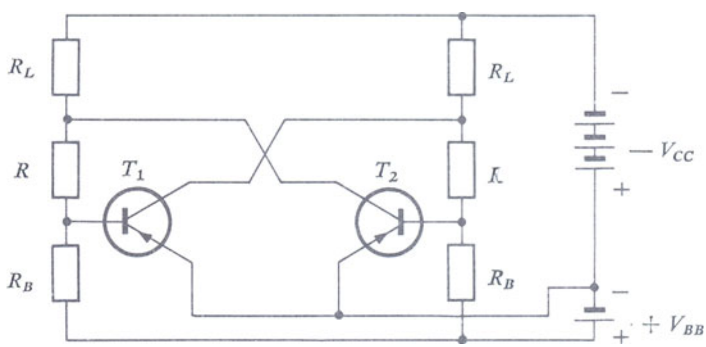


Fig. 425

It is, therefore, necessary to install, at different points in the switching chain, a certain number of circuits in order to restore the original shape to the pulse; this is the function of regeneration or signal shaping.

In practice the system able to perform this function should supply a single pulse for a signal applied across its input terminals; this is the function of the monostable multivibrator.

A multivibrator of this type is shown in Fig. 422. It consists as before of two transistors (T_1 and T_2) supplied by the same group of batteries ($-V_{CC}$); one transistor is biased by a battery $-V_{BB1}$ and the other by a battery $+V_{BB2}$. The difference between the couplings is that one is capacitive (C_2) and the other is resistive (R_1).

A circuit of this type possesses one stable state for example T_1 at saturation and T_2 at cut-off. When information is transmitted to it, it triggers, resulting in a new operating state (T_1 cut-off and T_2 saturation). It then automatically returns to its original state and remains there until the receipt of a new information signal.

6.4. Counting

Computing is one of the most important functions in switching. The number of pulses applied at the terminals of a stage is divided by 2. We do not propose here to discuss logic technique, since there are numerous specialized works on the subject; on the other hand it might be useful to remind ourselves of the principle of frequency division.

The counting function is illustrated simply in the synoptic diagram in Fig. 423. For every two pulses transmitted to the input, a single pulse appears in the output circuit.

The bistable multivibrator is an example of the type of circuit capable of performing this function. It consists of two transistors (T_1 and T_2, Fig. 424) supplied by the same group of batteries $-V_{CC}$ and biased by two batteries ($+V_{BB1}$ for transistor T_1 and $+V_{BB2}$ for transistor T_2). The connections consist of resistors (R_1 and R_2). In practice bistable multivibrators are completely symmetrical, which involves two identical transistors as well as biasing batteries ($V_{BB1} = V_{BB2} = V_{BB}$) and resistors ($R_{L1} = R_{L2} = R_L$; $R_1 = R_2 = R$; $R_{B1} = R_{B2} = R_B$). We will be studying this type of circuit (Fig. 425) in the following pages.

A system of this sort possesses two stable states:

(a) T_1 saturation, T_2 cut-off for the first state;

(b) T_1 cut-off, T_2 saturation for the second state.

This circuit remains in one of these positions in the absence of any information signal. As soon as a signal is applied to its terminals its state changes until the arrival of another pulse; it then returns to its original state. For two pulses applied to its input terminals, a single one is received in the output circuit.

6.5. Similarity of the three types of multivibrator

In addition to the three very different functions which can be performed by the various circuits so far considered, we should take note at this stage of the chief modifications appearing in the electrical circuits of these systems.

The basic diagram of a multivibrator is shown in Fig. 426. The components common to the three circuits are the two transistors (T_1 and T_2), the supply battery ($-V_{CC}$), the load and bias resistances (R_{L1}, R_{L2} and R_{B1}, R_{B2}) respectively.

Astable multivibrator (in blue)

The two biasing voltages are negative ($-V_{BB1}$ and $-V_{BB2}$) and the two connecting systems are capacitative (C_1 and C_2).

Monostable multivibrator (in green)

One coupling is capacitative (C_2) and the other resistive (R_1). The bias battery, on the capacitative coupling side, is negative ($-V_{BB1}$) and the other positive ($+V_{BB2}$). It would be possible to reverse the polarities of the two bias batteries, but this is rarely done in practice.

Bistable multivibrator (in red)

The bias batteries both supply a positive voltage ($+V_{BB1}$ and $+V_{BB2}$). The coupling components consist of resistors (R_1 and R_2). It should be noted that in every case in which the couplings are resistive, a capacitor C is connected in parallel with the corresponding resistances. The role of this capacitor was defined in the section entitled 'Acceleration circuit'.

6.6. Other switching circuits

We have so far illustrated the chief functions met with in switching by means of multivibrators only. In practice there are many circuits that fulfil these functions, including, of course, multivibrators themselves.

For example, in addition to capacitative coupling between the collector and base of transistors, the transmission of energy may also be made by means of a non-decoupled emitter resistance (Fig. 427).

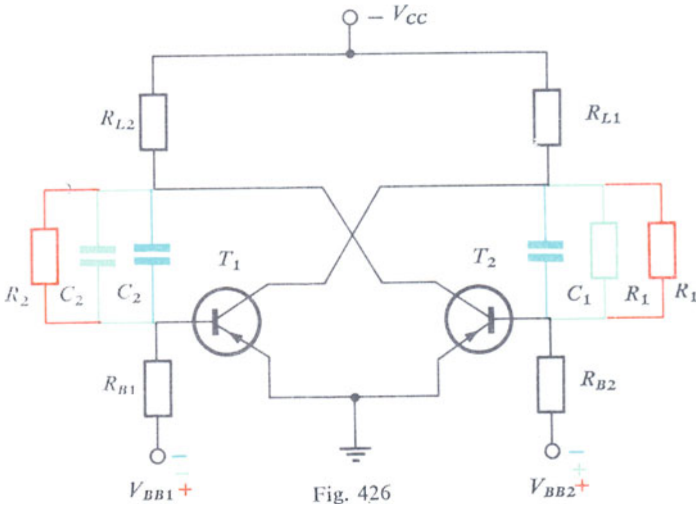


Fig. 426

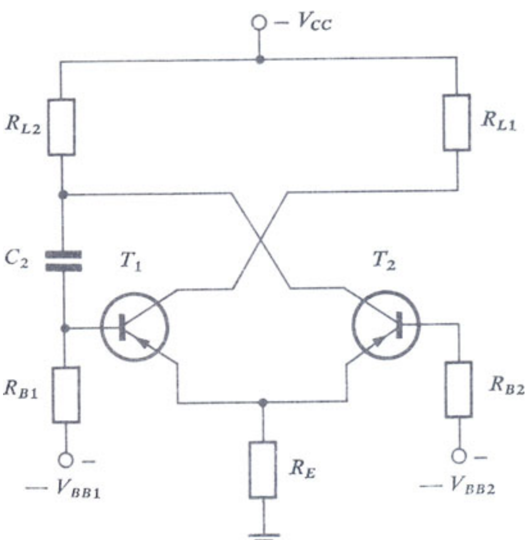


Fig. 427

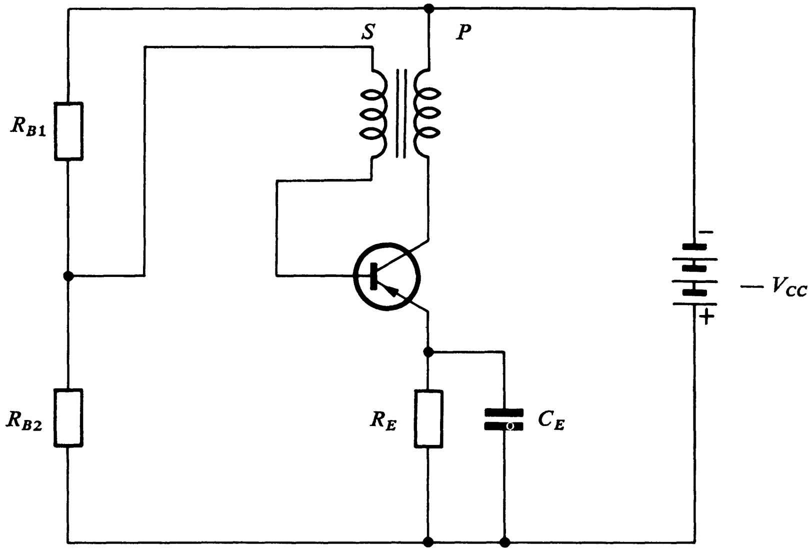


Fig. 428

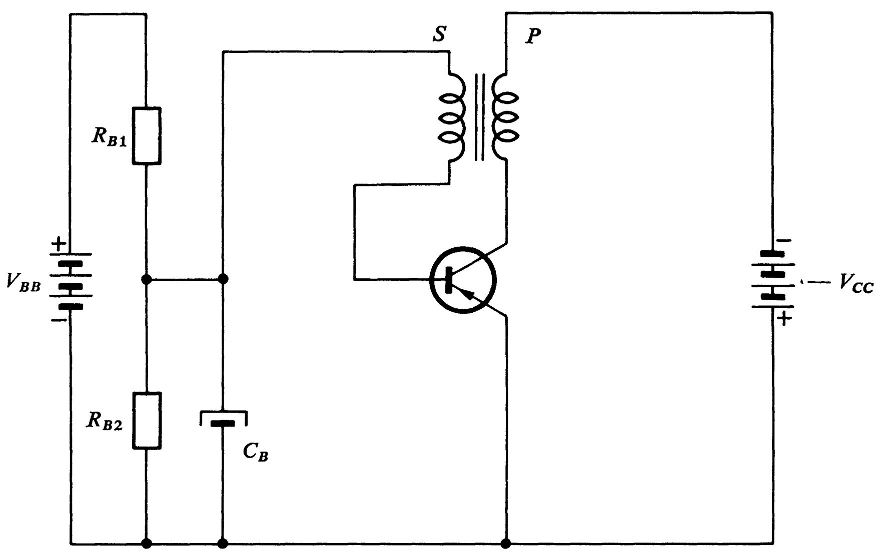


Fig. 429

In rapid or very rapid switching the systems consist of a much greater number of components so as to shorten the triggering time as much as possible.

There is however another large group of switching circuits; namely blocking oscillators which can perform the first two functions;

(a) signal generation;

(b) pulse shaping.

Signal generation. This function is performed by the astable or unstable blocking oscillator (Fig. 428), which consists of a single transistor associated with a transformer and an *RC* circuit. The role of this circuit is to allow triggering of the transistor from one state to another.

Pulse shaping. This is performed by the triggered blocking oscillator (Fig. 429). In this case a pulse at the input is associated with a pulse in the output circuit.

7. Astable Multivibrators

The role of the astable multivibrator is to supply energy in the form of rectangular signals from a d.c. source. The study of its operation can be divided into a number of sections:

(a) principle of the astable multivibrator;
(b) the physical explanation of the operation, with diagrams illustrating the signals at different points in the stage;
(c) calculation of the instantaneous values of the voltage and current with control of the conditions of cut-off and saturation;
(d) definition of the limiting values of voltage and current;
(e) possibility of shaping the signals received at the output;
(f) different types of circuit.

7.1. Theory and operating principle of an astable multivibrator

The basic diagram of an astable multivibrator was given earlier (Fig. 430). It consists of two transistors (T_1 and T_2) supplied by a group of batteries $-V_{CC}$ and biased by two batteries $-V_{BB1}$ and $-V_{BB2}$. Transistors T_1 and T_2 are loaded by resistors R_{L1} and R_{L2} respectively, whilst resistors R_{B1} and R_{B2} constitute the biasing systems, and C_1 and C_2 the coupling units.

In order to define the operation of a system of this nature all that is needed is to assume any given instant t corresponding, for example, to the following situation:

(a) T_1 saturated and the ratio between the collector current and the base current applying to the saturation condition: $-I_C/I_B < h_{FE}$;
(b) T_2 cut-off.

It is essential for the understanding of the transition from one state to another to determine the reasons why transistor T_2 is in the cut-off state at instant t. This transistor is of the *PNP* type and requires a positive voltage at its base for it to remain non-conducting (reverse bias of the two junctions causes the collector and emitter to be negative with respect to the base).

It will not be seen at once from the simplified diagram in Fig. 431 that the voltage at the base of the transistor T_2 may be positive. This voltage should apparently be negative since the collector of transistor T_1 is near to zero potential and the only voltage between the base and emitter of T_2 (between B_2 and earth) can be explained only by the charging of capacitor C_1 through resistor R_{B2}.

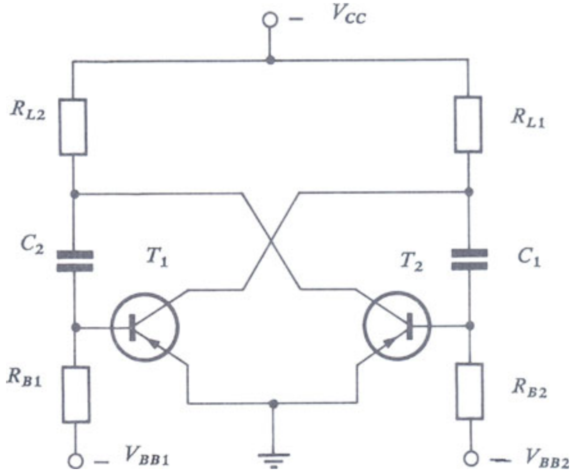


Fig. 430

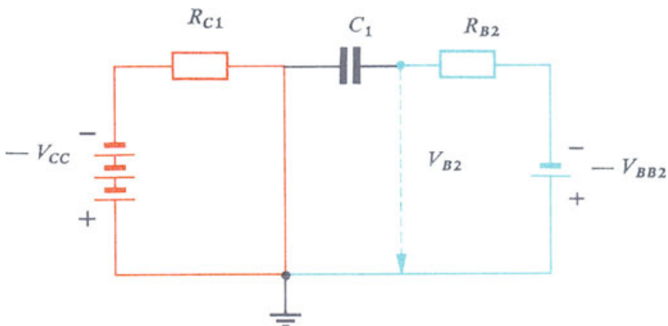


Fig. 431

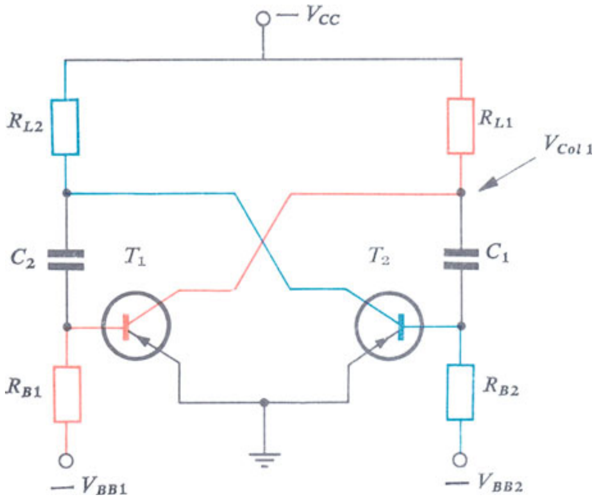


Fig. 432

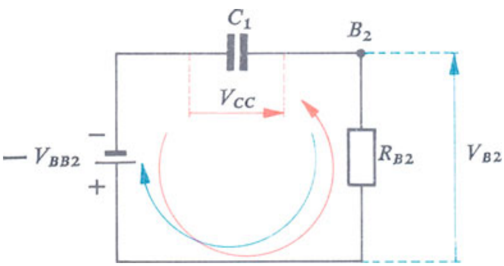


Fig. 433

Before reaching this state the multivibrator was in the opposing conditions:

(a) T_1 cut-off;

(b) T_2 saturation.

When the system has been triggered from this initial state to the final state the collector voltage of transistor T_1 changes suddenly from a maximum value equal to $-V_{CC}$, corresponding to the cut-off state, to 0 associated with the state of saturation. In Fig. 432 the voltage at point 'Col 1' has suddenly changed from $-V_{CC}$ to 0, that is a variation of the positive voltage equal to $+V_{CC}$.

This voltage appears instantaneously across the capacitance C_1, and the base of transistor T_2 suddenly becomes positive at this instant. Capacitor C_1 then discharges through resistor R_{B2} (Fig. 433) and recharges up to the value $-V_{BB}$ (voltage supplied by the corresponding battery).

In practice, the base of this transistor will never reach this voltage; as soon as it becomes definitely negative transistor T_2 becomes conducting and the multivibrator is triggered from one state to the other.

This explains the presence of a positive voltage at instant t at the base of the transistor when it is cut off, and also the possibility of this transistor changing from this operating condition, associated with the absence of current in the circuit, to approach the saturation state.

Triggering. Let us assume an instant t' corresponding to the appearance of a low negative voltage at the base of transistor T_2. This transistor will then start to conduct, and the voltage which was a maximum and equal to $-V_{CC}$ between its collector and earth, starts to decrease due to the voltage drop caused by this current in resistor R_{L2}. The corresponding point becomes less negative and this positive variation of the voltage is immediately applied between the base and emitter of transistor T_1.

The base of this transistor now becomes less negative and the corresponding collector current decreases; transistor T_1 is now no longer saturated. A negative voltage appears between the collector of this transistor and earth. This negative voltage variation is applied between the base and emitter of T_2 and accentuates its approach to conduction.

This process continues until transistor T_2 becomes saturated and transistor T_1 cut off.

It will be remembered that the condition of saturation requires for transistor T_2 a ratio between the collector and base currents lower than the current gain (in normal operation) of the transistor; the condition of cut-off implies for transistor T_1 the presence of a positive voltage between its base and emitter, that is between point B_1 and earth. This positive voltage is due to the strong voltage variation from $-V_{CC}$ to 0 between the collector and base of transistor T_2.

It should now be possible to limit the time during which this condition will last. All depends in fact on the initial duration of the cut-off state of transistor T_1, that is the discharge time of capacitor C_2 through resistor R_{B1}.

In a system of this nature the change of state is an exclusive function of the discharge of capacitors through resistors; the transistors must be prevented at all costs from passing from one state to another on their own. This does not seem possible when the transistor is cut-off, since this is caused by a positive voltage external to the transistor; on the other hand the slightest decrease in the collector current of the saturated transistor could cause uncontrolled triggering of the system.

7.2. Diagrammatic representation of the signals

We will be discussing below both the variations of the current and voltage appearing in the various 'meshes' of this stage. It is however interesting to represent, at the present stage of our study, the shapes of the various voltages which may be received at the 'active' electrodes of the two transistors.

An examination of the circuit in Fig. 434 will show:

(a) the dynamic neutrality of the emitters connected directly to earth;

(b) the possibility of defining the variations with time of voltage for the bases and collectors.

There is no real necessity to deal with these points in succession; we will therefore deal with them starting from the left and observing the conventional clockwise rotation.

Figure 435a gives the variations with time of the base–emitter voltage of transistor T_1. The limiting values of this voltage are located:

(a) positive, at the maximum value of the charging resistance of capacitor C_2 $(+V_{CC})$;

(b) negative, at the voltage supplied by the biasing battery $-V_{BB1}$.

Figure 435b shows the variations with time of the collector–emitter voltage of transistor T_2, which changes from a zero value during the period of conduction to a maximum value equal to $-V_{CC}$ during non-conduction.

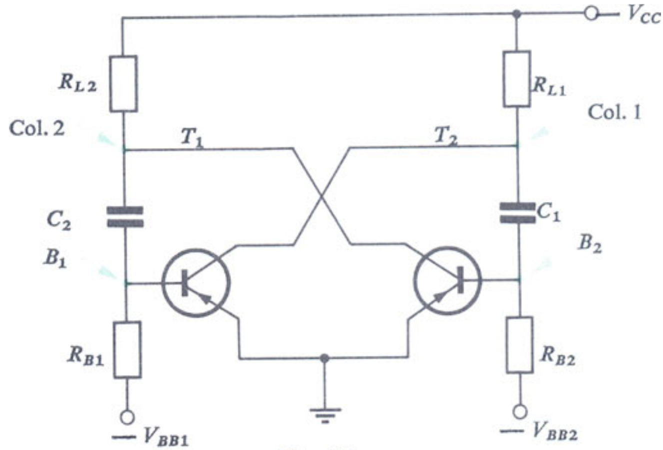
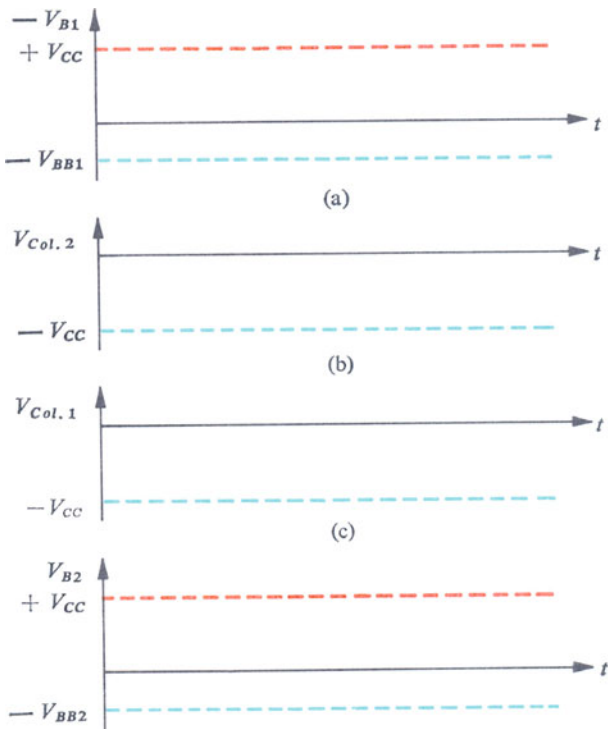
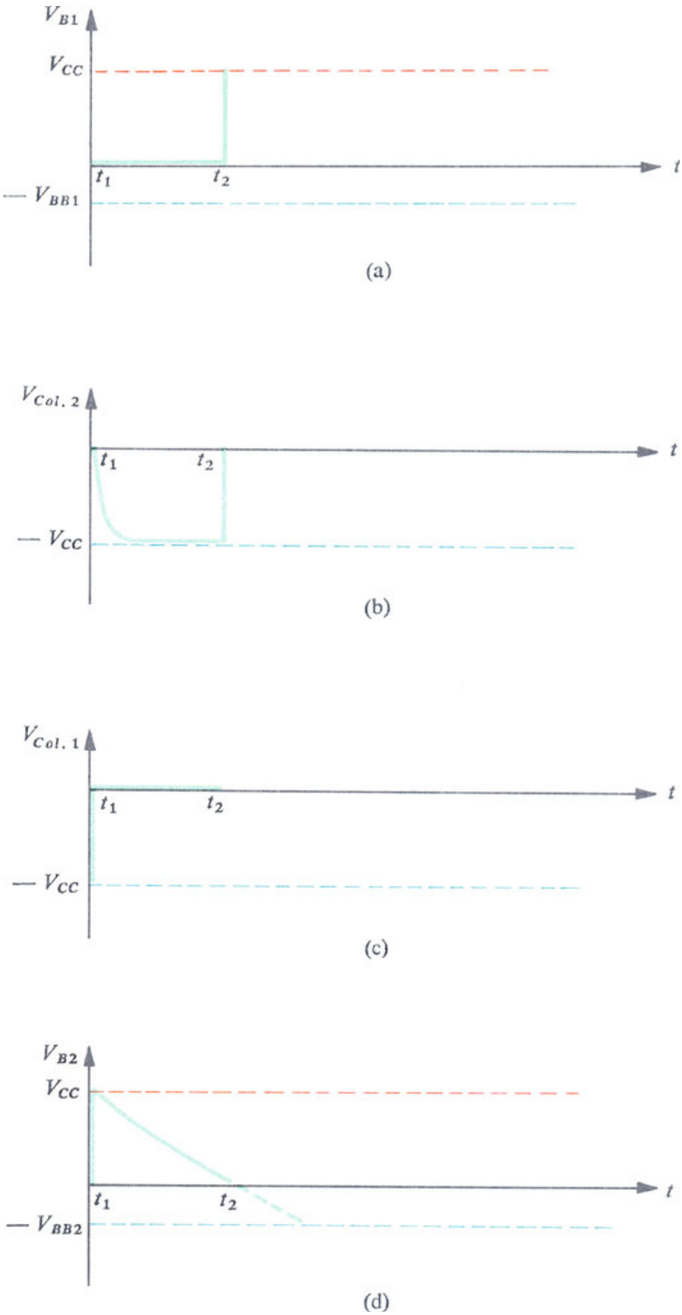


Fig. 434



(d)

Fig. 435



V_{B1}
V_{CC}
t_1
t_2
t
$-V_{BB1}$
(a)
$V_{Col.2}$
t_1
t_2
t
$-V_{CC}$
(b)
$V_{Col.1}$
t_1
t_2
t
$-V_{CC}$
(c)
V_{B2}
V_{CC}
t_1
t_2
t
$-V_{BB2}$
(d)

Fig. 436

Figure 435c shows the variations with limit of the collector-emitter voltage of transistor T_1. The extreme operating conditions of this transistor correspond to the values $-V_{CC}$ and 0 respectively.

Finally, Fig. 435d gives the variations with time of the base–emitter voltage of transistor T_2. The extreme variations of this voltage will be seen readily in Fig. 435a.

First pseudostable state

We will assume that at an instant t_1, at the origin of the coordinates in Figs 436a, b, c and d, transistor T_1 has just become saturated and transistor T_2 cut off. In these conditions the base–emitter voltage of transistor T_1 is slightly negative (near to zero in Fig. 436a) and the base–emitter voltage of transistor T_2 strongly positive and equal to $+V_{CC}$ (Fig. 436d).

Because of the saturation of transistor T_1 a large current immediately flows in its collector and there is practically no voltage between this electrode and earth (Fig. 436c). On the other hand since transistor T_2 is in the cut-off state there should normally be a maximum voltage equal to $-V_{CC}$ between the collector and earth. This is not so, however, for this voltage cannot increase except as a function of the charging of the capacitor C_2 through resistor R_{L2}. Consequently at instant t_1 the collector–emitter voltage of transistor T_2 is still zero (Fig. 436b).

Before passing on to the study of the variations with time of the various voltages, it seems necessary to define instant t_2 corresponding to the triggering of the multivibrator. The explanations given earlier showed that the change of state could only be explained by the appearance of a negative voltage between the base and emitter of the transistor in the cut-off state. The determination of instant t_2 therefore requires the study of the variations with time of the base–emitter voltage of transistor T_2.

This voltage changes from the value $+V_{CC}$ (at instant t_1) to zero (discharge of capacitor C_1 through R_{B2}) and then approaches the value $-V_{BB2}$, which however it never reaches because transistor T_2 starts to conduct as soon as the base–emitter voltage becomes slightly negative. Instant t_2 therefore corresponds to a zero value of this voltage (Fig. 436d).

From instant t_1 to instant t_2 transistor T_1 remains saturated, and in these conditions:

(a) the base–emitter voltage of this transistor remains near to zero (Fig. 436a);

(b) the collector–emitter voltage is also zero (Fig. 436c).

During this same period transistor T_2 is cut off. Its collector–emitter voltage changes gradually from zero to $+V_{CC}$. The variation of this voltage is dependent on the charging of capacitor C_2 through resistor R_{L2}. If the time constant of this circuit is short the capacitor charges very rapidly and the collector–emitter voltage almost immediately reaches the value of the supply voltage (Fig. 436b).

The various signals appearing at the different active points of the multivibrator from instant t_1 to instant t_2 are shown on the group of curves in Figs. 436a, b, c and d.

Second pseudostable state

From instant t_2 transistor T_2 is no longer cut off and changes to saturation, and the operation of transistor T_1 changes in the reverse direction, that is from saturation to cut-off.

At instant t_2 the triggering is explained by the appearance of a small negative voltage between the base and emitter of transistor T_2 (Fig. 437d). Because of the low value of the resistance offered by the input circuit of this transistor to the flow of current, we can liken this circuit to an ideal switch in the closed position; for this reason instant t_2 corresponds to the absence of voltage in this figure.

At the same instant the appearance of a very large current in the collector circuit of this transistor produces an immediate variation in the collector–emitter voltage from $-V_{CC}$ to zero (Fig. 437b).

The positive variation of the voltage is applied at once between the base and emitter of transistor T_1 (Fig. 437a), thus causing the transistor to be cut off.

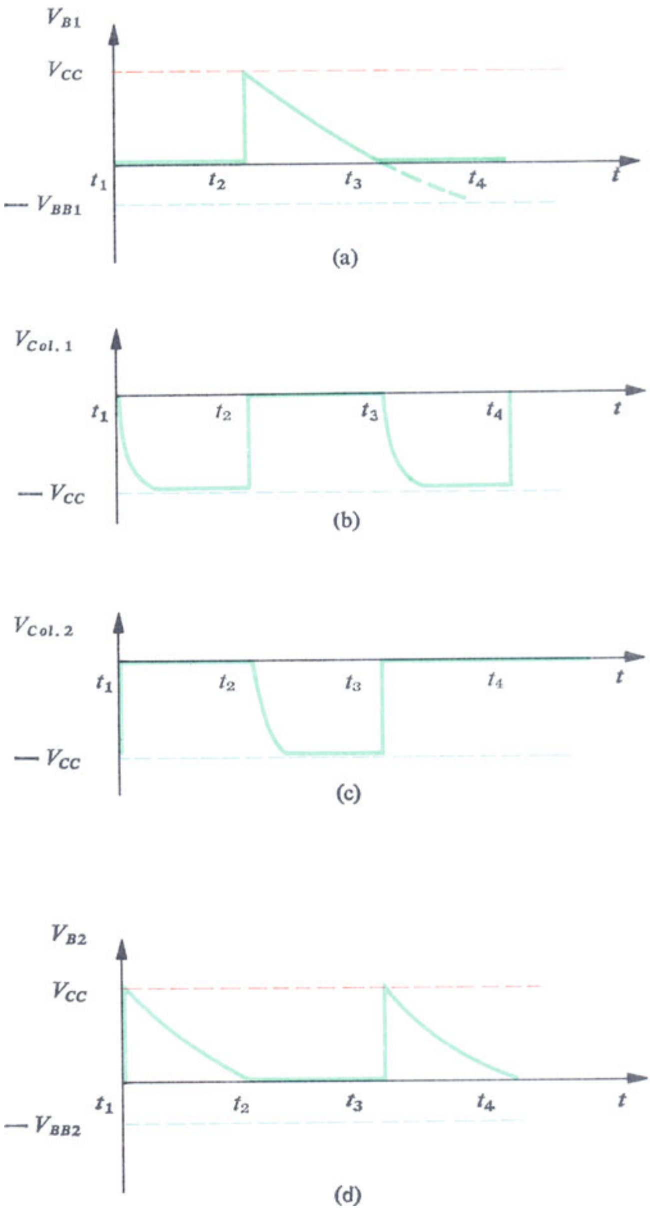
The collector–emitter voltage of transistor T_1 is still zero at this instant (Fig. 437c); it then increases as a function of the charging of capacitor C_1 through resistor R_{L1}.

As before, it is important to limit the duration of this second state. In fact, it is the start of conduction of transistor T_1 that determines the triggering of the system (instant t_3 in Fig. 437a).

From instant t_2 to instant t_3 the base–emitter voltage of transistor T_1 changes from $+V_{CC}$ to zero (Fig. 437a) in accordance with the discharge curve of capacitor C_2 through resistor R_{B1}.

During this period the collector–emitter voltage of this transistor increases from zero to $-V_{CC}$ (Fig. 437b) as shown in the charging curve of the capacitor C_1 through resistor R_{L1}.

The base–emitter voltage of transistor T_2 remains near zero (Fig. 437d)



V_{B1}
V_{CC}
t_1
t_2
t_3
t_4
t
$-V_{BB1}$
(a)
$V_{Col.1}$
t_1
t_2
t_3
t_4
t
$-V_{CC}$
(b)
$V_{Col.2}$
t_1
t_2
t_3
t_4
t
$-V_{CC}$
(c)
V_{B2}
V_{CC}
t_1
t_2
t_3
t_4
t
$-V_{BB2}$
(d)

Fig. 437

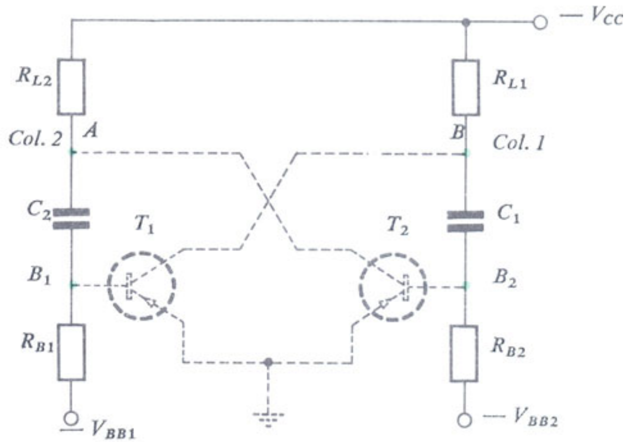


Fig. 438

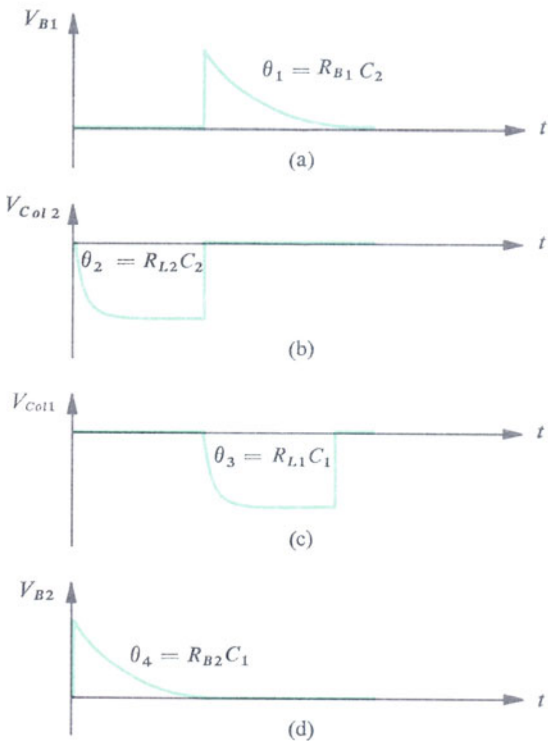


Fig. 439

and the collector–emitter voltage is zero since the collector current is a maximum (Fig. 437c).

These explanations could be continued as far as the definition of instants t_5, t_6, t_7, etc. In practice, however, all we need do is to reproduce the different voltage variations already mentioned in order to obtain the signals illustrated in Figs. 437a, b, c and d.

The variations of the base–emitter voltage of transistors bear no resemblance to a rectangular signal (Figs. 437a and d); on the other hand the variations of the collector–emitter voltage are not very different from the required rectangular form (Figs. 437c and d).

In an astable multivibrator the output signal may be received by the collectors of the two transistors (points A or B in Fig. 438).

Influence of the time constants on the shape and duration of the signal

Let us consider once more the voltage variations discussed on the previous page and referred to in Figs. 439a, b, c and d. Let the appropriate RC circuits be associated with each of these figures:

(a) $R_{B1}C_2$ circuit in Fig. 439a, corresponding to the variations of the base–emitter voltage of transistor T_1;

(b) $R_{L2}C_2$ circuit in Fig. 439b, to be associated with the different values of the collector–emitter voltage of transistor T_2;

(c) $R_{L1}C_1$ circuit in Fig. 439c, dependent on the collector–emitter voltage of transistor T_1;

(d) $R_{B2}C_1$ circuit in Fig. 439d, for the base–emitter voltage of transistor T_2.

Which circuits affect the pulse duration on the one hand; and which are those that affect the shape of the signal on the other hand?

Determination of the pulse duration. The state of conduction of one transistor, or rather the state of non-conduction of the other, is defined by the time taken for the disappearance of the positive pulse between the base and emitter of the transistor when cut off.

In the light of this reasoning the duration of the pulse between the collector and emitter of transistor T_2 is dependent on the period of non-conduction of this transistor ($R_{B2}C_1$ in Fig. 439d).

Similarly, the duration of the pulse on the collector of transistor T_1 is defined by the period of non-conduction of this transistor ($R_{B1}C_2$ in Fig. 439a).

It will be remembered that the time taken for the charging of a capacitor

C through a resistor R is in practice of the order of five times the time constant of the circuit:

$$t_C = 5\theta = 5\,RC.$$

Influence of the RC circuits on the shape of the signal. The signals received at the collectors of the transistors are shown in Figs 439b and c. These shapes approach more nearly the ideal shapes as the speed of charging of capacitors C_1 and C_2 through resistors R_{L1} and R_{L2} increases.

In Fig. 439a, the time constant of the $R_{L2}C_2$ circuit should be as short as possible.

The same applies to the $R_{L1}C_1$ circuit in Fig. 439c in order to take advantage of a very rapid voltage variation.

RC Circuits with long time constants

If we assume for the components R_{L1}, R_{L2} and C_1, C_2 values such that the time constants θ_1 and θ_2 are long with respect to the pulse durations, the variations of the collector–emitter voltage of each transistor become less and less rectangular in shape.

In order to illustrate better the differences appearing in this case, it will be useful to define afresh the shapes of the voltage available at the various 'active' points of the circuit.

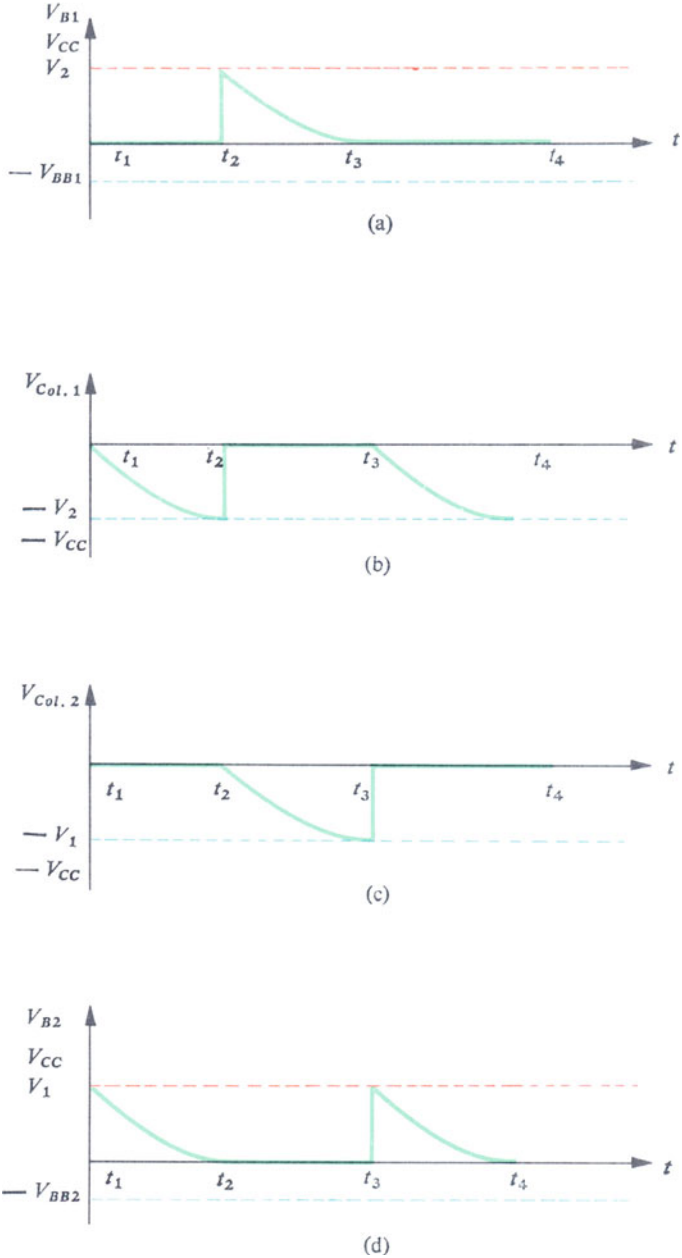
The changes made to the base–emitter voltage of transistor T_1 are very small except for a positive blocking pulse, the strength of which decreases as the time constant of the $R_{L2}C_2$ circuit increases (i.e. the voltage $+V_2$ in Fig. 440a).

The same reasoning applies to the base–emitter voltage of transistor T_2; the positive blocking pulse is smaller and equal to $+V_1(V_1 < V_{CC})$ in Fig. 440d.

The collector–emitter voltage of transistor T_2 never reaches the value of the supply voltage $-V_{CC}$; it is limited to the value $-V_2$ located between 0 and $-V_{CC}$ on the ordinate of Fig. 440b.

The collector–emitter voltage of transistor T_1 does not exceed the value V_1 which differs further from the value of the supply voltage as the time constant $R_{L1}C_1$ increases (Fig. 440c).

The signal available at the output of a multivibrator of this nature is now very different from the ideal rectangular shape required, and it will therefore always be necessary to take into account the influence of the RC circuits on the shape of these signals.



V_{B1}
V_{CC}
V_2
t
t_1
t_2
t_3
t_4
$-V_{BB1}$
(a)
$V_{Col.\,1}$
t
t_1
t_2
t_3
t_4
$-V_2$
$-V_{CC}$
(b)
$V_{Col.\,2}$
t
t_1
t_2
t_3
t_4
$-V_1$
$-V_{CC}$
(c)
V_{B2}
V_{CC}
V_1
t
t_1
t_2
t_3
t_4
$-V_{BB2}$
(d)

Fig. 440

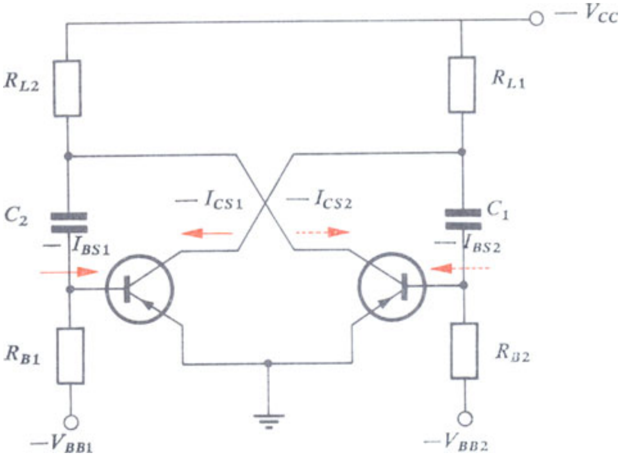


Fig. 441

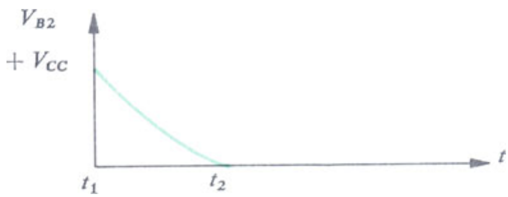


Fig. 442

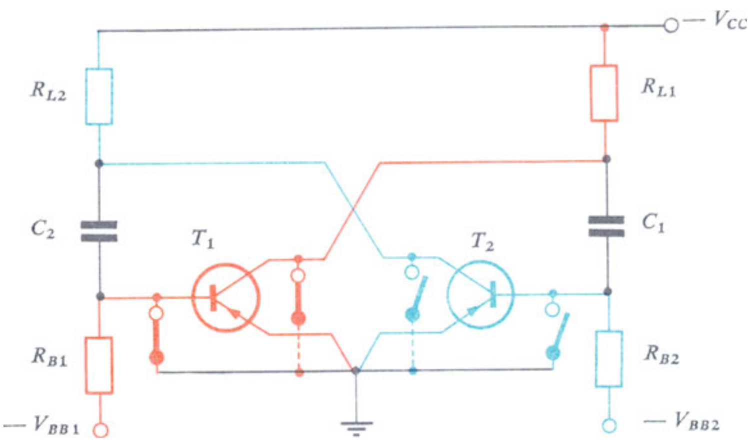


Fig. 443

7.3. Calculation of the instantaneous values of current and voltage

The various voltages appearing between the 'active' electrodes (collector and base) of transistors and earth can be illustrated diagrammatically. A study of this sort however is far from complete as it takes no account of the currents in these electrodes.

If we examine the circuit in Fig. 441 we see that:

(a) the collector of transistor No. 1 has a current ($-I_{CS1}$) flowing through it during its period of saturation;

(b) the base of this transistor also has a current ($-I_{BS1}$) flowing through it;

(c) the collector of transistor No. 2 has a current $-I_{CS2}$ flowing through it during its period of saturation;

(d) the base of this transistor has a current $-I_{BS2}$ flowing through it during this period.

The four currents must, therefore, be added to the above four voltages.

Static operating conditions

We have mentioned two pseudostable states in the previous pages. These conditions can be associated with the limiting operating states of transistors. It will be remembered that the first pseudostable state was associated with the following operating state of the transistors: (a) T_1 saturation; (b) T_2 cut-off.

The second pseudostable state therefore corresponds to opposing states of the transistors, namely: (a) T_1 cut-off; (b) T_2 saturation.

Before we pass on to the determination of variations with time of effective voltages and currents, it will be interesting to consider how they are affected by the values associated with the corresponding static operating conditions.

First pseudostable state. This state appears on the time scale at instant t_1 and is maintained until instant t_2 (Fig. 442).

Transistor T_1 is saturated, and it can therefore be likened to two switches in the closed position (Fig. 443). On the other hand, the cut-off state of the other transistor involves the use of two switches in the open position.

What are the voltages and currents whose values differ from zero during the whole of this period?

For transistor T_1 the instantaneous earthing of the collector and base involves a zero value for the two corresponding voltages; on the other hand, currents flowing through the switches representing the collector and base, namely:

(a) V collector $T_1 = 0$ from instant t_1 to instant t_2;

(b) V base $T_1 = 0$ during this same period;

(c) I collector $T_1 = -I_{CS1}$, collector saturation current;

(d) I base $T_1 = -I_{BS1}$ base saturation current.

Second pseudostable state. This state exists only from instant t_2 to instant t_3 on the time scale of Fig. 444.

Transistor T_1 is cut off (simplified representation by two open switches in Fig. 445) and transistor T_2 saturated (two closed switches).

In these conditions, the various values of voltages and currents are:

(a) for the collector of transistor T_2, a zero voltage and a current equal to its saturation value $(-I_{CS2})$;

(b) for the base of this transistor, a zero voltage and a current equal to $-I_{BS2}$.

We have deliberately omitted any suggestions for calculations concerning the transistor when cut off. In practice the open positions of the corresponding switches imply zero values of current; the voltages then no longer depend on the characteristics of the transistor but rather on the RC circuits used in the coupling units.

Dynamic operating conditions

The voltages and sometimes the currents do not immediately reach their final value at the instant of triggering. In order to obtain a better insight into the effective shape of the signals in a circuit of this nature, it seems necessary to define the equations with which we can calculate these parameters at any given instant.

The limiting values of current and voltage which we will be determining in more detail below should be forgotten for the moment, and the symbols involved in the present study should represent instantaneous values, hence the nomenclature:

(a) $v_{Col\,1(t)}$ and $v_{Col\,2(t)}$ for the collector and emitter voltages;

(b) $v_{B1(t)}$ and $v_{B2(t)}$ for the base and emitter voltages;

(c) $i_{Col\,1(t)}$ and $i_{Col\,2(t)}$ for the collector currents;

(d) $i_{B1(t)}$ and $i_{B2(t)}$ for the base currents.

First pseudostable state. At instant t_1 transistor T_1 has suddenly changed from cut-off to saturation; on the other hand transistor T_2 (the cause of this change of state) changes instantaneously from saturation to cut-off.

The blocking of this transistor is due to the appearance of a strong positive voltage between its base and emitter $(+V_{CC})$. Capacitor C_1 is then charged positively at this instant (Fig. 446).

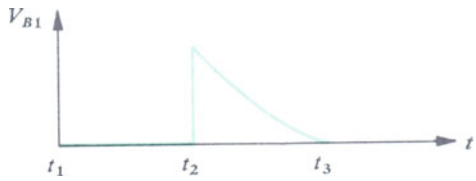


Fig. 444

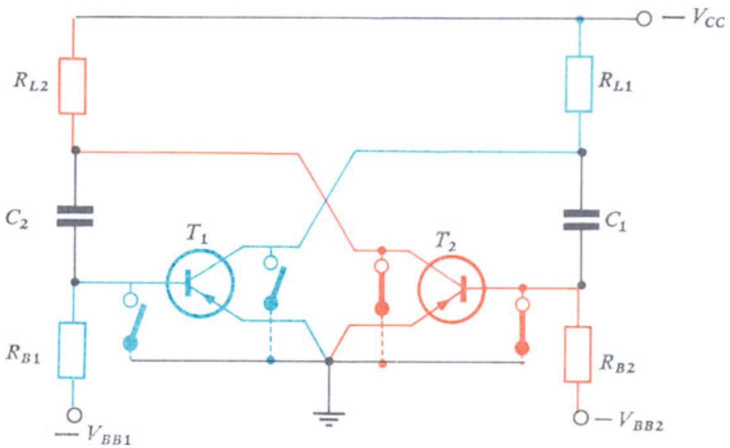


Fig. 445

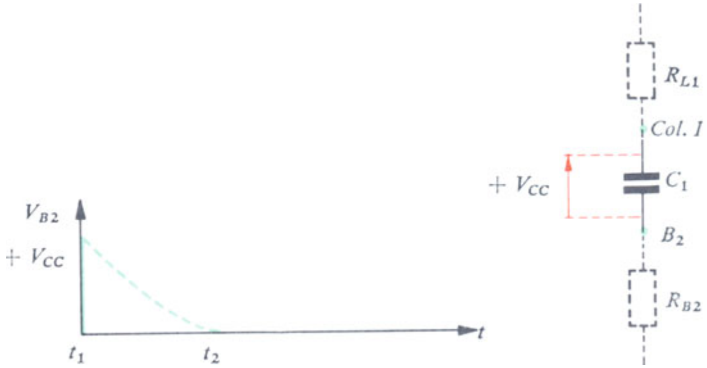


Fig. 446

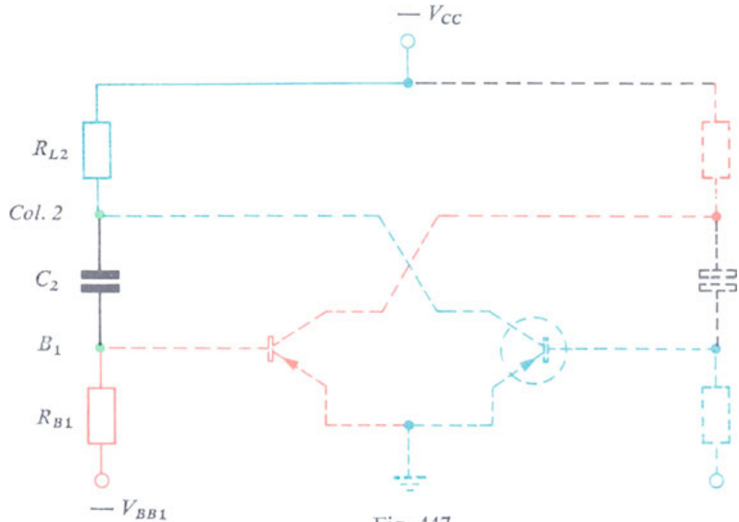


Fig. 447

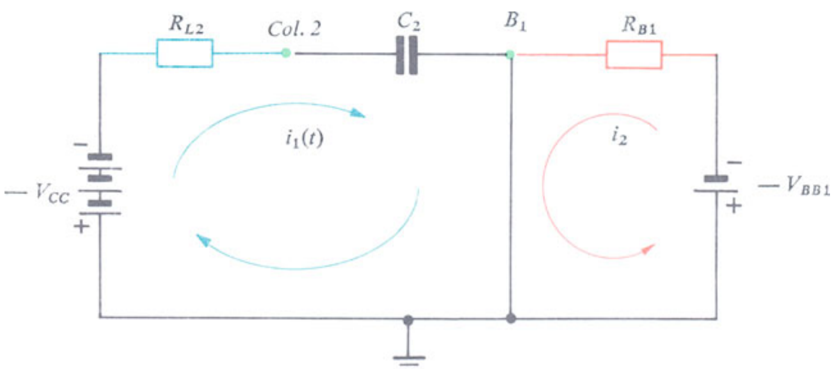


Fig. 448

The calculation of the instantaneous values of the voltages and currents may be facilitated by the choice of simplified equivalent circuits more suited to the type of operation concerned.

The branch at the right of Fig. 447 should be represented by the dynamic circuit of Fig. 448, which consists of two meshes:

(a) the collector circuit of transistor T_2 associated with the group of supply batteries $-V_{CC}$;

(b) the base circuit of transistor T_1 connected to the biasing battery $-V_{BB1}$.

In this figure point B_1 (representing the base of transistor T_1) is earthed, and its voltage is therefore always zero at whatever instant we are concerned.

On the other hand point Col_2 (collector of transistor T_2) is connected directly to the terminals of capacitor C_2, and its voltage depends exclusively on the charge in this capacitor through resistor R_{L2}.

The short-circuit connecting point B_1 to earth shows that it is necessary to calculate two separate currents in order to define the base current of transistor T_1. During the transient period this current is given by the relation:

$$i_{B1t} = i_{1t} + i_2$$

where the current i_{1t} represents the charging current of capacitor C_2 through resistor R_{L2}, that is an expression of the form:

$$i_{1t} = \frac{-V_{CC} + v_{C2t}}{R_{L2}}$$

and the current i_2, the almost constant bias current of transistor T_1, namely:

$$i_2 = -V_{BB1}/R_{B1}.$$

The relation giving the value of the base current and its variations with time is of the form:

$$i_{B1t} = i_{1t} + i_2 = \frac{V_{CC} + v_{C2t}}{R_{L2}} + \frac{V_{BB1}}{R_{B1}}.$$

Since this state exists only from instant t_1 to instant t_2 the calculation of the current may be limited to the duration just mentioned. In these conditions the above formula will no longer include the symbol t but introduces the representation of time $(t_2 - t_1)$ instead:

$$i_{B1}(t_2 - t_1) = \frac{V_{CC} + v_{C2}(t_2 - t_1)}{R_{L2}} + \frac{V_{BB1}}{R_{B1}}.$$

The collector current of transistor T_2 is zero during this period, since it is in the cut-off state.

We will now consider the branch on the right of Fig. 449. In order to facilitate the calculations it is again desirable to separate the various electrical elements into two meshes (Fig. 450):

(a) on the left, the collector current of transistor T_1 and the group of supply batteries $-V_{CC}$;

(b) on the right the base of transistor T_2 and its biasing battery $-V_{BB2}$.

A similar reasoning to that used on the previous page enables us to calculate the instantaneous values of the collector–emitter voltages of transistor T_1 and the base–emitter voltage of transistor T_2 as well as the collector and base currents of these two transistors.

Since the collector of transistor T_1 is connected directly to earth its voltage is zero during the whole duration of this pseudostable state, that is:

$$v_{CE1} = 0.$$

The base–emitter voltage of transistor T_2 depends above all on the discharge of capacitor C_1 through resistor R_{B2}. At the instant when this capacitor is completely discharged the multivibrator will trigger. This voltage is consequently given by the relation:

$$V_{BE2} = -V_{BB2} + (V_{CC} + V_{BB2}) \exp\left(-\frac{t_2 - t_1}{K_1 \circ R_{B2}}\right).$$

The collector current of transistor T_1 is equal to the difference of the two currents:

(a) the current i_3, practically constant, due to the voltage $-V_{CC}$ in resistor R_{L1};

(b) the current $i_{4(t)}$ a function of the discharge of capacitor C_1 through resistor R_{B2}.

These two currents are given respectively by the expressions:

$$i_3 = \frac{-V_{CC}}{R_{L1}} \quad \text{and} \quad i_4(t_2 - t_1) = \frac{V_{BB2} + v_{BE2}(t_2 - t_1)}{R_{B2}}.$$

The collector current of transistor T_1 in these conditions is equal to:

$$i_{C1}(t_2 - t_1) = i_3 - i_4(t_2 - t_1) = \frac{-V_{CC}}{R_{L1}} - \frac{V_{BB2} + v_{BE2}(t_2 - t_1)}{R_{B2}}.$$

Transistor T_2 is cut off; its base–emitter junction is strongly biased in the reverse direction and the base current is practically negligible; namely:

$$i_{B2} = 0.$$

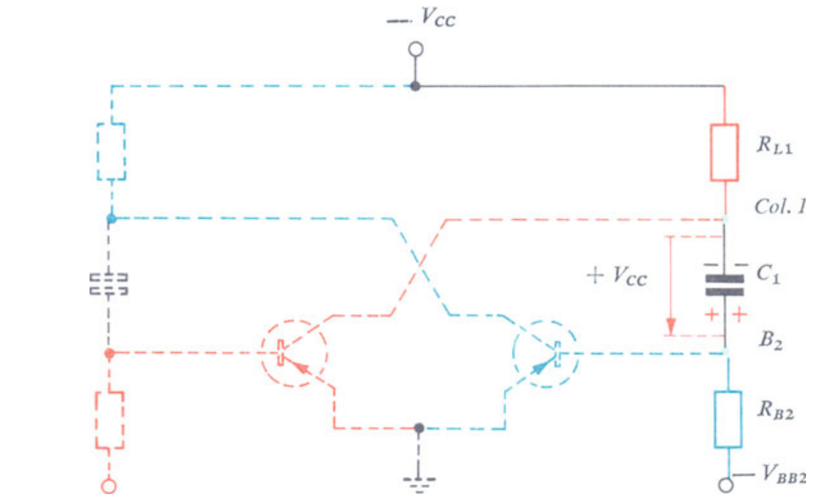


Fig. 449

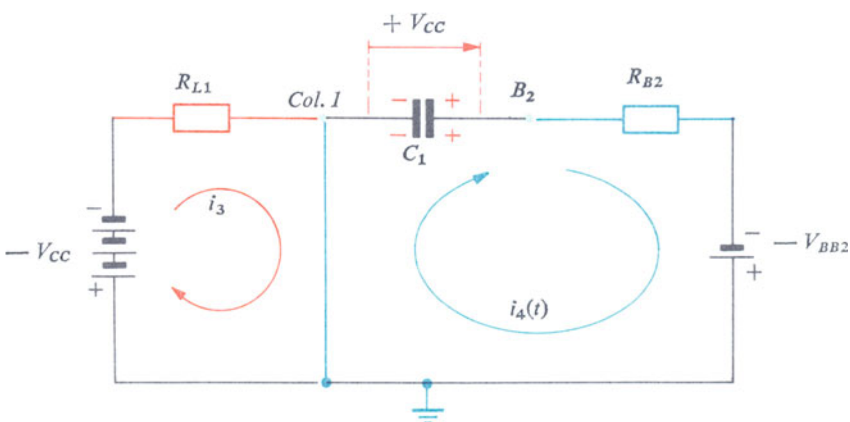


Fig. 450

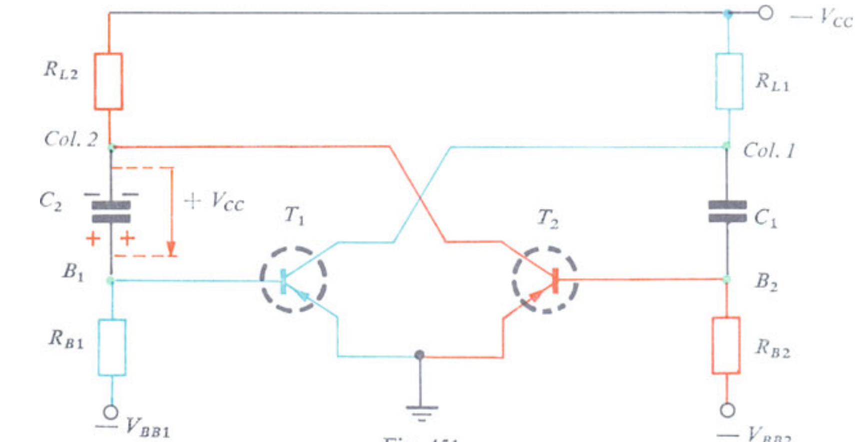


Fig. 451

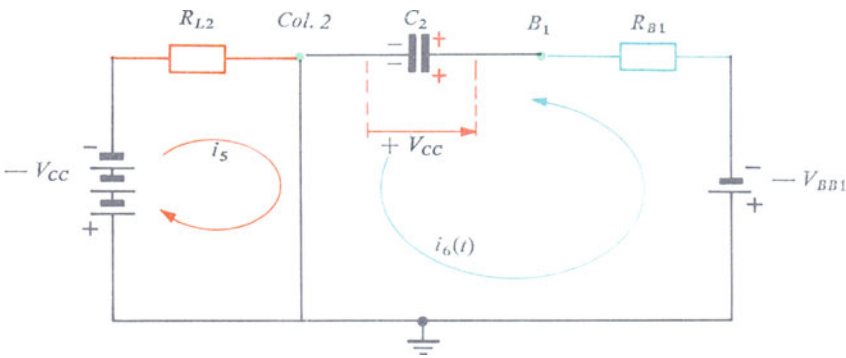


Fig. 452

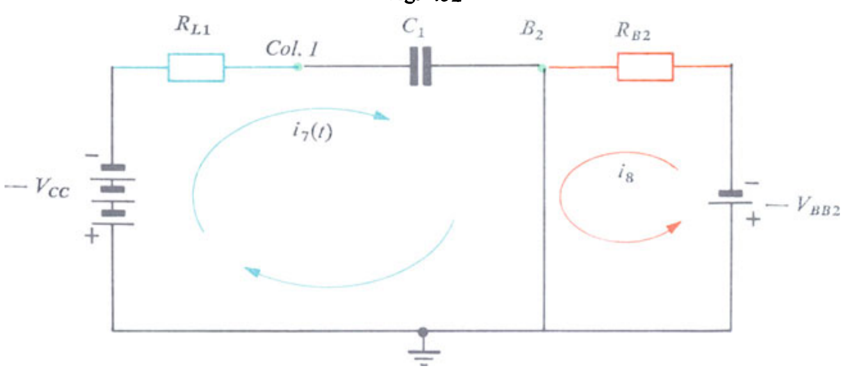


Fig. 453

Second pseudostable state. At instant t_2 the multivibrator triggers and the two transistors change their operating state; T_1 becomes cut-off and T_2 saturated. This persists until instant t_3.

The calculation of the instantaneous values of the voltages and currents may, therefore, be performed by the previous method provided the switches in the closed position are replaced by those in the open position.

We will first consider the left side of Fig. 451, the electrical representation of which can conveniently be replaced by the dynamic representation of Fig. 452.

The collector voltage of transistor T_2 is always zero, that is:

$$v_{CE2} = 0.$$

The base–emitter voltage of transistor T_1 is first strongly positive and then decreases until instant t_3 at which it disappears, and the multivibrator then triggers again. This voltage decreases in accordance with the expression:

$$V_{BE1(t_3-t_2)} = -V_{BB1} + (V_{CC} + V_{BB1}) \exp\left(-\frac{t_3-t_2}{C_2 R_{B1}}\right).$$

The collector current of transistor T_2 is equal to the sum of two currents:

(a) A constant current i_5 due to voltage $-V_{CC}$ in resistor R_{L2}:

$$i_5 = \frac{-V_{CC}}{R_{L2}}.$$

(b) The discharge current of capacitor C_2 through resistor R_{B1};

$$i_{6(t_3-t_2)} = \frac{V_{BB1} + v_{BE1(t_3-t_2)}}{R_{B1}}.$$

This current is consequently given by the expression:

$$i_{C2(t_3-t_2)} = \frac{-V_{CC}}{R_{L2}} - \frac{V_{BB1} + v_{BE1(t_3-t_2)}}{R_{B1}}.$$

The base current of transistor T_1 is zero due to the cut-off conditions of this transistor during this operating period.

Let us now revert to Fig. 451 and replace the branch on the right side by the equivalent circuit of Fig. 453.

The emitter–collector voltage of transistor T_1 is dependent on the charge of capacitor C_1 through resistor R_{L1}:

$$v_{CE1(t_3-t_2)} = -V_{CC}\left\{1 - \exp\left(\frac{t_3-t_2}{C_1 R_{L1}}\right)\right\}.$$

The base–emitter voltage of transistor T_2 is zero because the transistor is saturated:

$$v_{BE2} = 0.$$

There is no collector current in transistor T_1 (cut-off state):

$$i_{C1} = 0.$$

The base current of transistor T_2 is equal to the sum of two currents (Fig. 454):

(a) The charging current of capacitor C_1 through resistor R_{L1}:

$$i_{7(t_3-t_2)} = \frac{-V_{CC}+v_{C1(t_3-t_2)}}{R_{L1}}.$$

(b) A constant current i_8 due to the biasing voltage $-V_{BB2}$ in resistor R_{B2}:

$$i_8 = \frac{-V_{BB2}}{R_{B2}}.$$

In these conditions the base current of transistor T_2 is given by the expression:

$$i_{B2(t_3-t_2)} = \frac{-V_{CC}+v_{C1(t_3-t_2)}}{R_{L1}} + \frac{-V_{BB2}}{R_{B2}}.$$

Finally, the two 'pseudostable' states of an astable multivibrator may be defined quantitatively by a number of relations. The voltages may be calculated with the aid of these expressions, or more simply by means of the diagrams in Figs 455a, b, c and d. On the other hand if we wish to know the values of current precisely at every instant, we must use expressions suitable for the two possible states.

The object of the table below is therefore to regroup these expressions so as to make them suitable for immediate application.

	t_2-t_1 T_1: sat.; T_2: cut-off	t_3-t_2 T_1: cut-off; T_2: sat.
v_{BE1}	0	$-V_{BB1}+(v_{CC}+v_{BB1}) \exp\left(-\frac{t_3-t_2}{R_{B1}\cdot C_2}\right)$
v_{CE2}	$-V_{CC}\left\{1-\exp\left(\frac{t_2-t_1}{R_{L2}\cdot C_2}\right)\right\}$	0
v_{CE1}	0	$-V_{CC}\left\{1-\exp\left(\frac{t_3-t_2}{R_{L1}\cdot C_1}\right)\right\}$
v_{BE2}	$-V_{BB2}+(V_{CC}+V_{BB2}) \exp\left(-\frac{t_2-t_1}{R_{B1}\cdot C_1}\right)$	0
i_{B1}	$\frac{-V_{CC}+v_{(C2)(t_2-t_1)}}{R_{L2}}+\frac{-V_{BB1}}{R_{B1}}$	0
i_{C2}	0	$\frac{-V_{CC}}{R_{L2}}-\frac{V_{BB1}+v_{BE1(t_3-t_2)}}{R_{B1}}$
i_{C1}	$\frac{-V_{CC}}{R_{L1}}-\frac{V_{BB2}+v_{BE2(t_2-t_1)}}{R_{B2}}$	0
i_{B2}	0	$\frac{-V_{CC}+v_{(C1)(t_3-t_2)}}{R_{L1}}+\frac{-v_{BB2}}{R_{B2}}$

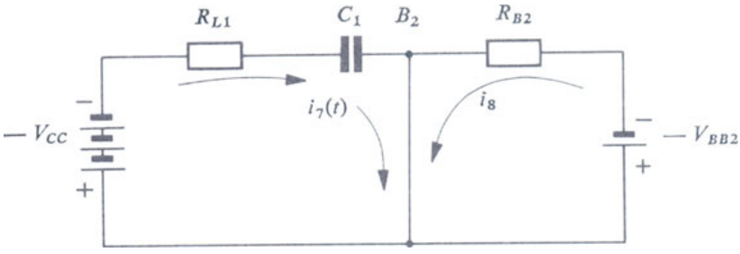


Fig. 454

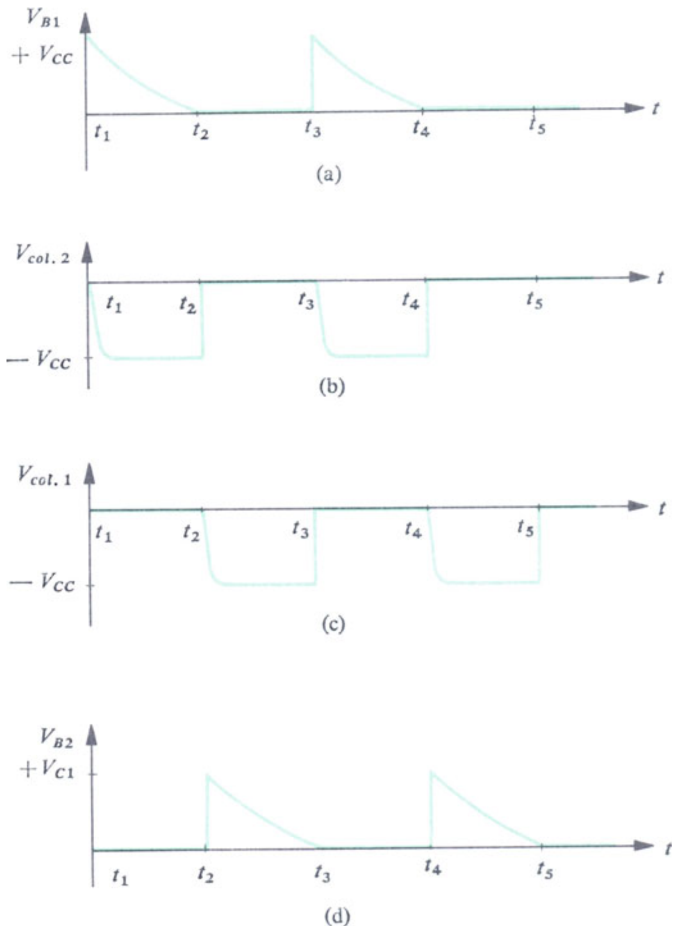


Fig. 455

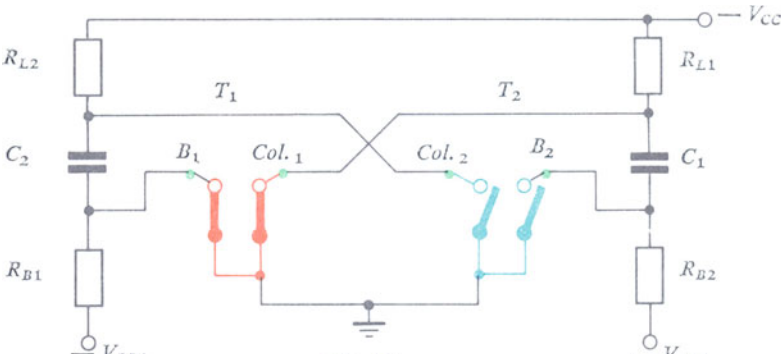


Fig. 456

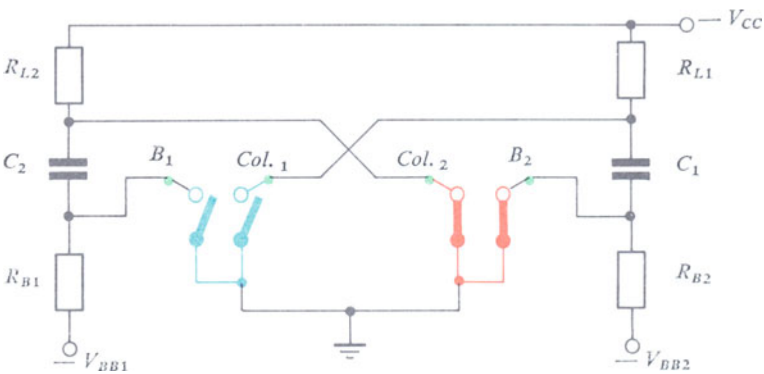


Fig. 457

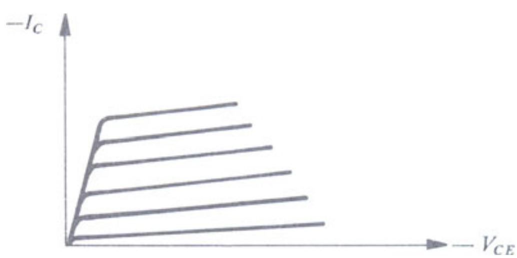


Fig. 458

Respect of the cut-off and saturation states

Satisfactory operation of a multivibrator depends essentially on the precise control of the shape and duration of the pulses constituting the output signal. The transition from cut-off to saturation or from saturation to cut-off should occur only due to the charge or discharge of one of the capacitors in the coupling systems.

It is, therefore, very important to prevent any possible variation in the operating state of a transistor in either the cut-off or saturation state. This is done quantitatively by means of the expressions on the previous page.

First stage (t_1 *to* t_2). Transistor T_1 is saturated and T_2 cut off (Fig. 456).

The saturation condition of the first of these transistors is defined in the static state by the following relation:

$$i_{C1}/i_{B1} < h_{FE}(T_1).$$

Since this transistor should remain saturated during the whole of this first stage, it is essential that the relation between the collector current and the base current should at all times be lower than the current gain of transistor T_1:

$$\frac{i_{C1(t_2-t_1)}}{i_{B1(t_2-t_1)}} < h_{FE}(T_1).$$

Respecting the cut-off condition of the second transistor depends entirely on the discharge of capacitor C_1 through resistor R_{B2}.

Second stage (t_2 *to* t_3). Transistor T_1 is cut off and transistor T_2 saturated (Fig. 457).

The first transistor remains cut off as long as is necessary for the discharge of capacitor C_2 through resistor R_{B1}.

The state of saturation of the second transistor should be maintained from instant t_2 to instant t_3, making it necessary during the whole of this period that the following relation should be respected:

$$\frac{i_{C2(t_3-t_2)}}{i_{B2(t_3-t_2)}} < h_{FE}(T_2).$$

Rapid switching

It is not always possible to liken transistors to groups of ideal switches in the open or closed positions representing their operating conditions. A simple glance at the network of output characteristics (Fig. 458) of one of the transistors shows:

(a) that saturation is equivalent to the connection between the collector and emitter of a certain leakage resistor (in red in Fig. 459);
(b) that the cut-off involves the connecting of a high-value resistor (in blue in the figure) in parallel with the switches in the open position.

An accurate calculation of the operating conditions of a multivibrator calls for the inclusion of these resistors in the equations.

Further, the use of transistors in rapid switching requires that we take into consideration the inertial effects represented electrically by capacitative effects.

The complete study of a multivibrator when operating at high frequencies involves the replacement of the switches in Fig. 460 by electrical units in the equivalent dynamic circuit of the switching transistor (Fig. 461).

The large variations in the parameters of this equivalent circuit make it necessary to divide the design of a multivibrator of this type into a certain number of phases. We do not intend in this book to enter into the details of such a method of calculation because of its very special nature.

7.4. Calculation of powers

It very seldom happens that multivibrators are produced for the supply of large signals. It is however of interest, even in a standard design, to be able to calculate the various amounts of power occurring in a circuit of this type.

There are three such powers:
(a) the power supplied by the battery (P_B);
(b) the output power (P_O);
(c) the power dissipated in the transistor (P_C).

These three powers are connected by the expression:

$$P_B = P_O + P_C.$$

In the chapter dealing with the study of the thermal behaviour of switching transistors, we pointed out the possibility of calculating these various powers and especially of determining accurately, for a given rectangular signal, the value of the power dissipated in the interior of the transistor. Knowing that this concept constitutes one of the limiting conditions of use, it is important always to take account of it when embarking on the design of a project.

Another important parameter for large signals is given by the ratio between the output power and the power supplied by the battery:

$$\eta = P_O/P_B.$$

The efficiency thus obtained should be as near as possible to 100% (an ideal case of course never realized).

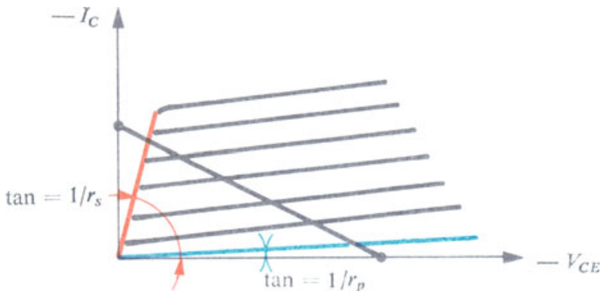


Fig. 459

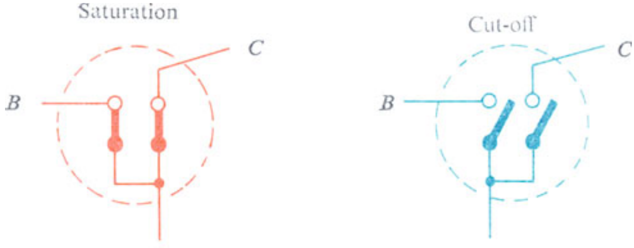


Fig. 460

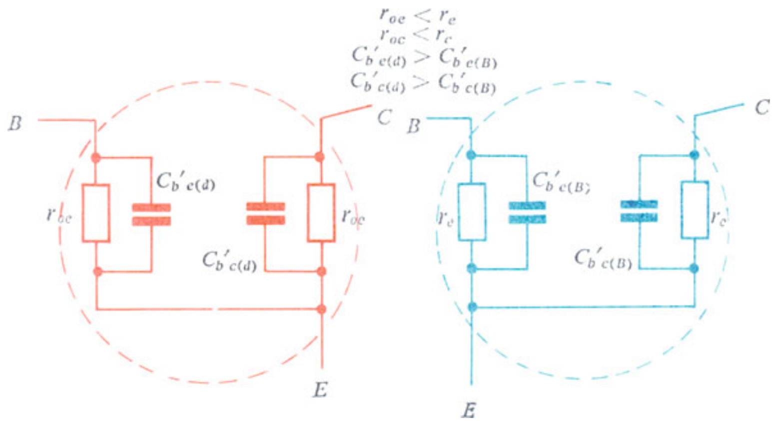


Fig. 461

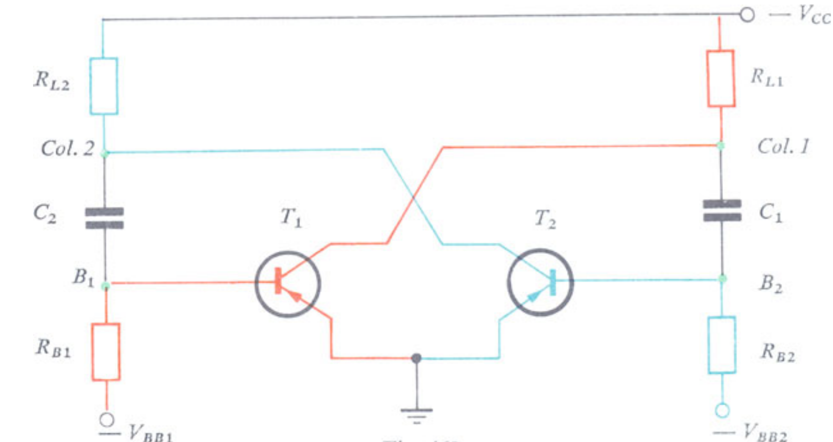


Fig. 462

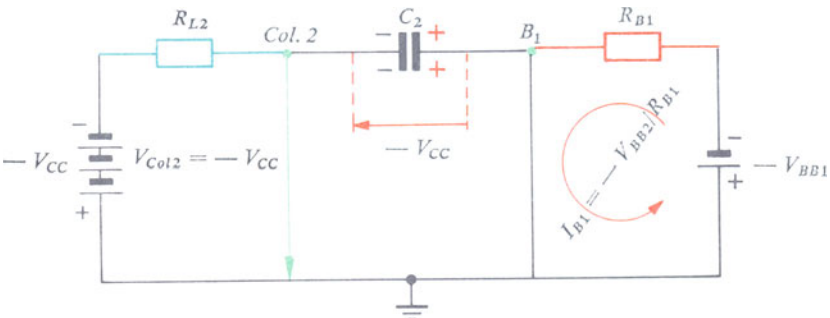


Fig. 463

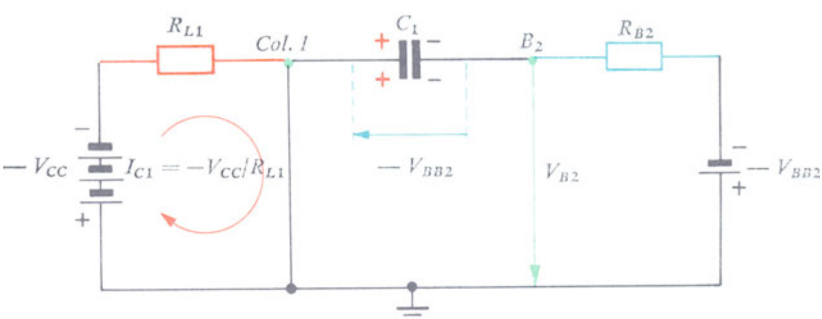


Fig. 464

7.5. Limiting values of the voltage and currents

The calculation of the instantaneous values of the currents and voltages in a multivibrator constitutes one of the most important stages in the design of a circuit of this type. It is, however, interesting to be able to define the maximum and limiting values of these voltages and currents even if they are never reached.

The interest of this study lies especially in the fact that it constitutes an introduction to the calculations essential for the production of monostable or bistable multivibrators.

First pseudostable state

At instant t_1 transistor T_1 is saturated and transistor T_2 cut off (Fig. 462).

Assuming that the multivibrator remains in this operating state indefinitely, then in these conditions the branch on the left of Fig. 462 may be represented by a more functional equivalent circuit (Fig. 463).

In accordance with the method already used in the study of transient effects it is possible to calculate the direct values of the voltages and currents:

(a) the collector–emitter voltage of transistor T_2 reaches the final value $-V_{CC}$;

(b) the base–emitter voltage of transistor T_1 remains equal to 0;

(c) the collector current of transistor T_2 is zero;

(d) the base current of transistor T_1 reaches the limiting value:

$$-V_{BB1}/R_{B1}.$$

The right side of Fig. 462 is now represented by the circuit in Fig. 464:

(a) the collector–emitter voltage of transistor T_1 is zero;

(b) the base–emitter voltage of transistor T_2 would be equal to $-V_{BB2}$, but this is not possible since this transistor could no longer be in the cut-off state;

(c) the collector current of transistor T_1 is a maximum and equal to $-V_{CC}/R_{L1}$;

(d) there is no base current in transistor T_2.

It appears from this study that the cut-off state of transistor T_2 is no longer respected, causing faulty operation of a system of this sort. On the other hand, the saturated state of transistor T_1 can be controlled; this makes it quite certain that when this transistor leaves its saturation state there will be no triggering.

Finally, let us remember that this study of the limiting values of the voltages and currents does not allow the quantitative determination of the satisfactory operation of an astable multivibrator, because the different electrical parameters concerned rarely reach such values. However, if we can be sure that the state of saturation will always be respected, the likelihood of triggering exclusively associated with the discharge of a capacitor (a component external to the transistor) is strengthened.

7.6. Signal shaping at the output

If we examine the variations of the collector–emitter voltage of transistors T_1 and T_2 (Fig. 465) it will be seen that as a function of the values of the time constants for the coupling circuits used, the output signals thus given are very different from the required rectangular shape.

There are several methods for shaping of these signals, and we will deal here with two of them.

Use of diodes in the collectors of the transistors

When the maximum amplitude is not the basic requirement it is possible to use only part of the voltage variation of the output signal and thus to take advantage of a shorter triggering time. The system shown in Fig. 466 operates in this way.

In addition to the normal components of a multivibrator of this type, we find two diodes connected in parallel between the collector of each transistor and an auxiliary biasing battery $-V_{CC}'$. This voltage is lower than that of the supply voltage.

Let us consider for example the variations of the collector–emitter voltage of transistor T_1 (Fig. 467).

During saturation the collector current is a maximum and the corresponding voltage is nearly 0. The diode D_1 connected between the collector and the bias battery $-V_{CC}'$ is biased in the reverse direction and presents to the flow of current a resistance which is theoretically nearly infinite.

When the transistor is switched from saturation to cut-off, the collector current decreases and the collector–emitter voltage increases; the cathode of diode D_1 becomes less and less positive with respect to its anode until the instant when it becomes biased in the forward direction.

In this case the resistance becomes very low and the collector of the transistor T_1 has its voltage limited to $-V_{CC}'$.

If we introduce this value of voltage in Fig. 467 and if we remember that it constitutes a limit never exceeded, the new variations of the collector–emitter voltage of these transistors (shown by the blue line in this figure) are now much more nearly rectangular.

A similar reasoning may be used to describe the improvement of the signals received between the collector of transistor T_2 and earth. Diode D_2 produces in this transistor variations of the collector–emitter voltage limited to $-V_{CC}'$.

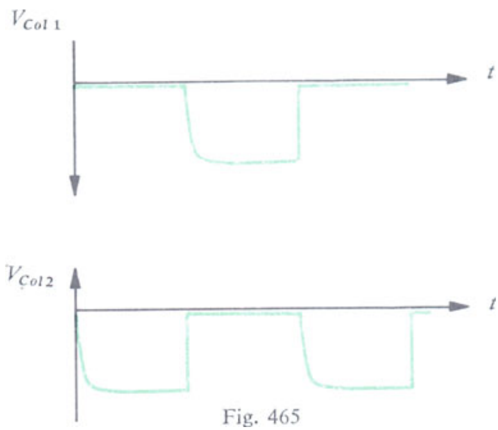


Fig. 465

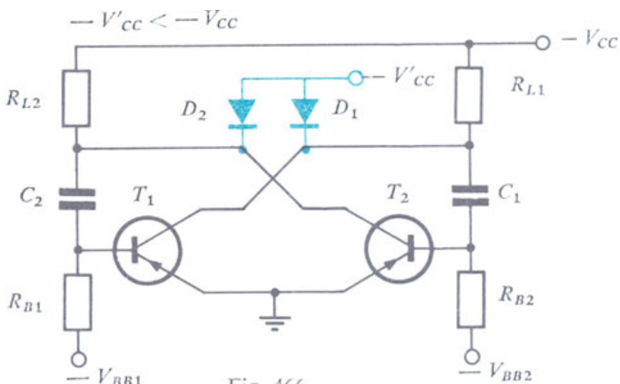


Fig. 466

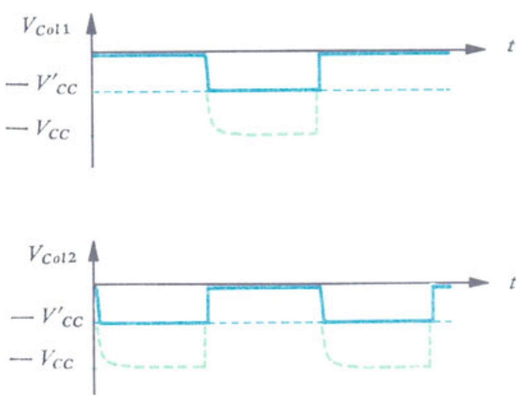


Fig. 467

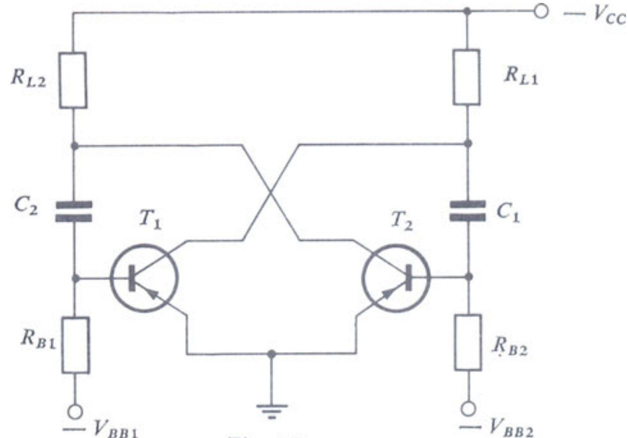


Fig. 468

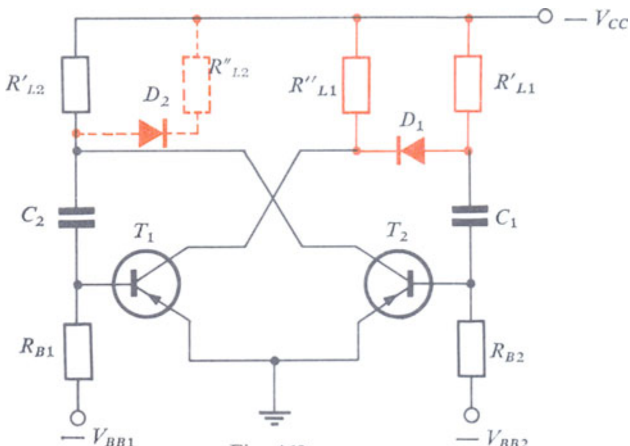


Fig. 469

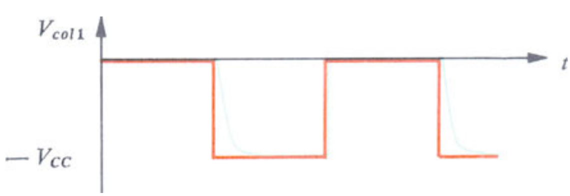


Fig. 470

The shape of the output signal approaches that of a rectangular waveform as the difference between the complementary voltage of the bias of the diodes ($-V_{CC}'$) and the supply voltage ($-V_{CC}$) increases. This improvement is always associated with a decrease in the amplitude of the signal.

Use of diodes associated with resistors

If we refer back to the basic diagram of the astable multivibrator in Fig. 468 we will see that the shape of the signal received in the collector circuit of transistor T_1, for example, is relatively bad due to the charging of capacitor C_1 through the resistor R_{L1}. We must therefore try to separate the coupling circuit from the charging of the capacitor during the period when the collector–emitter voltage is changing from zero to $-V_{CC}$.

In Fig. 469 load resistances of double the previous value are connected in the collectors of the transistors. Each group of resistance also contains a diode.

The special components of this new multivibrator are:

(a) the two load resistances of the transistor T_1 (R_{L1}' and R_{L1}'');

(b) the coupling diode between these two resistances (D_1);

(c) the two load resistances of transistor T_2 (R_{L2}' and R_{L2}'');

(d) the coupling diode between these two resistances (D_2).

The two load resistances of the collector of transistor T_1 are equal; their value is double that of R_{L1}. The same applies to the two load resistances of transistor T_2.

When transistor T_1 is in the saturated state, the corresponding collector–emitter voltage is zero, and the equivalent load resistance is equal to R_{L1} (the diode D_1 behaves practically as a short circuit).

Switching to cut-off causes a large variation of the collector–emitter voltage (from zero to $-V_{CC}$) in Fig. 470.

In the absence of a corrector circuit the variation of the collector–emitter voltage assumes the shape shown in green in the figure.

The fact of having placed the diode D_1 between the two resistors R_{L1}' and R_{L1}'' produces a separation between the coupling circuit and the load of transistor T_1. The negative voltage appearing between the anode and cathode of D_1 biases it in the reverse direction; the collector–emitter voltage of transistor T_1 therefore at once changes to the value $-V_{CC}$ (in red in Fig. 470).

The same reasoning applies to transistor T_2 at the instant when it is switched from the state of conduction to cut-off. The diode D_2, which constitutes a true short-circuit during the first period, then behaves as a switch in the open position. Capacitor C_2 is no longer connected to the collector of transistor T_2 and the collector–emitter voltage changes instantaneously from zero to $-V_{CC}$.

The advantage of a circuit of this type is that it improves the shape of the output signal considerably without changing its amplitude. In practice only one of these circuits is used.

7.7. Actual circuits

There are many types of 'astable multivibrator' circuits, but it is not our intention to consider them all in detail, but rather to discuss one or two practical examples.

Astable multivibrator with two capacitive couplings

A circuit of this type is shown in Fig. 471a, in which will be found the basic diagram discussed earlier.

Beside a corrector circuit (we are dealing here with slow switching) this multivibrator consists of:

(a) the two transistors T_1 and T_2;
(b) the two load resistances, $R_{L1} = R_{L2} = 1\ \text{k}\Omega$;
(c) two biasing resistances, $R_{B1} = R_{B2} = 18\ \text{k}\Omega$;
(d) two coupling capacitors $C_1 = C_2 = 10$ nF.

The supply voltage is equal to -6 V, and it acts at the same time as a biasing voltage.

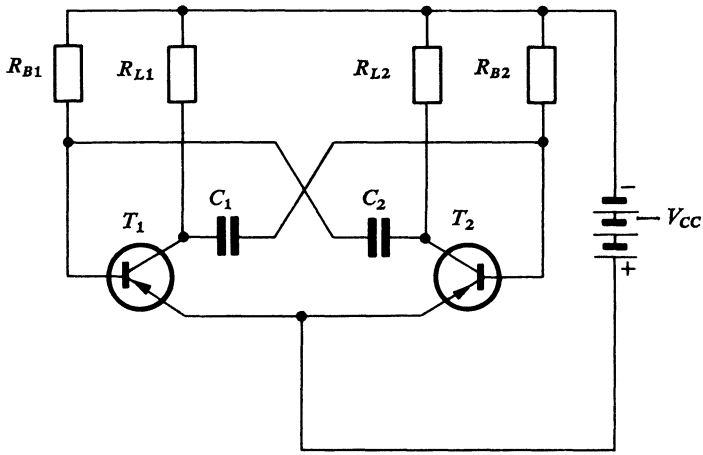
The operating principles of a circuit of this type have already been dealt with at length; it delivers variations of rectangular voltage at the output, or between the collectors and earth.

Astable multivibrator with emitter coupling

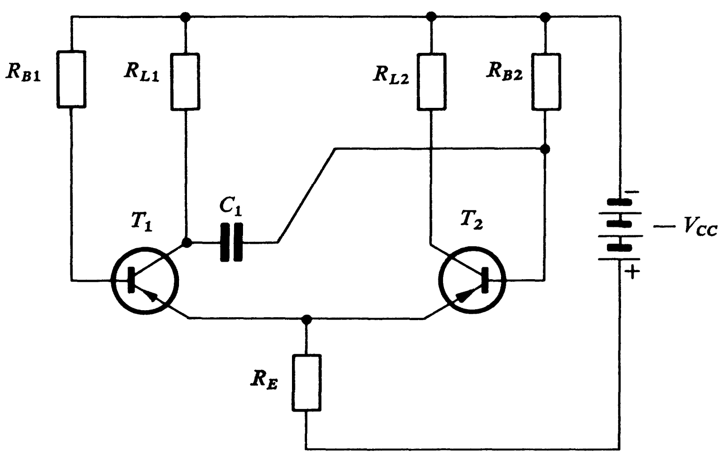
The operation of an astable multivibrator is based on the fact that any variation in the energy at the output of one of the transistors is at once transmitted to the input of the other transistor. The same reasoning applies to the second output and the other input.

This transfer of energy may be obtained by the use of a capacitative coupling between the collector of one transistor and the base of the other; it can also take place by means of a component common to the two transistors and sensitive to the variations of the current in the circuit.

In the circuit of Fig. 471b the common component is the resistor R_E in series with the emitters. This resistor is not decoupled; the voltage at its terminals is therefore dependent on the value of the current flowing through it.



(a)



(b)

Fig. 471

In order to strengthen the tendency to triggering, a supplementary coupling circuit is often associated with this type of energy transmission (capacitor C_1).

8. The Monostable Multivibrator

The control pulses essential for the satisfactory operation of systems using switching circuits are very rapidly distorted by the various units through which they have to pass. For example, the originating pulse shown in Fig. 472 is transformed, after passing through several stages, into a very distorted signal with practically no leading edge (Fig. 472b).

It is, therefore, necessary to reshape these pulses periodically, and this is the function of the monostable multivibrator.

A circuit of this type can be studied from a consideration of a static state, namely the steady state of this multivibrator.

Our examination is arranged in the following order:

(a) concept and principle of operation;
(b) calculation of the values of the voltage and current in the steady state;
(c) physical explanation of the operation with diagrammatic representation of the signals;
(d) discussion of the shapes of the voltage and current appearing during the dynamic period;
(e) problems concerning the control of a multivibrator of this nature;
(f) different types of circuit.

8.1. Concept and principle of operation of a monostable multivibrator

The basic diagram of a monostable multivibrator is shown in Fig. 473. It consists of two transistors (T_1 and T_2) supplied by a group of batteries ($-V_{CC}$). Each transistor is biased by a battery, namely $-V_{BB1}$ for transistor T_1 and $+V_{BB2}$ for transistor T_2. The various components are:

(a) the load resistances (R_{L1} and R_{L2});
(b) the bias resistors (R_{B1} and R_{B2});
(c) the coupling system (R_1 and C_2).

Definition of the steady state

The term 'monostable' used to describe this type of multivibrator implies the existence of one stable state.

In a multivibrator containing two transistors, one of them is saturated while the other is cut-off. It is therefore desirable at this stage, when examining Fig. 473, to state which transistor is saturated and therefore which one is cut-off.

The saturated state is defined as that in which the two junctions are

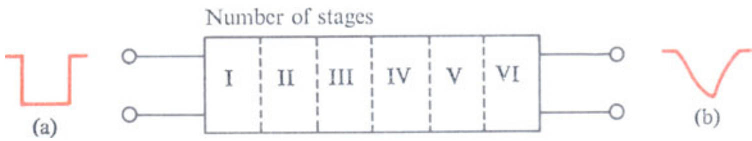


Fig. 472

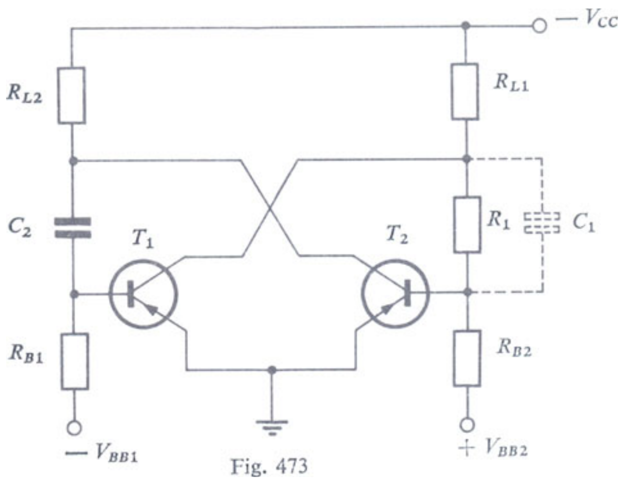


Fig. 473

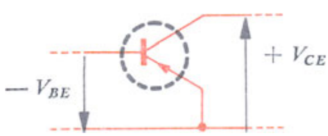


Fig. 474

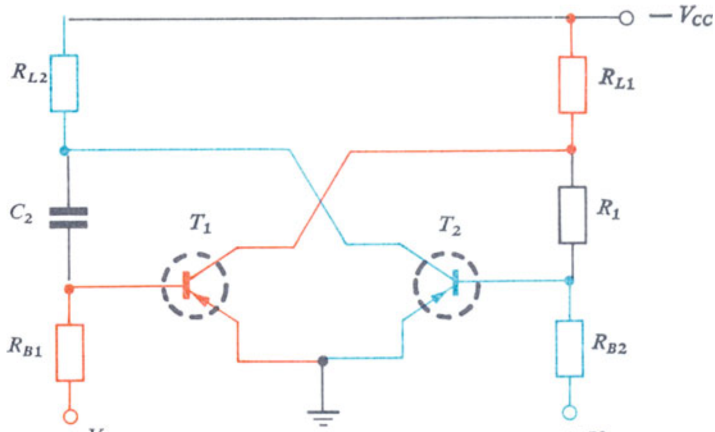


Fig. 475

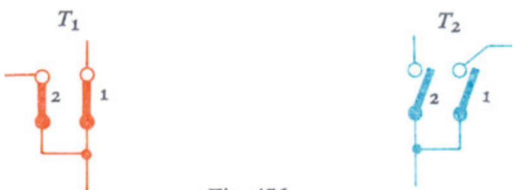


Fig. 476

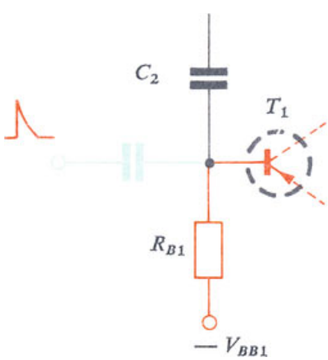


Fig. 478

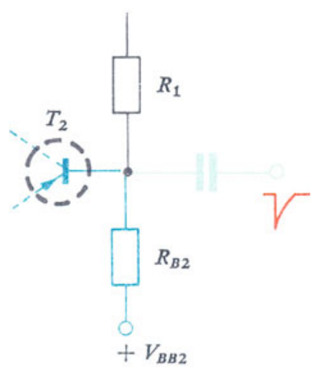


Fig. 477

biased in the forward direction. For a *PNP* transistor this condition is respected when the base is negative with respect to the emitter and collector (Fig. 474).

If we examine the bias circuits of the two transistors we see that only transistor T_1 is connected directly to the terminals of a source of negative voltage ($-V_{BB1}$). It would always be possible to give resistors R_{L1}, R_1 and R_{B2} such values that the base of transistor T_2 would be negative with respect to its emitter. In the steady state, the presence of the capacitor C_2 automatically causes the voltage between the base and the emitter of transistor T_1 to be negative. Finally, the values of the resistors R_{L1}, R_1 and R_{B2} will be calculated so as to cause transistor T_2 to be in the cut-off state.

In the stable state of such a multivibrator we have:

(a) T_1 saturation;

(b) T_2 cut-off.

Beside any information applied to or injected into the input of this circuit, it maintains this operating state. We can therefore now produce a simplified equivalent circuit in which transistor T_1 is represented by two switches in the closed position and transistor T_2 by two switches in the open position. This is illustrated in Fig. 476.

We will deal later with the calculation of the various currents and voltages, merely stating at this stage that the quantitative definition of a state of this sort will be very simple.

Influence of a pulse

There may be two types of pulse:

(a) negative pulses;

(b) positive pulses.

A negative pulse must be applied between the base and emitter of the transistor when cut-off (Fig. 477).

If this pulse is directed towards the input terminals of the saturated transistor it will merely strengthen this state and the multivibrator will not be triggered.

The positive pulse would normally be applied between the base and emitter of the saturated transistor (Fig. 478). The result would be the same as in the previous case if it is wrongly directed.

Let us consider the case of a negative pulse applied between the base and emitter of transistor T_2 (in the cut-off state). This pulse biases the base–emitter junction of this transistor in the forward direction; a current appears in the collector circuit and the collector–emitter voltage ($-V_{CE2}$), which was previously equal to $-V_{CC}$, now decreases (Fig. 475).

This positive voltage variation is immediately transmitted to the base of transistor T_1. The corresponding collector current starts to decrease whilst the collector voltage increases. This negative voltage variation is applied between the base and emitter of transistor T_2 causing another improvement in the collector current of this transistor.

The process continues and transistor T_2 very soon becomes saturated and T_1 cut-off.

The cut-off state of transistor T_1 is explained by the presence of a positive voltage between the base and emitter; this voltage appears at the terminals of the capacitor C_2. As the 'base–emitter' switch of transistor T_2 was open (Fig. 479), capacitor C_2 discharges through resistor R_{B1} and the battery $-V_{BB1}$. It then tends to recharge in the other direction.

Just as for the astable multivibrator, the trigger condition occurs at the instant when the cut-off transistor starts to conduct. It is then at this instant, when the positive voltage between the base and emitter of T_1 disappears, that the change of state occurs.

We need not remind ourselves in detail of the different effects that accompany the triggering of a multivibrator: all that is necessary is to refer to the earlier explanations, reversing the electrodes concerned.

This circuit then returns to its initial operating conditions, namely:

(a) T_1 saturated;

(b) T_2 cut-off.

The cut-off state of T_2 is not due to the positive charge of a capacitor but is caused by the values of the resistors in the bridge R_{L1}, R_1 and R_{B2}. This stage remains in this state until the arrival of another pulse.

If we examine the signal received at one of the collectors (T_1 or T_2) we see that only a single pulse is available. It is due to the drive pulse causing the first triggering of this multivibrator.

8.2. Calculation of the values of voltages and currents in the steady state

In the absence of any signal applied at the input of one of the transistors, the multivibrator operates in the conditions already defined: T_1 saturated and T_2 cut-off.

We can therefore pass from the full diagram of Fig. 479 to the simplified one of Fig. 480, in which the transistors are represented by two groups of switches:

(a) two open switches for transistor T_2;

(b) two closed circuits for transistor T_1.

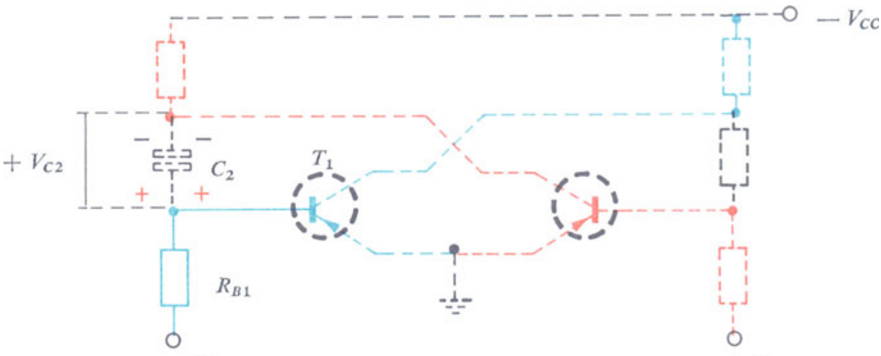


Fig. 478

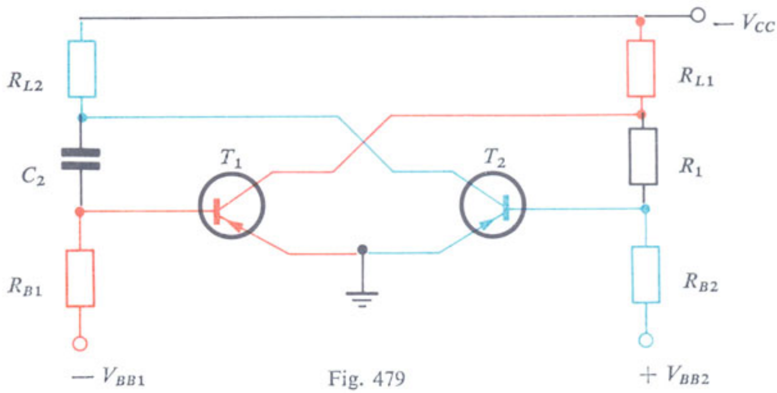


Fig. 479

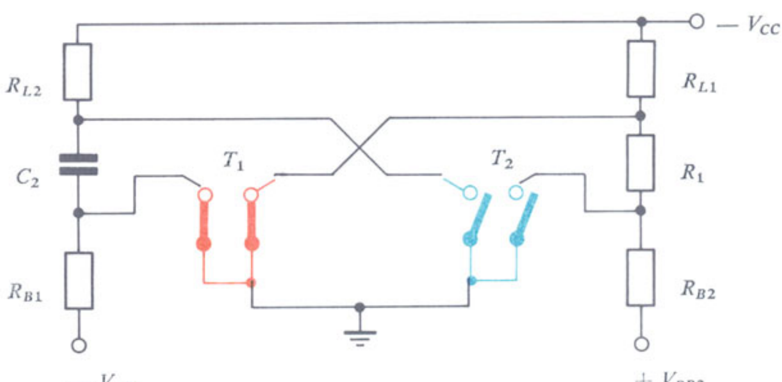
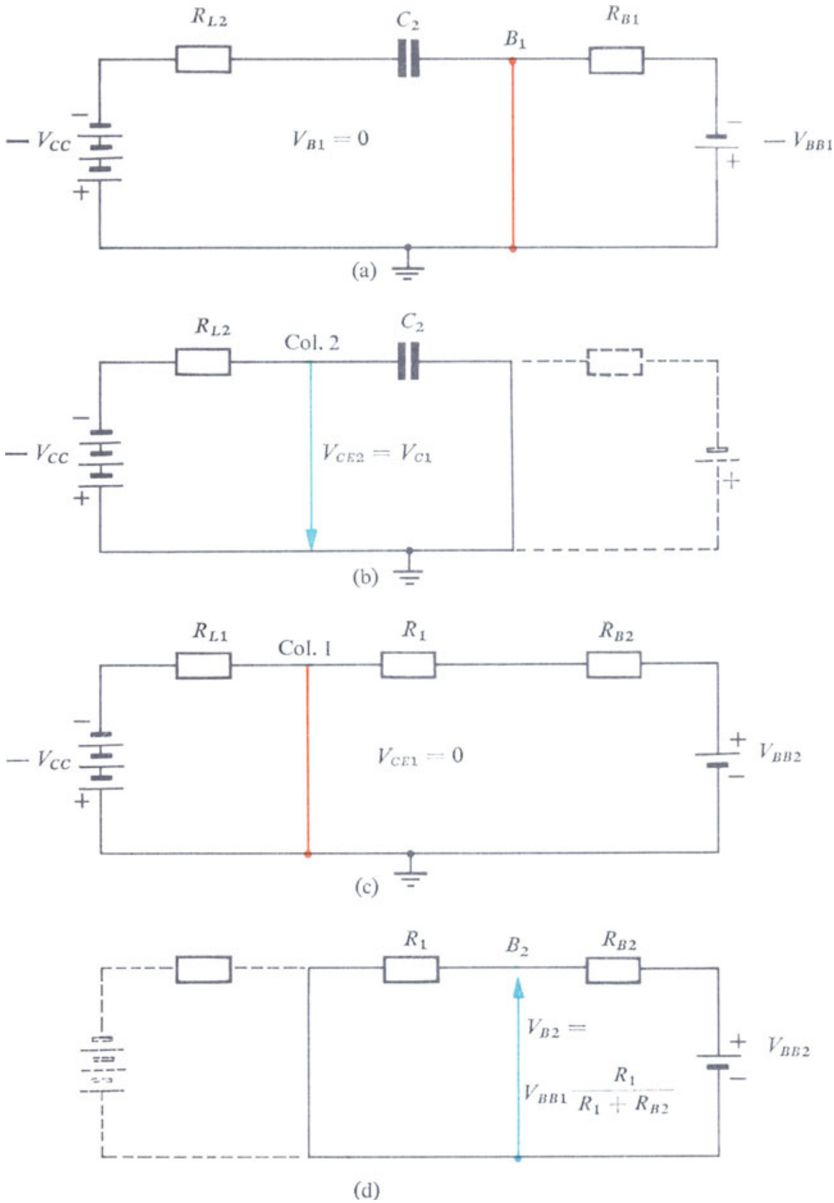


Fig. 480



R_{L2}
C_2
B_1
R_{B1}
$-V_{CC}$
$V_{B1} = 0$
$-V_{BB1}$
(a)
R_{L2}
Col. 2
C_2
$-V_{CC}$
$V_{CE2} = V_{C1}$
(b)
R_{L1}
Col. 1
R_1
R_{B2}
$-V_{CC}$
$V_{CE1} = 0$
V_{BB2}
(c)
R_1
B_2
R_{B2}
$V_{B2} = V_{BB1}\dfrac{R_1}{R_1 + R_{B2}}$
V_{BB2}
(d)

Fig. 481

The voltages and currents may be determined quantitatively by means of a simple mesh presentation as in Figs. 481a, b, c and d. The various electrical parameters to be calculated in connection with this class of circuit should be remembered. There are four active electrodes:

(a) the base of transistor T_1;
(b) the collector of transistor T_1;
(c) the base of transistor T_2;
(d) the collector of transistor T_2.

The voltage and current must be calculated for each of these electrodes.

Calculation of the voltages

There are four of these voltages: the base–emitter voltage of T_1, the collector–emitter voltage of T_1, the base–emitter voltage of T_2, and the collector–emitter voltage of T_2.

Base–emitter voltage of T_1. It will be seen in Fig. 481a that the switch representing the base–emitter junction of transistor T_1 is closed; in these conditions the voltage at its terminals is zero:

$$V_{BE1} = 0.$$

Collector–emitter voltage of transistor T_1. The switch representing the output circuit of transistor T_1 is also closed (Fig. 481c) and the voltage is zero:

$$V_{CE1} = 0.$$

Base–emitter voltage of transistor T_2. Let us consider Fig. 481d. Since the base–emitter switch of transistor T_2 was open, a voltage due to the battery $+V_{BB2}$ appears at the terminals of resistor R_1. This voltage is equal to the product of the resistance and the current flowing through it, namely:

$$V_{BE2} = V_{BB2} \frac{R_1}{R_1 + R_{B2}}.$$

Collector–emitter voltage of transistor T_2. In Fig. 481b the 'collector–emitter' switch of transistor T_2 is open. The collector–emitter voltage is, therefore, equal to the charging voltage of capacitor C_2. In continuous operation the capacitor is fully charged and the value of the voltage is the same as that of the supply.

$$V_{CE2} = -V_{CC}.$$

Calculation of the currents

A current is associated with each electrode, and it is, therefore, necessary to calculate four currents: the base current of T_1, the collector current of T_1, the base current of T_2 and the collector current of T_2.

Base current of transistor T_1. The 'base–emitter' switch of transistor T_1 is closed (Fig. 482a), and a current flows in the circuit. This current is a function of the voltage applied to its terminals and of the resistor connected in the circuit:

$$-I_{B1} = -V_{BB1}/R_{B1}.$$

Collector current of the transistor T_1. In Fig. 482b the switch representing the output circuit of T_1 is closed; the collector current, therefore, comes from the group of supply batteries $-V_{CC}$ and varies inversely with the resistance R_{L1}:

$$-I_{C1} = -V_{CC}/R_{L1}.$$

Base current of transistor T_2. As this transistor is cut-off, the switch representing the input circuit is open, resulting in a zero value of current:

$$I_{B2} = 0.$$

Collector current of transistor T_2. The 'collector–emitter' switch of transistor T_2 being also open, the corresponding current is 0:

$$I_{C2} = 0.$$

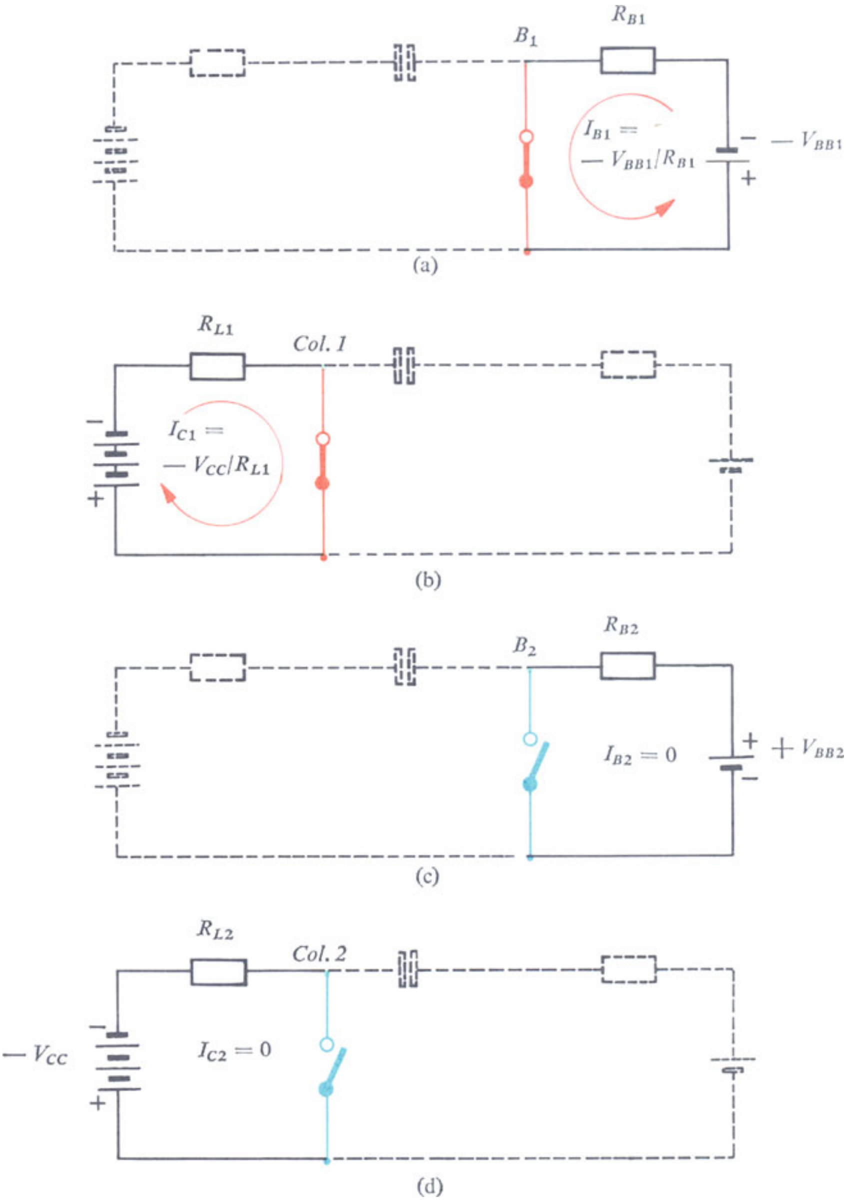
Summary of voltages and currents.

	transistor T_1	transistor T_2
V_{CE}	0	$-V_{CC}$
V_{BE}	0	$V_{BB2}\dfrac{R_1}{R_1+R_{B2}}$
I_C	$-V_{CC}/R_{L1}$	0
I_B	$-V_{BB1}/R_{B1}$	0

Checking of the conditions of saturation and cut-off

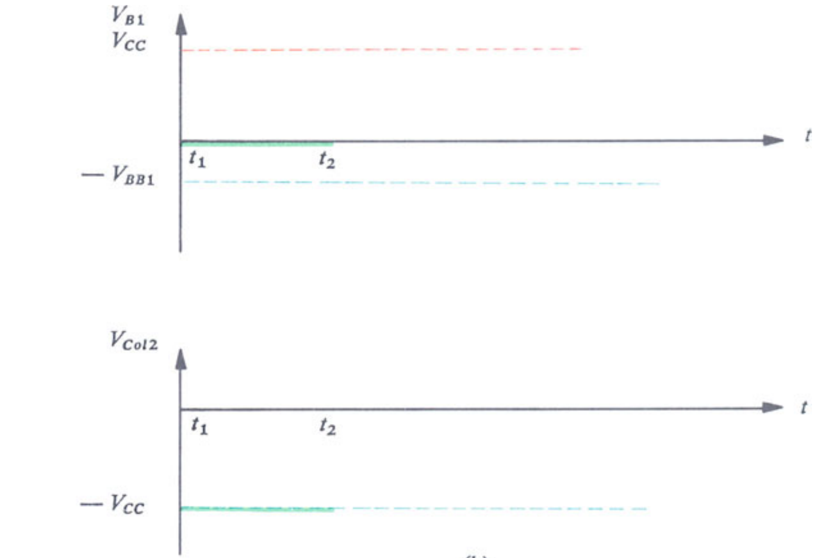
If we know the values of the collector and base currents of transistor T_1 when saturated, we can now ensure that this condition is respected by applying the following formula:

$$-I_{C(T_1)}/-I_{B(T_1)} < h_{FE(T_1)}.$$

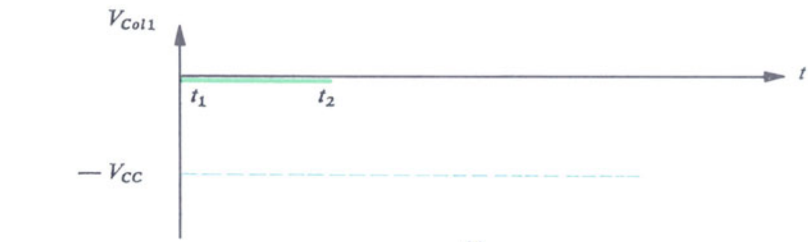


R_{B1}
B_1
$I_{B1} =$
$-V_{BB1}/R_{B1}$
$-V_{BB1}$
(a)
R_{L1}
Col. 1
$I_{C1} =$
$-V_{CC}/R_{L1}$
(b)
R_{B2}
B_2
$I_{B2} = 0$
$+V_{BB2}$
(c)
R_{L2}
Col. 2
$-V_{CC}$
$I_{C2} = 0$
(d)

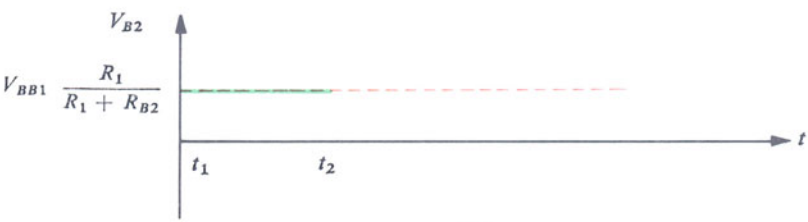
Fig. 482


V_{B1}
V_{CC}
t
t_1
t_2
$-V_{BB1}$
V_{Col2}
t
t_1
t_2
$-V_{CC}$

(b)


V_{Col1}
t
t_1
t_2
$-V_{CC}$

(c)


V_{B2}
$V_{BB1}\ \dfrac{R_1}{R_1 + R_{B2}}$
t
t_1
t_2

(d)

Fig. 483

It is very simple to check the cut-off state of transistor T_2: all that is needed is that the base should be positive with respect to the collector and emitter. Since V_{BE2} is equal to $V_{BB2}R_1/(R_1+R_{B2})$ and V_{CE2} is equal to $-V_{CC}$, transistor T_2 is in the cut-off state.

8.3. Physical explanation of the operator with diagrammatic presentation of the signals

Just as for the monostable multivibrator, we will attempt to define the voltage variations between the various active points of the circuit and earth.

Figure 483a shows the variations with time of the base–emitter voltage of transistor T_1. The limiting values of this voltage are:

(a) positively, the value of the charging voltage of capacitor C_2: $+V_{CC}$;

(b) negatively, the voltage due to the bias battery $-V_{BB1}$.

Figure 483b shows the variations of the collector–emitter voltage of transistor T_1. The limiting values of this voltage are:

(a) 0 when the transistor is saturated;

(b) $-V_{CC}$ when cut-off.

The variations of the collector–emitter voltage of transistor T_2 may be determined with the aid of Fig. 483c. This voltage also varies between the two limiting values 0 and $-V_{CC}$.

Finally, Fig. 483d shows the variations of the base–emitter voltage of transistor T_2. The maximum positive value is a function of the bias voltage V_{BB2} and is given by the equation:

$$V_{BE2} = V_{BB2}\,\frac{R_1}{R_1+R_{B2}}.$$

When transistor T_2 is conducting (the input circuit being likened to an ideal switch in the closed position) the base–emitter voltage of this transistor is assumed to be near to zero.

Representation of the stable state. From instant t_1 to the instant preceding t_2, no external signal is applied to the bases of the transistors, and the multivibrator is in its stable position. It will be seen at once from the table on the previous page that:

(a) the collector–emitter voltage of T_1 is zero; it is represented by a horizontal line coincident with the abscissa;

(b) the base–emitter voltage of T_1 is also equal to zero; it is represented by the line on the abscissa;

(c) the collector–emitter voltage of transistor T_2 is equal to $-V_{CC}$;

(d) the base–emitter voltage of transistor T_2 is positive and equal to:

$$V_{BB2}\,\frac{R_1}{R_1+R_{B2}}.$$

Definition of the dynamic period. The application of a negative pulse between the base and emitter of transistor T_2 (Fig. 484d) causes the appearance of a current in the collector of this transistor and a decrease in the negative voltage previously present between this point and earth (Fig. 484c). Any variation in either direction between the collector and emitter of transistor T_2 is at once retransmitted between the base and emitter of transistor T_1.

In these conditions at instant t_2 the base–emitter voltage of transistor T_1 becomes positive until it reaches the value $+V_{CC}$ which corresponds to the variation of the collector–emitter voltage of T_2.

The state of transistor T_1 changes from saturation to cut-off, and its collector–emitter voltage (which was previously zero) increases and approaches the value $-V_{CC}$ in accordance with the charging curve of capacitor C_2 through resistor R_{L2} (Fig. 484c).

The cut-off period of transistor T_1 depends on the discharge time of capacitor C_2 through resistor R_{B1}; this voltage variation is shown in Fig. 484a which also gives instant t_3 at which transistor T_1 again becomes conducting.

From instant t_2 to instant t_3, the collector–emitter voltage of transistor T_1 changes from zero to the value $-V_{CC}$ and remains there (Fig. 484c).

The collector–emitter voltage of transistor T_2 remains at zero during the whole of this period (Fig. 484b).

The base–emitter voltage of this transistor remains at zero from t_2 to t_3 (Fig. 484d).

From instant t_3 transistor T_1 becomes saturated and its collector–emitter voltage varies from $-V_{CC}$ to 0 (Fig. 484c).

This positive voltage variation is applied at once between the base and emitter of transistor T_2 (Fig. 484d), which becomes cut-off.

The collector–emitter voltage of transistor T_2 increases from 0 to the value $-V_{CC}$ (Fig. 484b).

In the absence of any further information the voltages remain at their various values associated with the steady operating state. If a second pulse is applied at instant t_4 the same process is repeated.

It will be seen at once from Figs 484a, b, c and d that for a pulse applied at the input, a pulse appears at the output circuit (between the collector and emitter of the transistors).

This explanation could be repeated on the basis of a positive pulse applied between the base and emitter of transistor T_1.

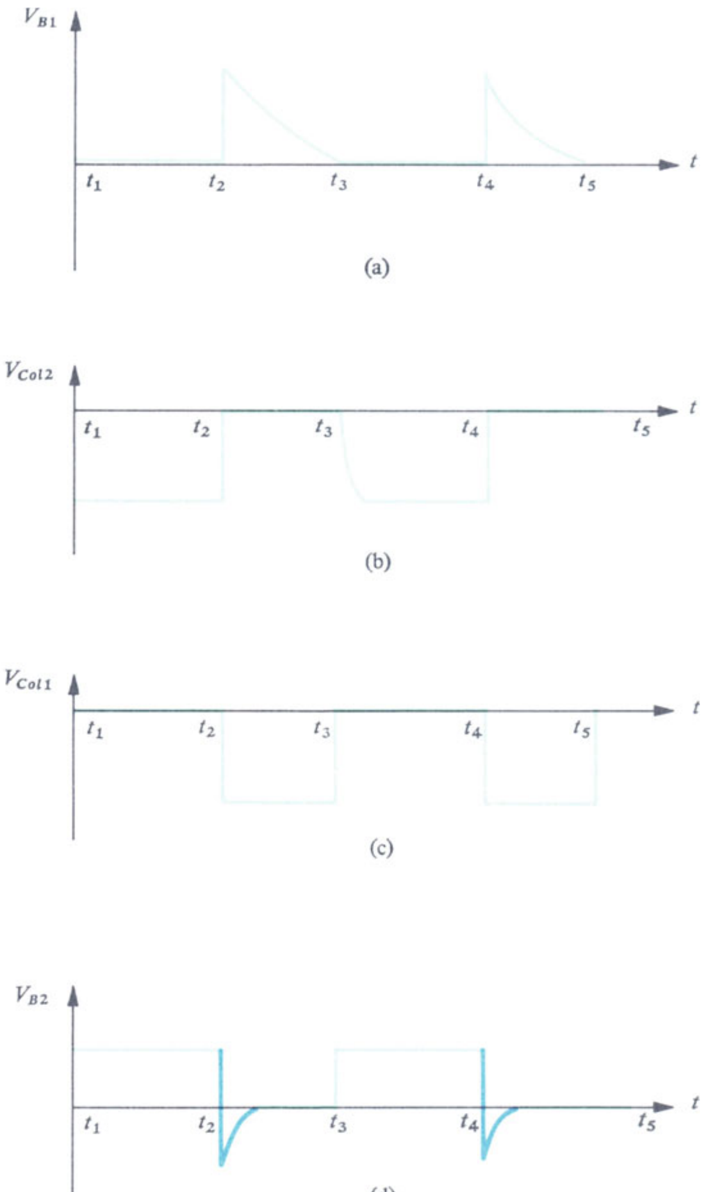

V_{B1}
t
t_1
t_2
t_3
t_4
t_5
(a)
V_{Col2}
t
t_1
t_2
t_3
t_4
t_5
(b)
V_{Col1}
t
t_1
t_2
t_3
t_4
t_5
(c)
V_{B2}
t
t_1
t_2
t_3
t_4
t_5
(d)

Fig. 484

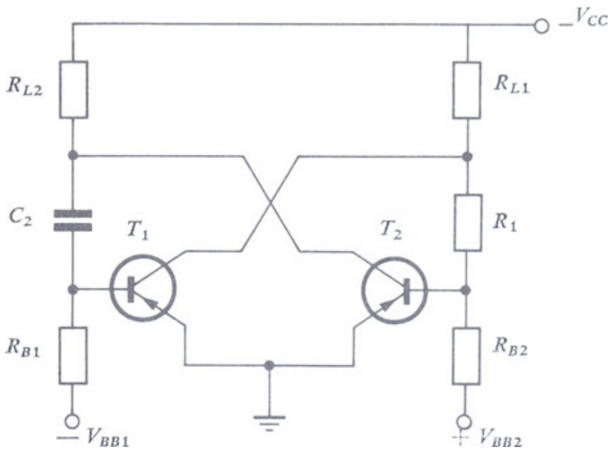


Fig. 485

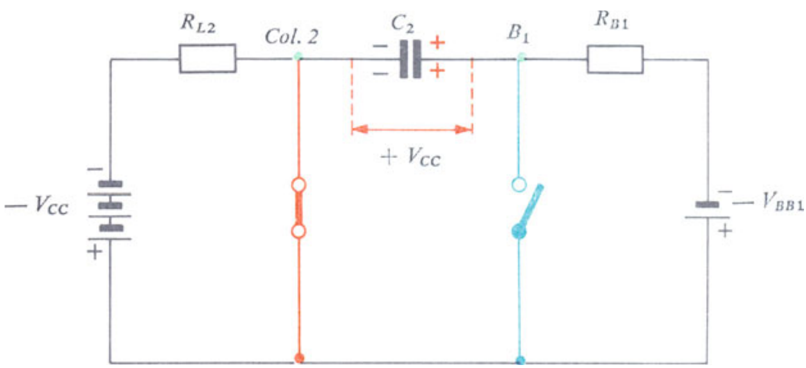


Fig. 486

8.4. Calculation of the voltages and currents during the dynamic period

During the triggering period it is necessary for us to know the variations with time of the voltage and current in this circuit (Fig. 485).

This can be done readily if we extract from the overall circuit a certain number of circuits especially concerned with each electrode for which the values of voltage and current must be calculated.

Figure 486 shows the group of components of this stage that occur in the expression for the base–emitter voltage and the base current of transistor T_1 as well as in the collector–emitter voltage and collector current of transistor T_2.

Base–emitter voltage of transistor T_1. At instant t_2 capacitor C_2 is charged by the voltage V_{CC} which has made the base positive with respect to this voltage. The capacitor discharges through resistor R_{B1} and the base–emitter voltage of this transistor varies in accordance with the same function, namely:

$$V_{BE1} = -V_{BB1} + (V_{CC} + V_{BB1}) \exp\left(-\frac{t_3 - t_2}{C_2 R_{B1}}\right).$$

Base current of transistor T_1. Since this transistor is cut-off, the base–emitter circuit is open, and the current is:

$$I_{B1} = 0.$$

Collector–emitter voltage of transistor T_2. Transistor T_2 is saturated and its output resistance is near to 0. This point in Fig. 486 is connected directly to earth, giving:

$$V_{CE2} = 0.$$

Collector–current of transistor T_2. During the transient period, that is from instant t_2 to instant t_3, the collector current of transistor T_2 is equal to the difference of the two currents:

(a) the current $i_{1(t)}$, or the discharge current of capacitor C_2 through resistor R_{B1} and the battery $-V_{BB1}$;

(b) the current i_2 due to the group of supply batteries $-V_{CC}$ in resistor R_{L2}.

$$i_{1(t_3-t_2)} = \frac{-V_{BB1} + V_{BE1(t_3-t_2)}}{R_{B1}} \quad \text{and} \quad i_2 = \frac{-V_{CC}}{R_{L2}}.$$

For these conditions $I_{C2(t3-t2)}$ is given by the relation:

$$I_{C2(t_3-t_2)} = \frac{-V_{CC}}{R_{L2}} - \frac{V_{BB1} + V_{BE1(t_3-t_2)}}{R_{B1}}.$$

With the aid of the circuit in Fig. 487 we can now construct a simplified diagram for the calculation of the collector–emitter voltage, the collector current of transistor T_1 and the base–emitter voltage and base current of transistor T_2.

Collector–emitter voltage of transistor T_1. From instant t_2 to instant t_3 transistor T_1 is in the cut-off state. Its collector–emitter voltage is due to the group of supply batteries and equal to:

$$V_{CE1} = -V_{CC}\frac{R_1}{R_1+R_{L1}}.$$

Collector current of transistor T_1. The cut-off of this transistor mentioned above involves the total absence of current in the circuit:

$$-I_{C1} = 0.$$

Base–emitter voltage of transistor T_2. The saturation state of this transistor corresponds to a value of input resistance near to 0, resulting in a zero voltage between the base and emitter:

$$V_{BE2} = 0.$$

Base current of transistor T_2. The closed position of the switch shows that two currents flow through the base circuit of this transistor:

(a) the current i_3 dependent on the supply voltage $-V_{CC}$ and the two resistors R_{L1} and R_1;

(b) the current i_4 which is a function of the bias voltage $+V_{BB2}$ and the resistor R_{B2}:

$$I_{B2} = I_3 - I_4 = \frac{-V_{CC}}{R_{L1}+R_1} - \frac{V_{BB2}}{R_{B2}}.$$

Summary of the voltages and currents in dynamic operation.

	transistor T_1 (t_3-t_2)	transistor T_2 (t_3-t_2)
V_{CE}	$-V_{CC}\frac{R_1}{R_1+R_{L1}}$	0
V_{BE}	$-V_{BB1}+(V_{CC}+V_{BB1}) \exp\left(-\frac{t_3-t_2}{R_{B1}\cdot C_2}\right)$	0
I_C	0	$\frac{-V_{CC}}{R_{L2}} - \frac{-V_{BB1}+V_{BE1(t_3-t_2)}}{R_{B1}}$
I_B	0	$\frac{-V_{CC}}{R_{L1}+R_1} - \frac{V_{BB2}}{R_{B2}}$

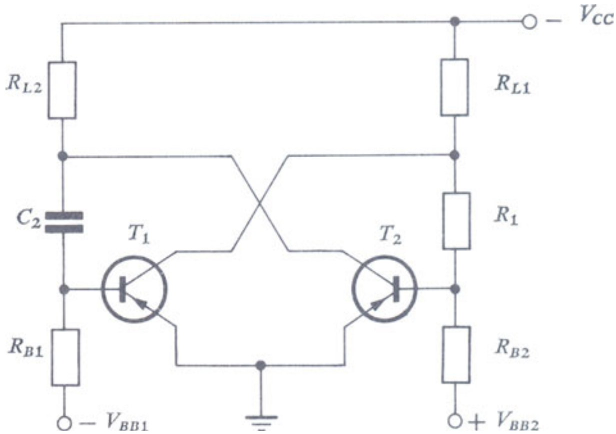


Fig. 487

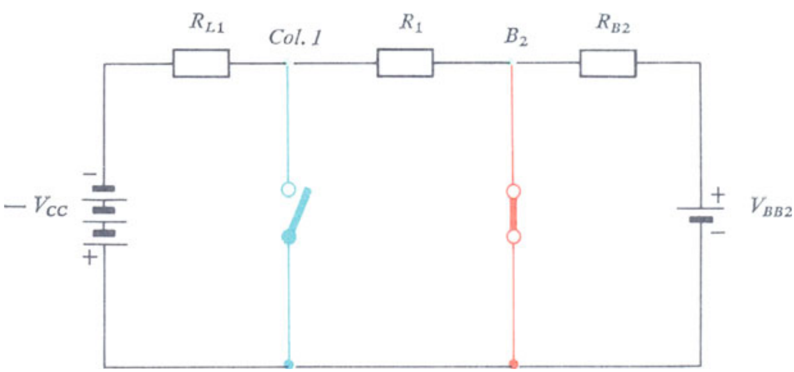


Fig. 488

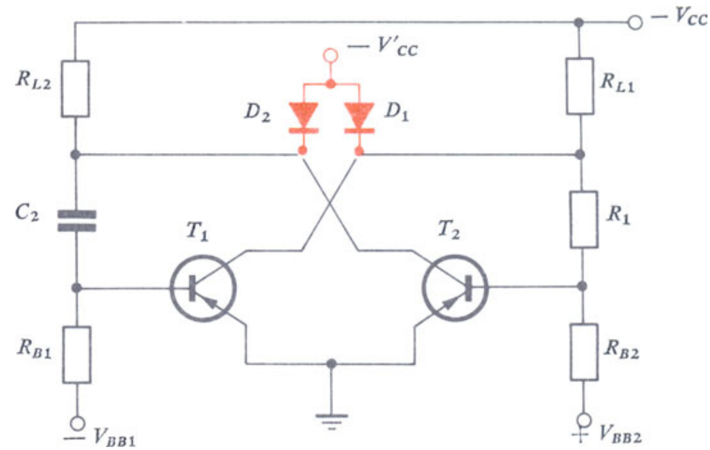


Fig. 489

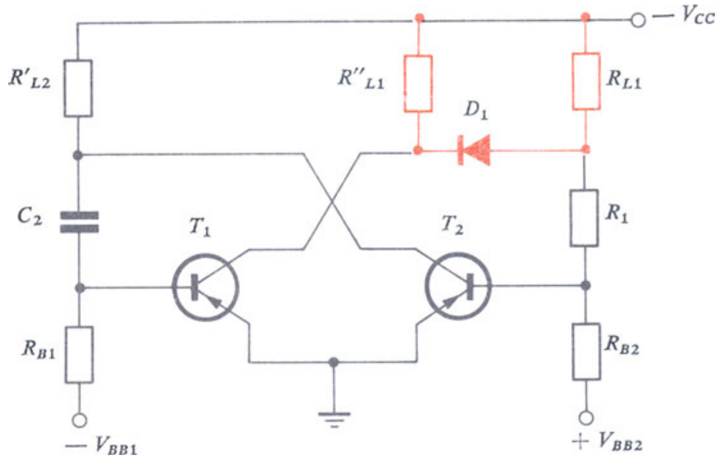


Fig. 490

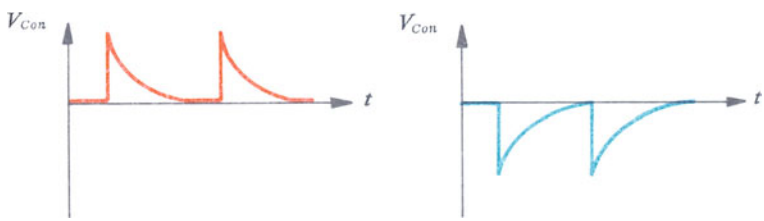


Fig. 491

8.5. Correction circuit

Just as for the astable multivibrator, it may be desirable to have a diode correction circuit in order to improve the shape of the signal received at the output circuit.

In the first system an auxiliary bias source $-V_{CC}'$ is used; (it gives a voltage lower than the supply voltage $-V_{CC}$). This source is connected to the collector of the transistors by means of two diodes (Fig. 489).

The principle of operation of a system of this nature was discussed when we were dealing with the astable multivibrator, and it is, therefore, not necessary to repeat it. It should be remembered, however, that it has the major disadvantage that it reduces the amplitude of the signal available at the output.

In the second system the loads of the transistor in the static state are divided by means of diode circuits (Fig. 490). This allows a considerable improvement in the leading edge of the output pulse.

8.6. Problems concerned with the triggering of a Monostable multivibrator

The operation of the circuit was explained earlier by means of the application of a negative voltage pulse between the base and emitter of transistor T_2, or a positive voltage pulse between the base and emitter of transistor T_1.

A system of this sort can be triggered either by:

(a) successive positive pulses (in red in Fig. 491); or

(b) negative pulses (in blue in the same figure).

The electrical coupling circuit between the source of pulses (pulse generator, the previous stage, etc.) and the input of the transistor concerned may consist simply of a capacitor. Since the pulses are always in the same direction there is no problem in their operation with respect to the inputs of the two transistors.

In rapid switching it is sometimes necessary to use more complex coupling systems, especially those with diodes; but the principle of operation of a circuit of this nature does not differ from that described earlier.

Notes on the shape and duration of the control pulses

The change of state of a transistor implies that a pulse of sufficient amplitude has been applied to the terminals of the input circuit. If we remember that at instant t_2 the transistor is in an extreme operating state and that this state is associated with a delay time for the variation of the current (t_d for the transistor in the cut-off state and t_s for the transistor when saturated) it is necessary that pulses with a duration greater than these delay times should be available. This situation is shown in Figs. 492a and b.

In the first a negative voltage pulse reaches transistor T_2 which is cut-off; the duration of this pulse is longer than the delay time of the current in the circuit:

$$t_p > t_d.$$

In the second, a positive voltage pulse reaches the base of transistor T_1 which is saturated; the duration of this pulse is longer than the desaturation time of the transistor:

$$t_p > t_s.$$

Figure 493 illustrates the general shape of the pulses transmitted in circuits of this type. They first cause a very sudden variation of the voltage in the input circuit, and then decrease rather slowly.

Let us consider the case of the control of a transistor in the cut-off state driven by a negative voltage pulse (Fig. 494).

At instant t_2 the large variation of the negative voltage produced in the input circuit of transistor T_2 by the control pulse tends to cause the transistor to become saturated. The corresponding collector current can only start to vary from instant t_2'; $t_2' - t_2$ constitutes the limits of the delay time of the circuit.

It is therefore necessary that at instant t_2' the control signal should have an amplitude still large enough to allow the collector current to increase and transistor T_2 to become saturated.

The same reasoning applies when transistor T_1 is driven by a positive pulse which switches it from saturation to cut-off (Fig. 495).

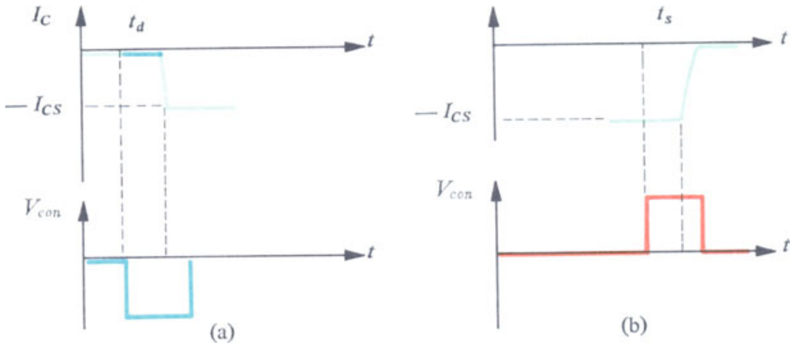


Fig. 492

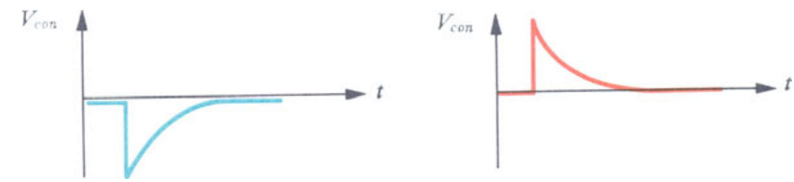


Fig. 493

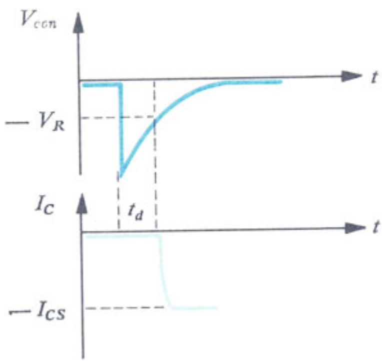


Fig. 494

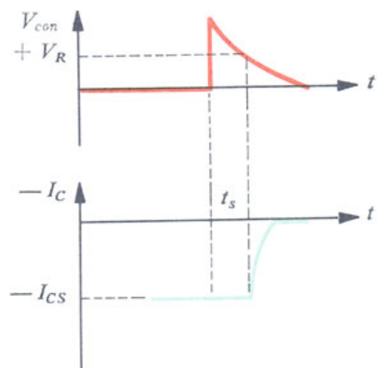


Fig. 495

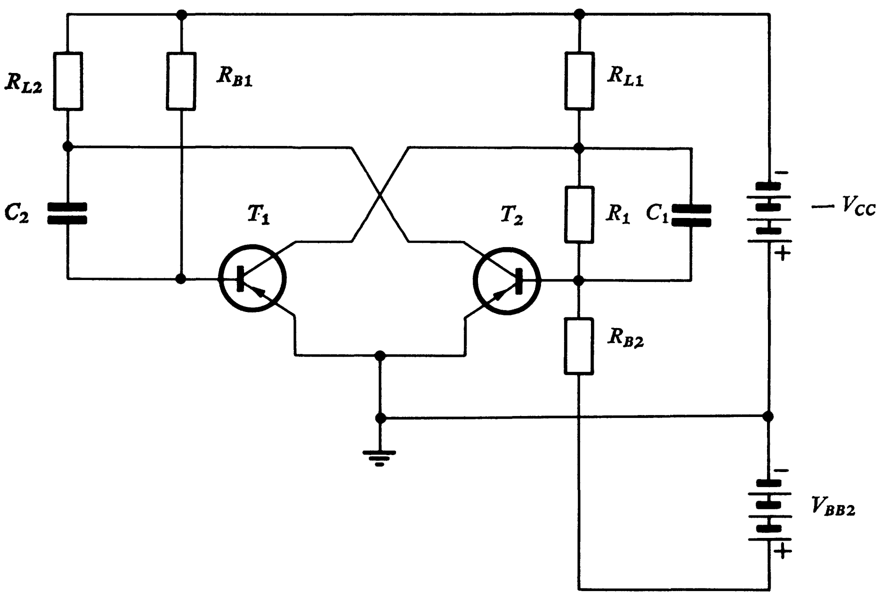


Fig. 496

8.7. An actual circuit

Just as for the astable multivibrator, we could discuss a large number of monostable multivibrators. We will revert to the basic diagram studied earlier, and consider the multivibrator in Fig. 496.

It consists of two transistors of the same type loaded by 1 kΩ resistors. and biased:

(a) for the first transistor (T_1) by a resistor of 18 kΩ connected to the supply $-V_{CC}$;

(b) for the other transistor by a resistor of 33 kΩ connected to the positive electrode of another group of batteries $+V_{BB}$.

The coupling circuits are:

(a) in the first case, capacitative ($C_2 = 10$ nF);

(b) in the second case, resistive ($R_1 = 5{\cdot}6$ kΩ). This resistor is often associated with a capacitor with a value near to 500 pF and acting as an acceleration capacitance.

In a circuit of this nature transistor T_1 is permanently saturated and transistor T_2 cut-off.

We need not repeat the principle of operation of this stage; on the other hand it would be useful to associate coupling systems with it so as to be able to transmit the control pulses to it.

As a function of the type of application the coupling may be connected:

(a) either at the input of transistor T_1 for positive pulses;

(b) or at the input of transistor T_2 for negative pulses.

The coupling system may consist of a simple capacitor, but preference is given to the semiconductor diode.

The operation of a diode coupling system will be studied in great detail when we are discussing the bistable multivibrator.

9. The Bistable Multivibrator

The basic function in logic circuits is the division of the number of pulses, and this role is undertaken by the bistable multivibrator, as shown in Fig. 497. The bistable multivibrator shown synoptically supplies a single pulse at its output for every two pulses applied across the input.

A circuit of this type has two stable states; these can therefore be studied separately from the viewpoint of their purely static behaviour.

Just as for astable and monostable multivibrators, we will discuss the various aspects in the following order:

(a) concept and principle of operation;
(b) calculation of the values of voltage and current for the two stable states;
(c) diagrammatic representation of the signals;
(d) diagrammatic method of calculating the stable states;
(e) transient effects during the triggering periods;
(f) problems concerning the triggering of the coupling systems used;
(g) different types of circuit.

9.1. Concept and principle of operation of a bistable multivibrator

The basic diagrams for the two previous types of multivibrator, astable and monostable, provide the basis for the discussion of the circuit shown in Fig. 498.

This stage consists of two transistors T_1 and T_2. Because of the symmetry nearly always met with in this type of multivibrator the two resistive chains consist of: the load resistors, and the coupling and the bias resistors which are identical; namely R_L, R and R_B.

The two transistors are supplied by the same group of batteries $-V_{CC}$; and are biased by the battery $+V_{BB}$.

In the absence of pulses applied at the input of the stage this multivibrator is in a very definite operating state, namely:

(a) transistor T_1 saturated;
(b) transistor T_2 cut-off.

The state of saturation of transistor T_1 involves a negative value of voltage between its base and emitter; in fact, ignoring the value of the input resistance is equivalent to a comparison of the base–emitter circuit with a short circuit, giving a value near to zero for this voltage. The same applies to the collector–emitter voltage.

Transistor T_2 is cut-off, and its base is positive with respect to the emitter and collector.

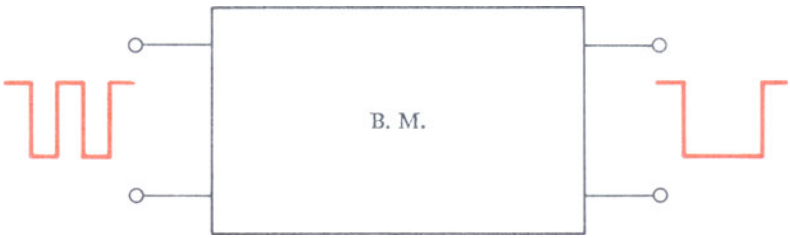


Fig. 497

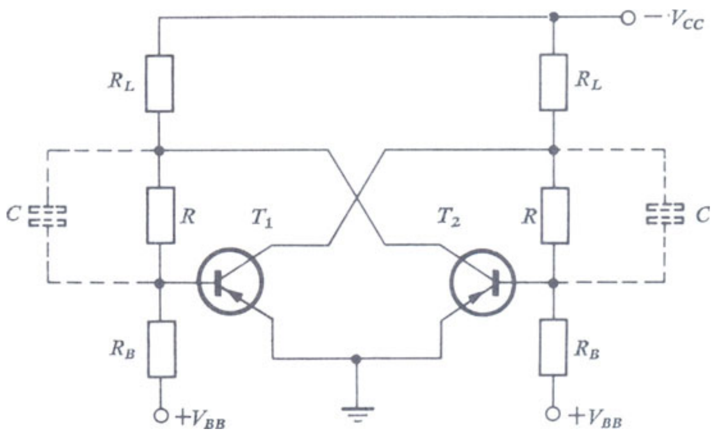


Fig. 498

Influence of a pulse

There are two possible methods for the control of a bistable multivibrator: either to apply a positive pulse between the base and emitter of the saturated transistor (T_1); or to apply a negative voltage pulse between the base and emitter of the transistor in the cut-off state (T_2).

Case of a pulse between the base and emitter of the saturated transistor

If a negative pulse is applied at a given instant between the base and emitter of transistor T_1 (when saturated) it tends to reduce the collector current (Fig. 499); the collector–emitter voltage changes from zero to a negative value which is immediately transmitted to the base of transistor

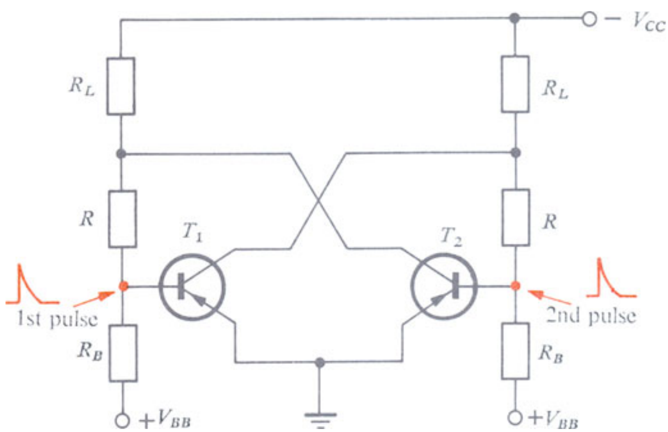


Fig. 499

T_2. This transistor starts to conduct, resulting in a decrease in its collector–emitter voltage. This positive variation boosts the effect of the exciting pulse on transistor T_1 which is consequently driven more and more rapidly into non-conduction.

This process continues, and a new stable state appears in which transistor T_1 is cut-off and transistor T_2 is in the saturated state.

The multivibrator is now in its second stable state. The input and output resistance of transistor T_2 are then practically zero, whilst the input and output resistance of transistor T_1 are very large.

If no pulse is applied to the input terminals of the circuit, the multivibrator remains permanently in this new operating state.

Influence of a second positive pulse between the base and emitter of the saturated transistor

We will now assume a second positive pulse, with a coupling system such that this pulse is applied between the base and emitter of transistor T_2 which is then saturated.

The positive voltage appearing between the base and emitter of this transistor causes a decrease in the collector current, which in turn causes an increase in the collector–emitter voltage, that is a negative voltage variation applied at once between the base and emitter of transistor T_1 which is cut-off.

Transistor T_1 starts to conduct and its collector–emitter voltage becomes less negative. The resulting positive variation is instantaneously transmitted to the base of transistor T_2, where it strengthens the action of the exciting pulse. Transistor T_1 becomes saturated and transistor T_2 cut-off. This second operating state is maintained until the arrival of a third positive pulse.

If we make an evaluation of the signals received at the output of the circuit with respect to the pulses applied at its input terminals, we see in Fig. 500 that for two input pulses a single pulse only is transmitted. The 'counting' function is thus respected and binary operations can therefore easily be performed.

9.2. Calculation of the values of the voltages and currents for the two stable states

A multivibrator of this nature may be in either of two stable states; we will consider the first of these states as an example:

(a) T_1 saturated;

(b) T_2 cut-off.

First stable state

By means of the complete circuit diagram for the bistable multivibrator (Fig. 501) we can derive two simplified equivalent circuits for the easy calculation of the values of the voltages and currents in each of the two active electrodes of the chain of resistances on the left.

In Fig. 502 a group of supply batteries $-V_{CC}$ is connected directly across the terminals of a load resistance R_L and the bias battery $+V_{BB}$ is connected across resistor R_B.

The particulars to be calculated are:

(a) the collector–emitter current of transistor T_2;

(b) the collector current of transistor T_2;

(c) the base–emitter current of transistor T_1;

(d) the base current of transistor T_1.

Collector–emitter voltage of transistor T_2. Transistor T_2 is in the cut-off state and its collector is connected to earth by the coupling resistor R. The collector–emitter voltage of this transistor is therefore equal to the voltage across resistor R, namely:

$$V_{CE2} = -V_{CC}\,\frac{R}{R+R_L}.$$

Collector current of transistor T_2. The cut-off state involves the total absence of current in the collector of the transistor concerned:

$$-I_{C2} = 0.$$

Base–emitter voltage of transistor T_1. Transistor T_1 is saturated; its base is connected directly to earth and the base current is a function of the group of currents flowing through this part of the circuit:

$$-V_{BE1} = 0.$$

Base current of transistor T_1. It will be seen from Fig. 503 that the base current is equal to the difference of two currents:

(a) a current I_1 due to the group of batteries $-V_{CC}$ in resistors R_L and R;

(b) a current I_2 due to the battery $+V_{BB}$ in resistor R_B.

The values of these two currents are respectively:

$$i_1 = \frac{-V_{CC}}{R+R_L} \quad \text{and} \quad i_2 = \frac{V_{BB}}{R_B}$$

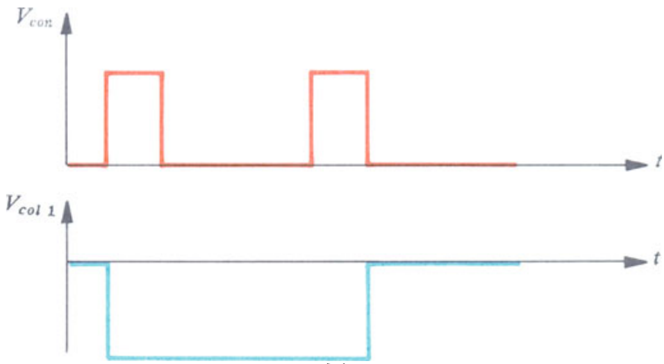


Fig. 500

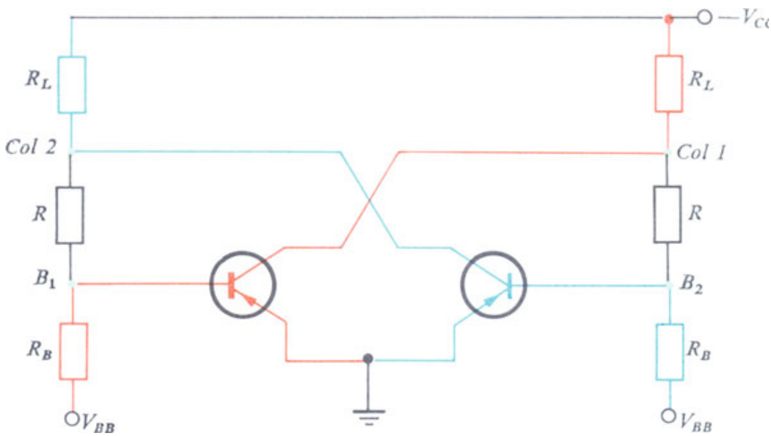


Fig. 501

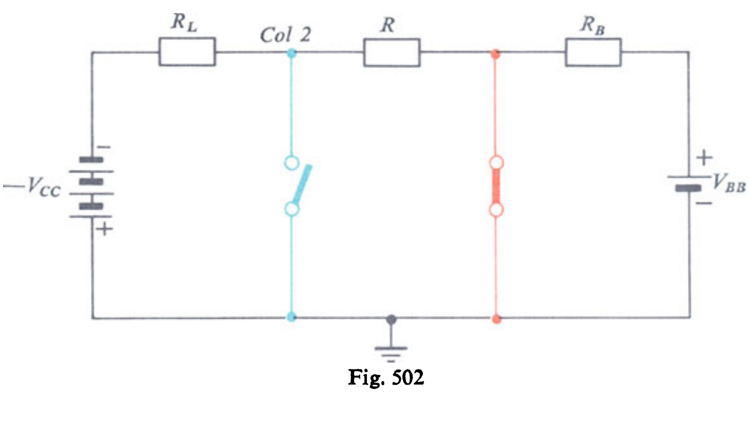


Fig. 502

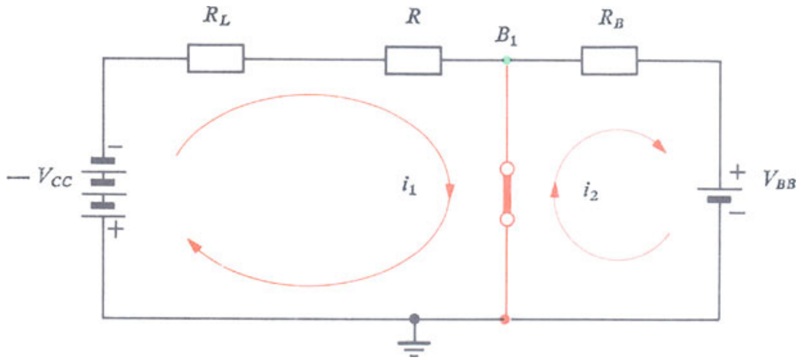


Fig. 503

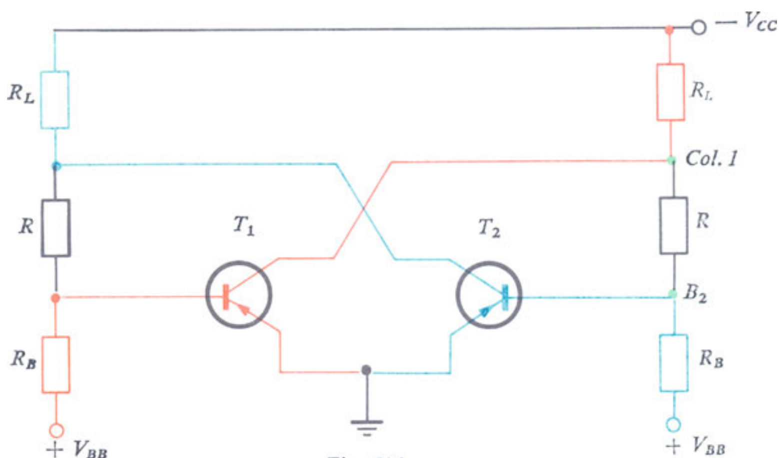


Fig. 504

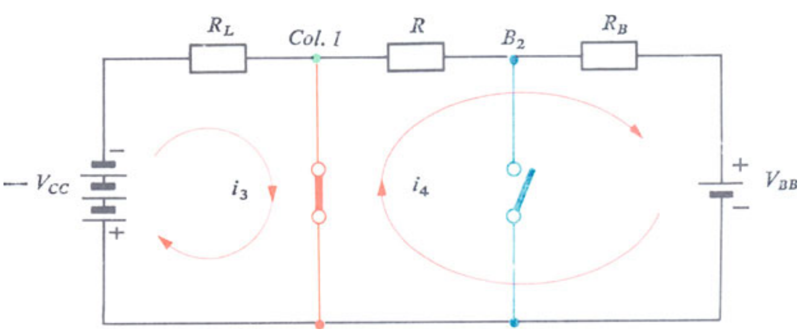


Fig. 505

and since $I_{B1} = i_1 - i_2$ $\qquad I_{B1} = \frac{-V_{CC}}{R+R_L} - \frac{V_{BB}}{R_B}.$

The values of the voltages and currents thus calculated remain constant until the arrival of another pulse.

Calculation of the currents and voltages in the chain of resistors on the right

The reader is referred back to the complete circuit of the bistable multivibrator in Fig. 504 and to the equivalent circuit of the chain of resistors on the right of the diagram, shown in Fig. 505.

The group of supply batteries $-V_{CC}$ is connected directly across the load resistance R_L, and the bias battery $+V_{BB}$ across a group of resistors in series: R_B and R. The particulars to be calculated are in the order:

(a) the collector–emitter voltage of transistor T_1;
(b) the collector current of transistor T_1;
(c) the base–emitter voltage of transistor T_2;
(d) the base current of transistor T_2.

Collector–emitter voltage of transistor T_1. Transistor T_1 is saturated, its output resistance is near to 0 and its collector–emitter voltage is zero:

$$-V_{CE1} = 0.$$

Collector current of transistor T_1. It will be seen from the simplified circuit of Fig. 505 that the collector current of T_1 is equal to the difference of the two currents:

(a) a current i_3 due to the group of batteries $-V_{CC}$ in resistor R_L;
(b) a current i_4 due to the battery $-V_{BB}$ in the resistors R_B and R.

The values of these two currents are:

$$i_3 = \frac{-V_{CC}}{R_L} \quad \text{and} \quad i_4 = \frac{V_{BB}}{R+R_B}$$

and since $I_{C1} = i_3 - i_4$

$$I_{C1} = -\frac{V_{CC}}{R_L} - \frac{V_{BB}}{R+R_B}.$$

Base–emitter voltage of transistor T_2. Transistor T_2 is cut-off, its input resistance is near to infinity and its base–emitter voltage is a function of the voltage drop at the terminals of resistor R. Only the bias battery V_{BB} occurs in the expression for the current flowing through the circuit (Fig. 506) and the voltage is, therefore, positive and equal to:

$$V_{BE2} = V_{BB} \frac{R}{R+R_B}.$$

Base current of transistor T_2. Since the input resistance of transistor T_2 is very high practically no current can flow and:

$$I_{B2} = 0.$$

Just as for the voltages and currents shown on the left of the diagram for the bistable multivibrator, these voltages and currents remain at the same value until the arrival of another pulse.

Second stable state

The same reasoning could be adopted for the calculation of the values of the voltages and currents corresponding to the new operating state of the multivibrator, that is:

(a) T_1 cut-off;

(b) T_2 saturated.

Figures 507 and 508 show the simplified diagrams by means of which the calculations may be performed.

Collector–emitter voltage of transistor T_2

$$V_{CE2} = 0.$$

Collector current of transistor T_2

$$I_{C2} = \frac{V_{CC}}{R_L} - \frac{V_{BB}}{R_B+R}.$$

Base–emitter voltage of transistor T_1

$$V_{BE1} = V_{BB} \cdot \frac{R}{R+R_B}.$$

Base current of transistor T_1

$$I_{B1} = 0.$$

Collector–emitter voltage of transistor T_1

$$V_{CE1} = -V_{CC} \cdot \frac{R}{R+R_L}.$$

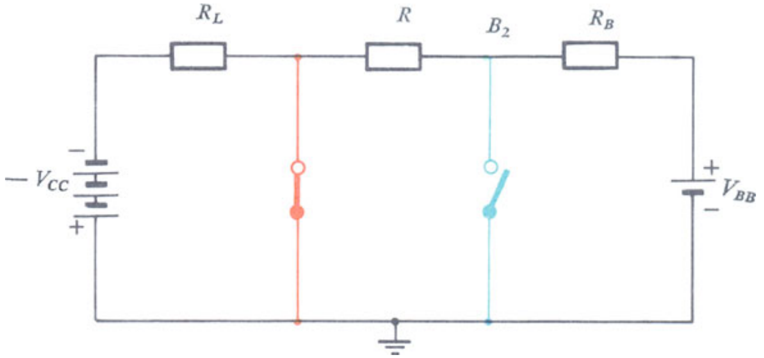


Fig. 506

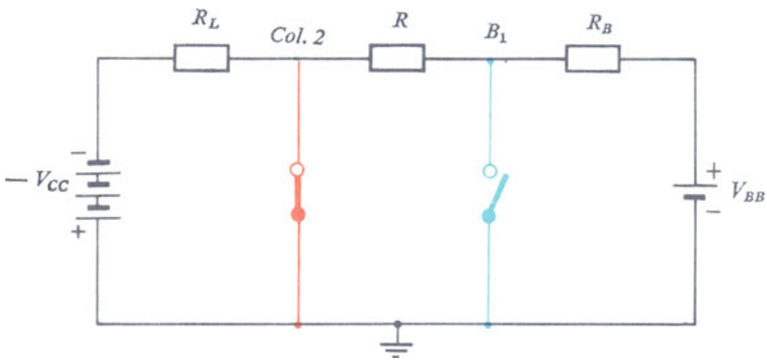


Fig. 507

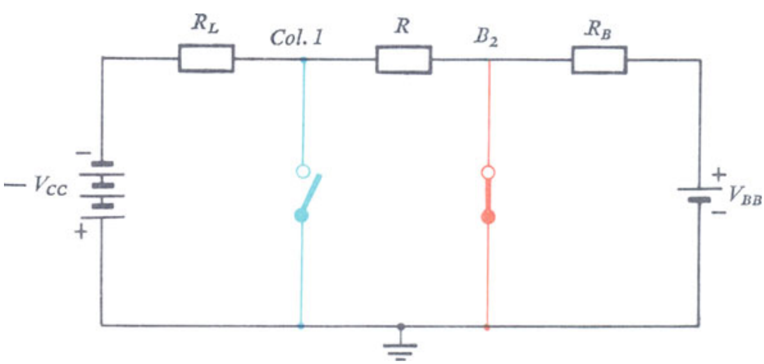


Fig. 508

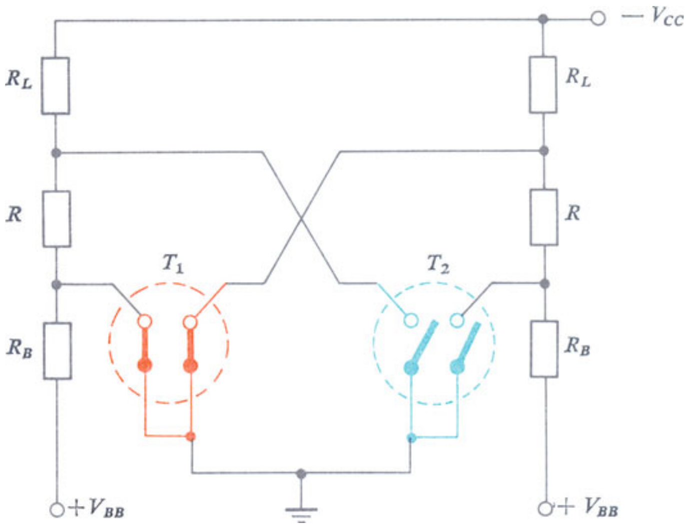


Fig. 509

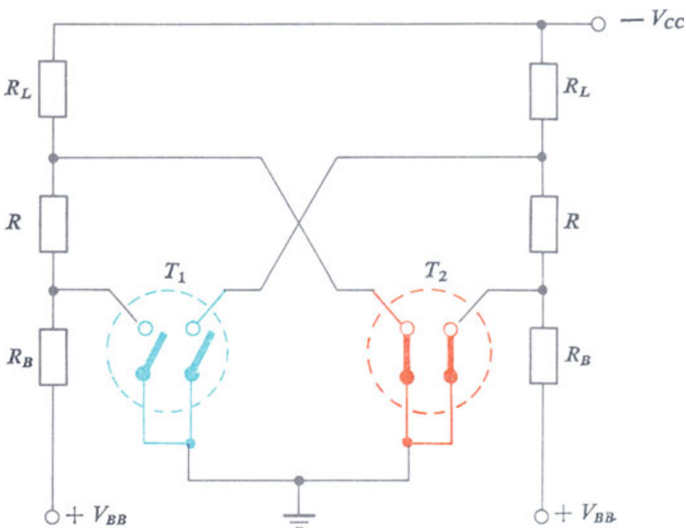


Fig. 510

Collector current of transistor T_1

$$I_{C1} = 0.$$

Base–emitter voltage of transistor T_2

$$V_{BE2} = 0.$$

Base current of transistor T_2

$$I_{B2} = \frac{V_{CC}}{R+R_L} - \frac{V_{BB}}{R_B}.$$

The values of these voltages and currents remain constant until the arrival of a second pulse.

Summary of the voltages and currents in a bistable multivibrator

First stable state. (Fig. 509).

	T_1 saturated	T_2 cut-off
V_{CE}	0	$-V_{CC}\dfrac{R}{R+R_L}$
V_{BE}	0	$V_{BB}\dfrac{R}{R+R_B}$
I_C	$\dfrac{V_{CC}}{R_L} - \dfrac{V_{BB}}{R+R_B}$	0
I_B	$\dfrac{V_{CC}}{R+R_L} - \dfrac{V_{BB}}{R_B}$	0

Second stable state. (Fig. 510).

	T_1 cut-off	T_2 saturated
V_{CE}	$-V_{CC}\dfrac{R}{R+R_L}$	0
V_{BE}	$V_{BB}\dfrac{R}{R+R_B}$	0
I_C	0	$\dfrac{V_{CC}}{R_L} - \dfrac{V_{BB}}{R+R_B}$
I_B	0	$\dfrac{V_{CC}}{R+R_L} - \dfrac{V_{BB}}{R_B}$

9.3. Diagrammatic representation

The diagrammatic representation of the signals available at the output terminals of a bistable multivibrator can be divided into four stages:

(a) first stable state;

(b) influence of a pulse;

(c) second stable state;

(d) influence of a second pulse.

Just as for the astable and monostable multivibrators, the variations of the voltage between the different active points of the circuit and earth can be illustrated by a group of figures:

(a) base–emitter voltage of transistor T_1 (Fig. 511a);

(b) collector–emitter voltage of transistor T_2 (Fig. 511b);

(c) collector–emitter voltage of transistor T_1 (Fig. 511c);

(d) base–emitter voltage of transistor T_2 (Fig. 511d).

First stable state

Transistor T_1 is saturated and T_2 is cut-off. The base–emitter voltage of the first transistor is zero (Fig. 511a) due to the forward bias of this junction. The collector–emitter voltage of transistor T_2 is a function of the voltage drop caused by the flow of a current equal to

$$\frac{V_{CC}}{R_L} - \frac{V_{BB}}{R+R_B}$$

in resistor R; namely $-V_{C2}$, the value of this voltage (Fig. 511b).

The collector–emitter voltage of transistor T_1 is zero (Fig. 511c) because of the saturation conditions of this transistor.

The base–emitter voltage of transistor T_2 is due to the battery $+V_{BB}$ and is equal to the voltage across resistor R, namely V_{B2} in Fig. 511d.

The values of these four voltages remain constant until a pulse is applied to one of the bases of the transistors.

Influence of a pulse. The problems connected with the triggering of a bistable multivibrator will be studied in greater detail at a later stage. There are in fact two possible methods of triggering the transistors, namely:

(a) the application of a positive voltage pulse between the base and emitter of the saturated transistor (T_1);

(b) the application of a negative voltage pulse between the base and emitter of the transistor (T_2) in the cut-off state.

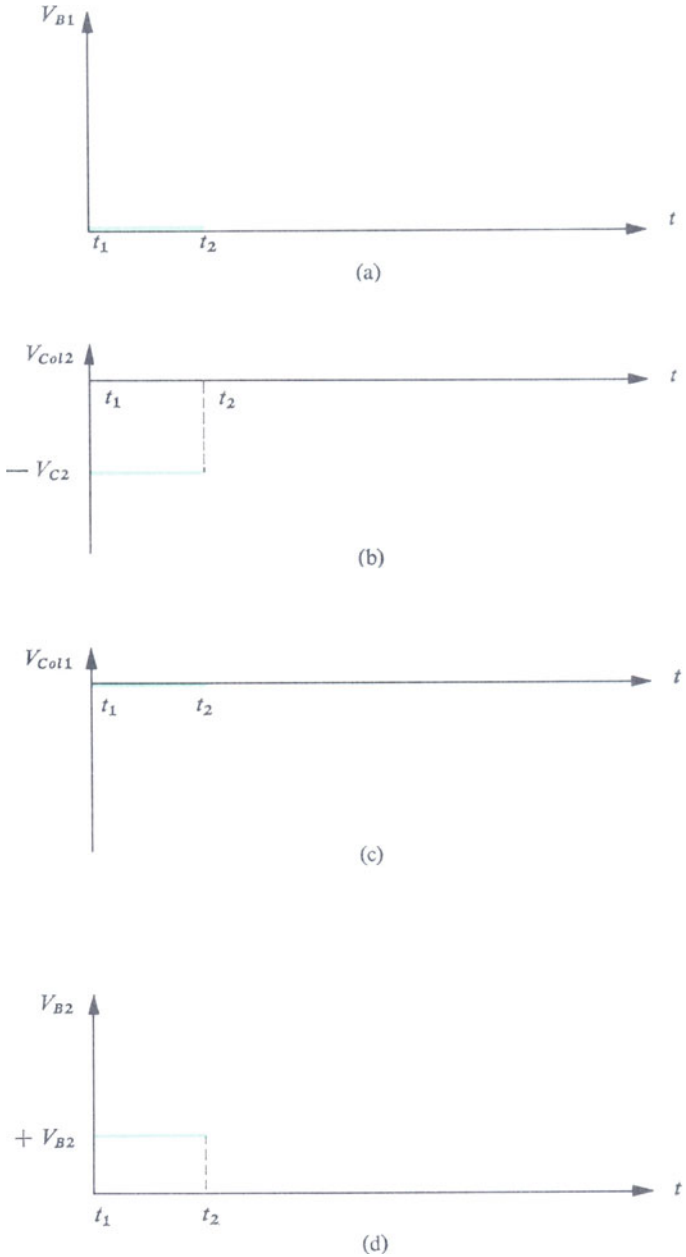
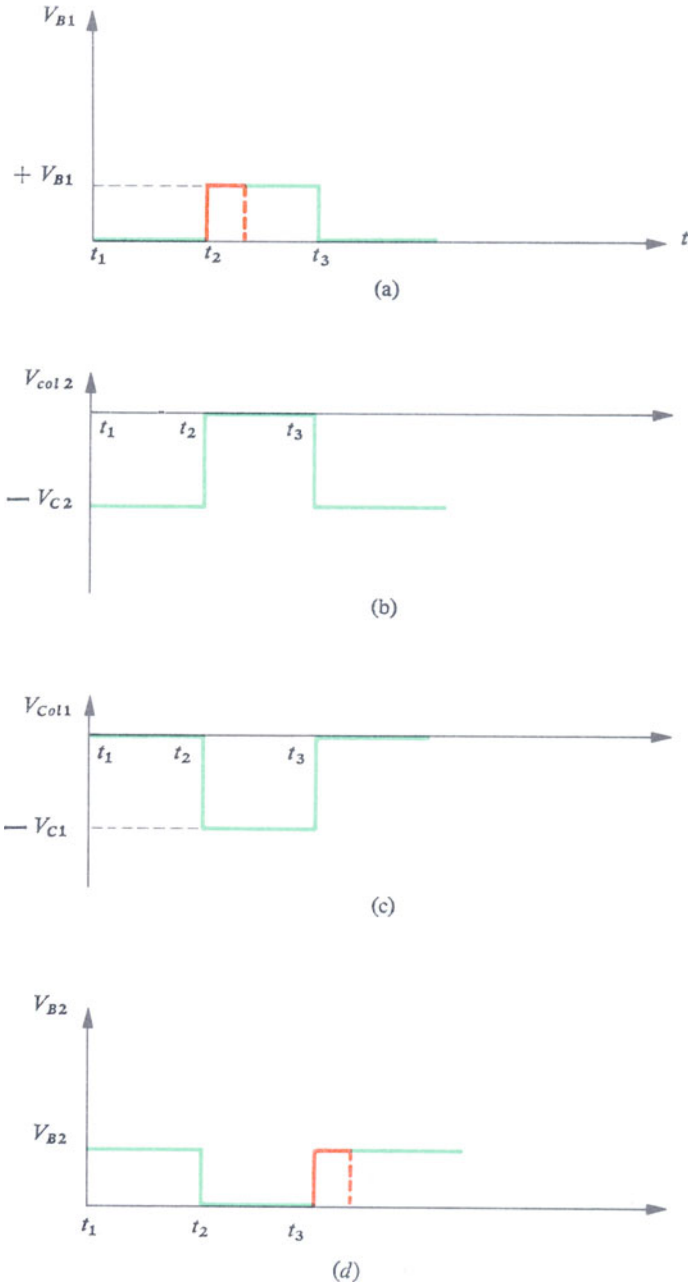


Fig. 511



V_{B1}
$+V_{B1}$
t_1
t_2
t_3
t
(a)
$V_{col\,2}$
t_1
t_2
t_3
$-V_{C2}$
(b)
V_{Col1}
t_1
t_2
t_3
$-V_{C1}$
(c)
V_{B2}
V_{B2}
t_1
t_2
t_3
(d)

Fig. 512

We will consider the first case as an example and assume that the abscissae of these four figures at instant t_2 correspond to the application of a positive pulse between the base and emitter of transistor T_1.

This positive voltage pulse appears at this instant in Fig. 512a and the base of transistor T_1 becomes positive with respect to the emitter.

The collector current of this transistor starts to decrease, and a negative voltage consequently appears between its collector and emitter (Fig. 512c).

Part of this instantaneous negative variation of the collector–emitter voltage of T_1 is applied between the base and emitter of transistor T_2 (Fig. 512d), which starts to conduct. It becomes practically saturated almost at once, causing a large decrease in the collector–emitter voltage (Fig. 512b). This positive variation is transmitted to the base of transistor T_1 and drives the transistor into cut-off.

It has been assumed that all the variations are instantaneous, and the delay and triggering times of these two transistors have been left out of consideration. They will be introduced into the discussion later in order to study the transient effects in greater detail.

Second stable state

Transistor T_1 is cut-off and T_2 saturated.

The base–emitter voltage of transistor T_1 is positive (Fig. 512a) and equal to $+V_{B1}$.

The collector–emitter voltage of transistor T_2 is zero (Fig. 512b) because the transistor is saturated.

The collector–emitter voltage of transistor T_1 is strongly negative and is a function of the voltage appearing across resistor R (Fig. 512c).

The base–emitter voltage of transistor T_2 is zero (Fig. 512b). The base–emitter junction of transistor T_2 is biased in the forward direction.

Second pulse. In principle the driving pulses of a bistable multivibrator are all in the same direction; our first choice was a positive voltage pulse in order to explain the triggering of this circuit and we will therefore examine the influence of a second positive pulse appearing for example at instant t_3 on the abscissae of Figs 512a, b, c and d.

A special problem should be considered at once; it consists in the explanation of the importance of the direction to be given to the pulses as a function of the operating state of each transistor.

A coupling circuit connected only to the base of the first transistor (T_1) would allow the application of this pulse only to the input of the stage which is already in the cut-off state. We must in these conditions have a special coupling circuit capable of selecting, in accordance with the direction of the pulse, the input circuit to be triggered.

We will therefore assume a suitable coupling circuit and a positive voltage pulse applied between the base and emitter of transistor T_2 at instant t_3.

The base of this transistor becomes positive at once (Fig. 513d) and the collector current decreases.

The collector–emitter voltage of transistor T_2 now becomes negative (Fig. 513b). This negative variation of the voltage is applied between the base and emitter of transistor T_1 (Fig. 513a), which begins to conduct and rapidly becomes saturated.

The collector–emitter voltage of transistor T_1 disappears instantaneously (Fig. 513c), and part of this positive voltage variation is transmitted to the base of transistor T_2 where it strengthens the cut-off condition.

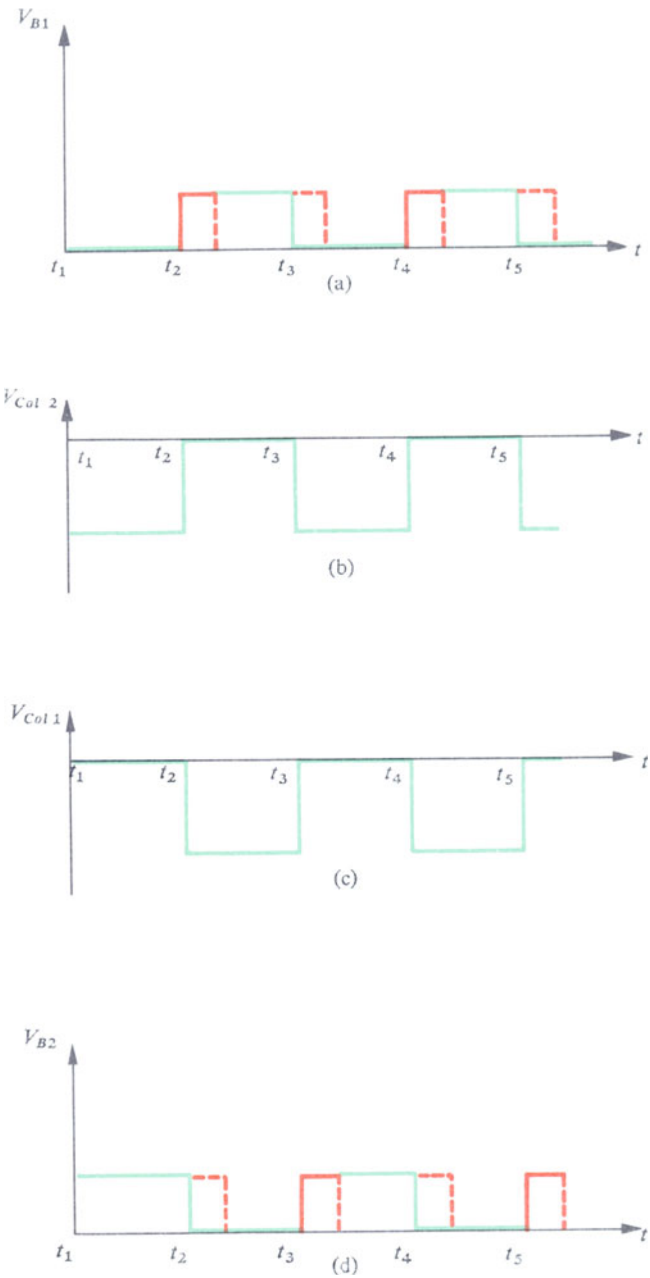
In the absence of fresh pulses the values of the various voltages remain constant at the values mentioned when discussing the first stable state. For this second pulse we have ignored the transient effects; it will be remembered that they will form the subject of a special study in the following paragraphs.

It will be seen at once from Figs 513b and c that for two pulses applied across the input of this circuit a single pulse is available at the collectors. Frequency division has therefore been achieved.

9.4. Diagrammatic method for the design of a bistable multivibrator

In the design of bistable multivibrators we can nearly always make a very rapid determination of the various values of the resistors to be inserted into such a circuit.

Before we pass on to the possibility of a purely diagrammatic solution of this problem, it might be of interest to revert to the earlier calculations made in the quantitative study of the two stable states and to modify the various voltages in such a way that the standard formulae for the calculation of the currents in the transistor during the period of conduction can be used at once.



V_{B1}
t_1
t_2
t_3
t_4
t_5
t
(a)
$V_{Col\ 2}$
t_1
t_2
t_3
t_4
t_5
t
(b)
$V_{Col\ 1}$
t_1
t_2
t_3
t_4
t_5
t
(c)
V_{B2}
t_1
t_2
t_3
t_4
t_5
t
(d)

Fig. 513

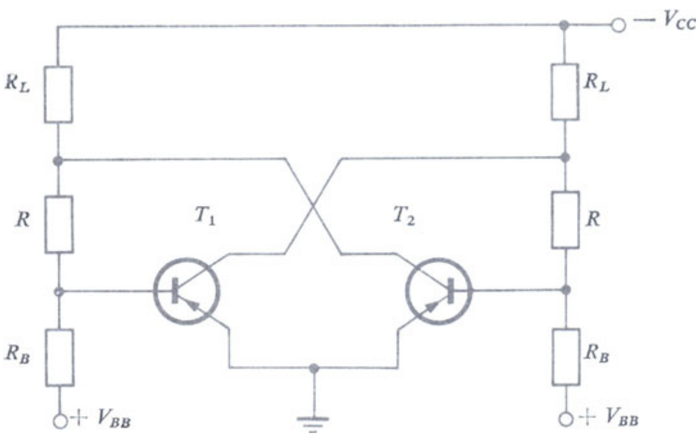


Fig. 514

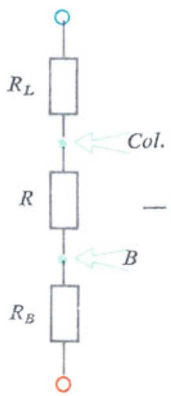


Fig. 515

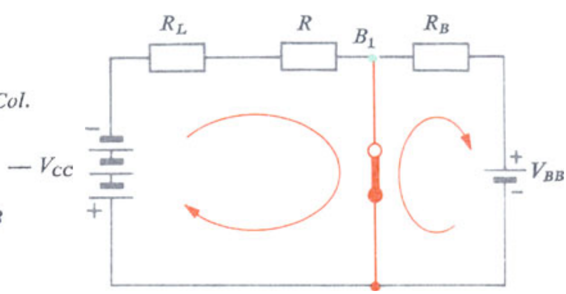


Fig. 516

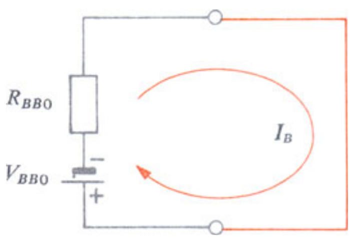


Fig. 517

Definition of the open-circuit voltages of the imaginary collector and base generators and of their internal resistance

Let us examine the complete circuit of a bistable multivibrator (Fig. 514) and consider separately the left hand section of the figure representing the three resistive components R_L, R and R_B. At the top of this chain of resistors (Fig. 515) we find the group of supply batteries $-V_{CC}$, and at the bottom the bias battery $+V_{BB}$.

Between resistor R_L and resistor R will be seen the connection to the collector of one of the transistors; and between R and R_B the connection to the base of the other transistor.

Referring back to the study of the first stable state, we see that transistor T_1 is saturated and T_2 cut-off.

The base current may be calculated easily by means of the equivalent circuit of Fig. 516, and is given by the relation:

$$I_B = \frac{V_{CC}}{R_L+R} - \frac{V_{BB}}{R_B}.$$

On the other hand, the collector circuit of transistor T_2 is open, as it is in the cut-off state.

In Figure 515 we could consider for the connection point B the possibility of replacing the group of resistors and batteries by a simple generator supplying an open-circuit voltage equal to $-V_{BBO}$ and having an internal resistance equal to R_{BBO}.

If this generator is short-circuited (as in Fig. 517) the short-circuit current is the same as the base current of the saturated transistor.

Open-circuit voltage supplied by the generator. This voltage is equal to the sum of two voltages:

(a) the voltage across resistor R_B due to the group of batteries $-V_{CC}$;

(b) the voltage across the group R_L and R due to the presence of battery V_{BB}.

The voltage across resistor R_B is given by the expression:

$$V_{(R_B)} = -V_{CC}\frac{R_B}{R_L+R+R_B}.$$

The voltage across the group R_L and R is equal to:

$$V_{(R+R_L)} = V_{BB}\frac{R+R_L}{R_L+R+R_B}.$$

The open-circuit voltage given by the generator shown in Fig. 518 is consequently given by the relation:

$$V_{BBO} = V_{(RB)} + V_{(R+R_L)} = -V_{CC}\frac{R_B}{R_L+R+R_B} + V_{BB}\frac{R+R_L}{R_L+R+R_B}.$$

Internal resistance of this generator. With respect to the base point in Fig. 519 it will be seen that there are two groups of resistors in parallel, namely:

(a) resistor R_B;

(b) resistors R and R_L.

The equivalent resistance of this group is given by the expression:

$$R_{BBO} = \frac{R_B(R+R_L)}{R_L+R+R_B}.$$

The interest of these two items is that it is now possible to calculate the value of the base current of the transistor in the saturated state; all that is needed is to calculate the ratio between the open circuit voltage mentioned above and the internal resistance of this generator, whence:

$$I_{BS} = \frac{V_{BBO}}{R_{BBO}}.$$

Open-circuit voltage given by the collector generator. The chain of resistors on the right of Fig. 520 (shown in Fig. 521) resembles that already considered in Fig. 519.

On the other hand, since the base of the transistor is open and the collector connected directly to earth (collector of transistor T_1) the corresponding value of this current may be calculated by the conventional method, namely:

$$I_C = \frac{V_{CC}}{R_L} - \frac{V_{BB}}{R+R_B}.$$

As in the previous case, this collector point can be considered as connected across the terminals of a generator supplying an open-circuit voltage equal to $-V_{CCO}$ and having an internal resistance represented by R_{CCO}.

The open-circuit voltage available at the output of a generator of this nature is a function (Fig. 522):

(a) of the voltage due to the generator in the group of resistors R and R_B;

(b) of the voltage due to the battery $+V_{BB}$ and appearing across the load resistor R_L.

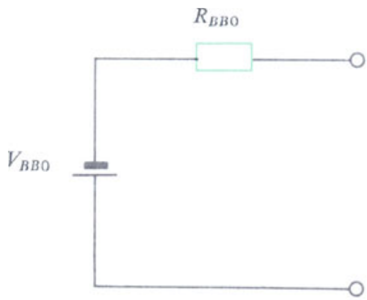


Fig. 518

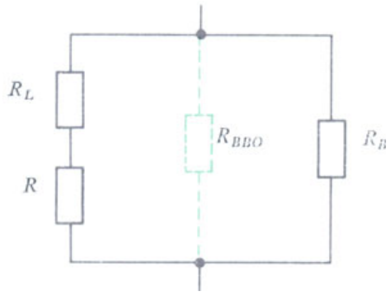


Fig. 519

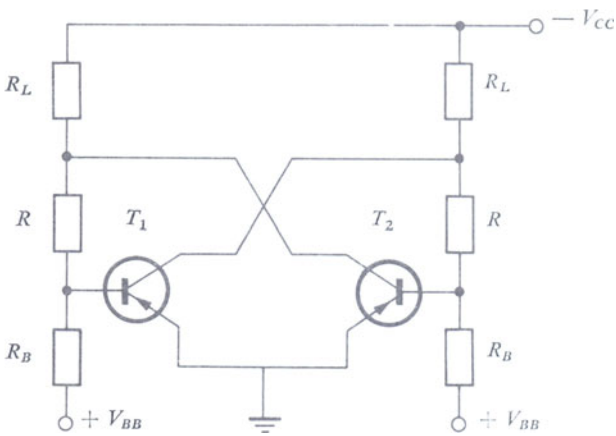


Fig. 520

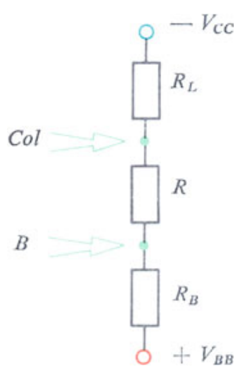


Fig. 521

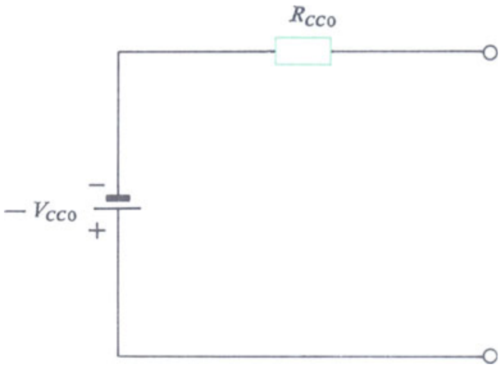


Fig. 522

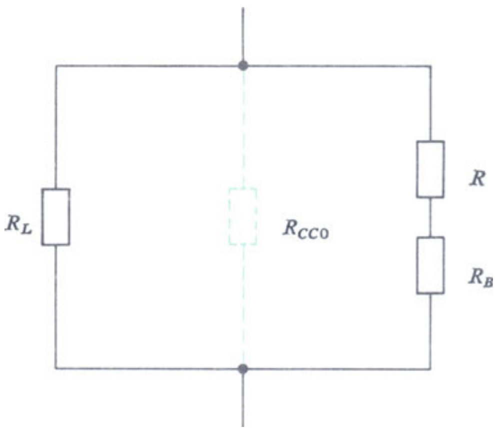


Fig. 523

The first of these voltages is given by the equation:

$$V_{(R+R_B)} = -V_{CC}\frac{R+R_B}{R_L+R+R_B}.$$

The second by the expression:

$$V_{(R_L)} = V_{BB}\frac{R_L}{R_L+R+R_B}.$$

The total available voltage when the generator is in the open position is then equal to:

$$V_{CCO} = V_{(R+R_B)}+V_{(R_L)} = -V_{CC}\frac{R+R_B}{R_L+R+R_B}+V_{BB}\frac{R_L}{R_L+R+R_B}.$$

Internal resistance of the collector generator. Seen from the collector, the chain of resistors appears as an association of two groups of resistors in parallel (Fig. 523):

(a) resistor R_L;

(b) resistors R and R_B in series.

The equivalent resistance of a group of this nature is given by the equation:

$$R_{CCO} = \frac{R_L(R+R_B)}{R_L+R+R_B}.$$

Since we know the open-circuit voltage supplied by the generator and its internal resistance, we can now divide one of these two quantities by the other and therefore calculate the short-circuit current of the generator, that is the collector current of the transistor in the saturated state. This is given in the form:

$$I_{CS} = \frac{V_{CCO}}{R_{CCO}}.$$

The interest of being able to calculate immediately the values of the base and collector currents of a transistor in the conducting state lies in the fact that it enables us to check at once whether the condition of saturation for this transistor is respected.

This is given by the relation:

$$I_C/I_B < h_{FE}.$$

All that is necessary is to replace I_C and I_B in this equation in order to obtain an expression for the immediate calculation of the saturation condition of the transistor.

Diagrammatic solution for a bistable multivibrator

The advantage of diagrammatic solutions is that they avoid more or less complicated calculations and often allow a more nearly ideal method for the design of a circuit.

This method, which is often used in laboratories, consists in the transposition of the various quantitative data on the voltages and resistances into proportional indications on the ordinates and abscissae.

In the bistable multivibrator, shown in Fig. 524, the electrical information concerned is:

(a) the supply voltage $-V_{CC}$;

(b) the bias voltage V_{BB};

(c) the resistors R_L, R and R_B.

In Fig. 525 the voltages are shown on the ordinate. With respect to zero, the positive bias voltage is at a higher level (V_{BB} in the figure) and the collector–emitter voltage at a lower level ($-V_{CC}$).

Draw the perpendicular from point $-V_{CC}$ on the ordinate and then let us try to graduate the straight line thus obtained in values of resistance.

In the multivibrator stage, the load resistor R_L is in direct contact with the negative pole of the group of supply batteries $-V_{CC}$. If the lowest horizontal line is graduated in values of resistance, the load resistor R_L may be located there, starting from the point $-V_{CC}$.

In Fig. 524 coupling resistor R follows the load resistor R_L on the same axis.

Finally, resistor R_B is at the end of the chain already mentioned, and the corresponding value of this resistor may be placed at the right-hand end of the line.

The data mentioned at the beginning of this paragraph are clearly shown in the diagram just constructed.

Use of the diagrams. In addition to the purely static parameters already located on the diagram, it is interesting to include the information which can be obtained from the diagram itself.

Because of the saturation condition of one of the transistors (T_1 for instance) we must calculate the base and collector currents. It is not possible to locate directly a value of current in a figure of this nature, but on the other hand it is certainly easy to calculate the voltages and the equivalent resistance values.

In the above pages we have emphasized the interest of the calculation of the open-circuit voltages and the equivalent resistances of the generator.

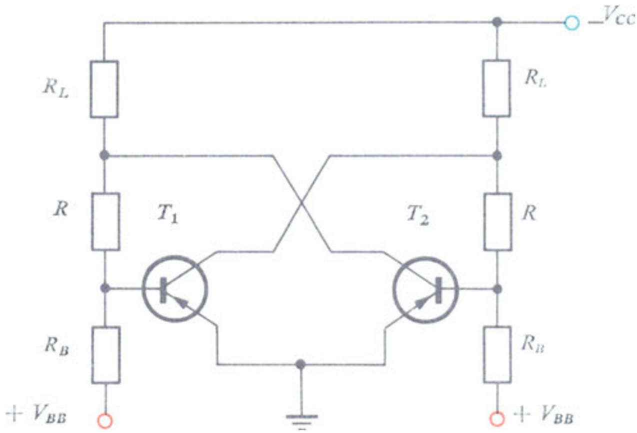


Fig. 524

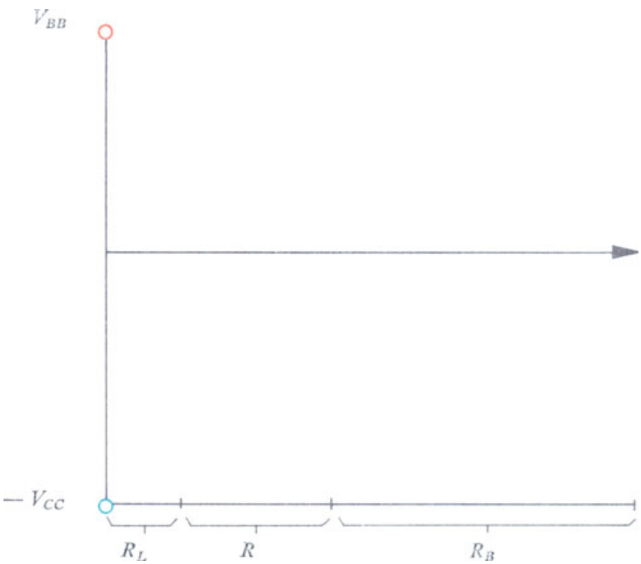


Fig. 525

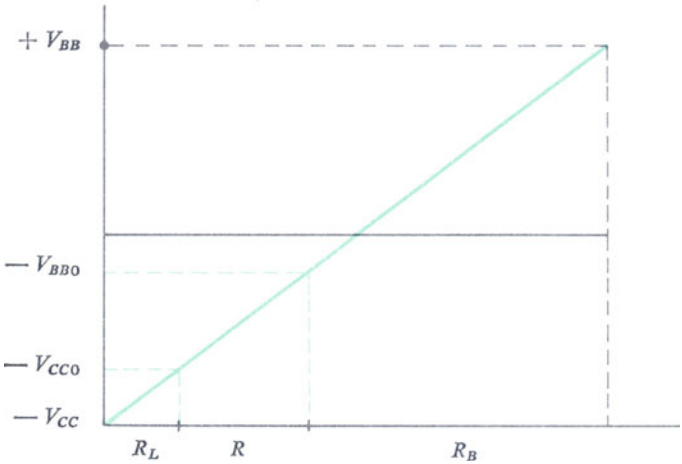


Fig. 526

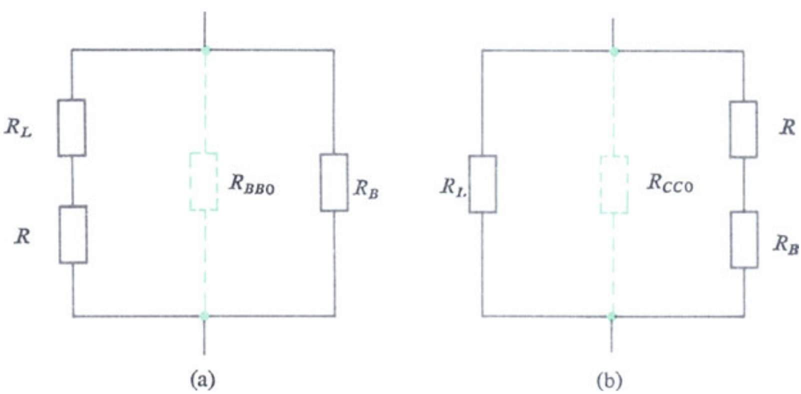


Fig. 527

By means of the diagram in Fig. 526 we can determine the value of the two voltages $-V_{CCO}$ and $-V_{BBO}$ accurately, as well as the equivalent resistances R_{CCO} and R_{BBO}.

Determination of the open-circuit voltages. In Fig. 526 we will draw a horizontal line passing through point $+V_{BB}$ and a vertical line at the end of resistor R_B. This constitutes a rectangle in which we can draw a diagonal from the bottom left hand end to the opposite end (top right).

Raise a perpendicular to the abscissa to the point of connection of resistors R_L and R. This line cuts the diagonal at point A. All that is needed now is to project this point on to the ordinate in order to obtain a value of voltage corresponding to the open-circuit voltage given by the collector generator $-V_{CCO}$.

In the same way, raise a perpendicular from the abscissa to the point of connection of resistors R and R_B. This perpendicular cuts the diagonal and the projection of this point on to the ordinate gives the open-circuit voltage supplied by the generator located in the base circuit. This voltage is equal to $-V_{BBO}$.

Calculation of the equivalent resistances. It is difficult to make this calculation by means of the diagram used for the quantitative determination of the open-circuit voltages. It is therefore preferable to return to an electrical representation, namely the two groups of resistors shown in Figs 527a and b.

The internal resistance of the generator supplying the collector of the transmitter in the conducting state is represented by the connection in parallel of two groups of resistors. The equivalent resistance was calculated earlier, and found to be equal to:

$$R_{CCO} = \frac{R_L(R+R_B)}{R_L+R+R_B}.$$

There is no need to repeat the explanation of the calculation of the internal resistance of the generator in the base circuit of the transistor in the saturation state. The position is shown clearly in Fig. 527b, which makes it easy to understand the following equation at once:

$$R_{BBO} = \frac{R_B(R+R_L)}{R_L+R+R_B}.$$

As before, a knowledge of the voltages and currents enables us to control the saturation conditions of transistor T_1, for example, if this transistor is in its conducting state.

If transistor T_1 is saturated, transistor T_2 should be cut-off. For a *PNP* transistor the cut-off state (involving the reverse bias of the two junctions) requires a negative collector–emitter voltage and a positive base–emitter voltage.

If we refer back to the diagram in Fig. 528 we may try to use it to determine the values of the base–emitter voltage and the collector–emitter voltage of transistor T_2.

Collector–emitter voltage of the transistor in the cut-off state. In a circuit of this type a collector–emitter voltage appears when the base of the other transistor is connected direct to earth. In the diagram the earth should be associated with the level O, that is with the initial abscissa.

Since the voltage at the point of connection between resistors R and R_B is zero, we may take this point as reference on the abscissa and raise a perpendicular only as far as the intersection with the zero potential axis. The operating point of the multivibrator must pass through this point (Fig. 528); it consists of two sections of a line:

(a) one, joining the end (bottom left) to point C;

(b) the other, going from point C to the end at the top right.

If we raise a perpendicular to the abscissa to the point of connection of resistors R_L and R, the line thus obtained cuts the operating line (in blue in the figure) at point D. The projection of point D on to the ordinate at once gives us the value of the collector–emitter voltage of the transistor in the cut-off state ($-V_{C2}$ in Fig. 528).

Base–emitter voltage of the transistor in the cut-off state. The method of calculating this voltage is exactly the same as the previous one. The collector of the transistor in the saturated state involves a zero value of voltage at the point of connection between resistors R_L and R. The perpendicular to the abscissa, raised at the point of intersection, reaches the horizontal centre line at point Q.

We can now draw a second group of operating lines for the multivibrator. All that is needed is to follow the method shown in Fig. 529.

The base–emitter voltage of the transistor in the cut-off state is obtained by projecting the point of intersection of the perpendicular on to the abscissa to the point R, R_B, and of the operating line on to the ordinate. If the resistors are properly adjusted the positive voltage between the base and emitter ($+V_{B2}$) is high enough to cause the transistor to be in the cut-off state corresponding to the requirements of the particular design.

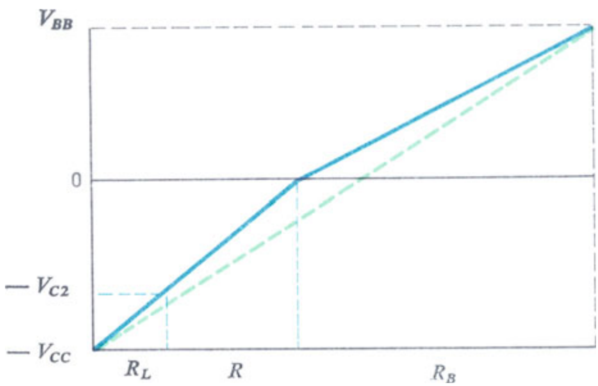


Fig. 528

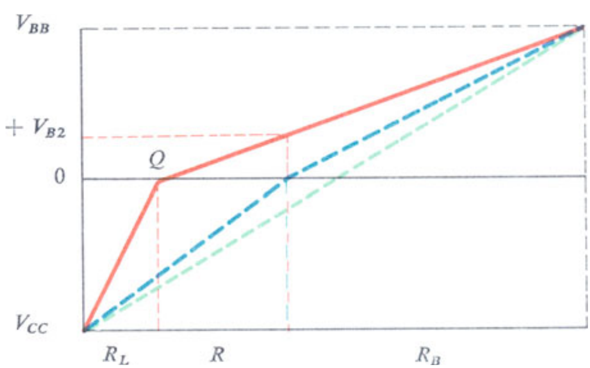


Fig. 529

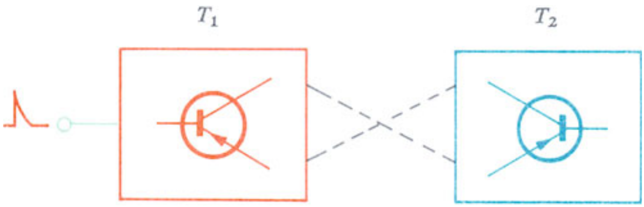


Fig. 530

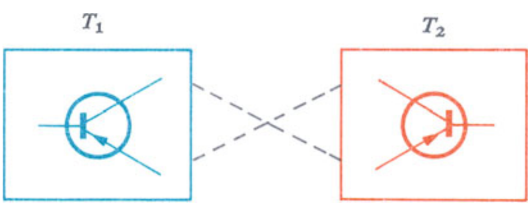


Fig. 531

Illustration and simplified study of the transient effects

Transient effects occur during transition from one stable state to another.

When we were discussing the principle of operation of a circuit of this nature we ignored this effect, by assuming that the multivibrator changed instantaneously from one state to the other.

For example, the synoptic diagram in Fig. 530 represents the first stable state of the multivibrator; transistor T_1 is saturated and T_2 cut-off.

When a pulse is received the circuit triggers and the operating state becomes that shown in Fig. 531; transistor T_1 is then cut-off and T_2 saturated.

We now have to define the method of passing from one state to the other. We will assume that the multivibrator is driven by a positive pulse applied between the base and emitter of the saturated transistor, that is transistor T_1 in Fig. 530.

The pulse makes the base of T_1 positive with respect to its emitter; as the collector is still negative, the transistor is placed in the inverse operating state. On the other hand transistor T_2 is still in its cut-off period at this instant.

The current variations in the circuit cause the appearance of a collector current in transistor T_2 which passes into the normal operating state when transistor T_1 is still in a state of inverse operation.

The removal of the charges stored in the base and the reduction of the collector current cause transistor T_1 to be in the cut-off state. Transistor T_2 is still operating normally.

Finally, transistor T_2 reaches saturation level, and transistor T_1 being already cut-off, the second stable state is reached. The table below summarizes the different stages occurring during this transient period.

	T_1	T_2
1st stable state	saturation	cut-off
1st transient period	reverse	cut-off
2nd transient period	reverse	normal
3rd transient period	cut-off	normal
2nd stable state	cut-off	saturation

9.5. Triggering of a bistable multivibrator

In the above explanation of the triggering of a bistable multivibrator from one state to the other we drew attention to the necessity of being able to direct the pulse to the most suitable input circuit.

Refer back to the synoptic diagram in Fig. 532. If we assume the initial stable state (T_1 saturated, T_2 cut-off) and if we have positive pulses, the first pulse should be applied between the base and emitter of the saturated transistor (T_1).

The multivibrator triggers and the two transistors change their state (T_1 cut-off and T_2 saturated).

The second positive pulse should now be applied to the base of transistor T_2, which is in the saturated state. We should therefore require a coupling circuit to direct the new pulse to the base of this transistor and to prevent its being applied simultaneously between the base and emitter of transistor T_1, which is already in the cut-off state. This rules out a capacitive coupling system like that shown in Fig. 533.

The capacitors are unable to direct the trigger pulse; in these conditions the same positive pulse is applied simultaneously to the inputs of the two transistors.

What is the drawback of this simultaneous distribution of the two items of input information?

The positive pulse applied between the base and emitter of transistor T_2 tends to limit the value of the corresponding collector current and to trigger transistor T_2 from saturation to cut-off. On the other hand, the same pulse applied between the base and emitter of transistor T_1 drives the transistor more strongly into cut-off and renders the triggering of the multivibrator much more problematical.

Consequently the only solution is to use coupling components which will allow the information to pass in one case and oppose its passage in the other. Components answering these requirements are well known; they are, of course, diodes, which have been studied in detail in this book.

Clamping diode circuit

In the circuit of Fig. 534, two coupling circuits each consisting of a diode are added to the standard components of the bistable multivibrator: namely D_1 the diode driving transistor T_1 and D_2 the diode connected to the input of transistor T_2. These two diodes are connected to the same pulse generator, which may also form one of the stages coming before the multivibrator under consideration.

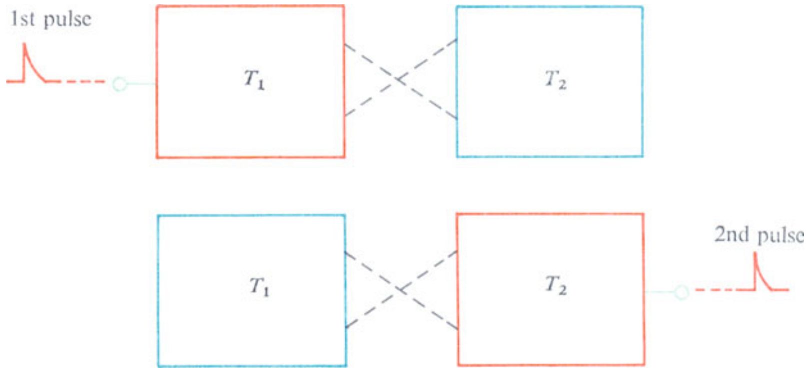


Fig. 532

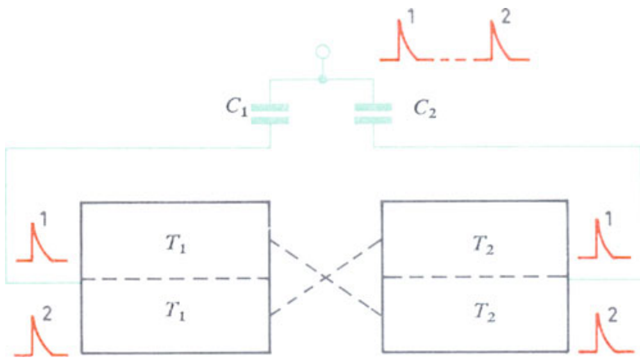


Fig. 533

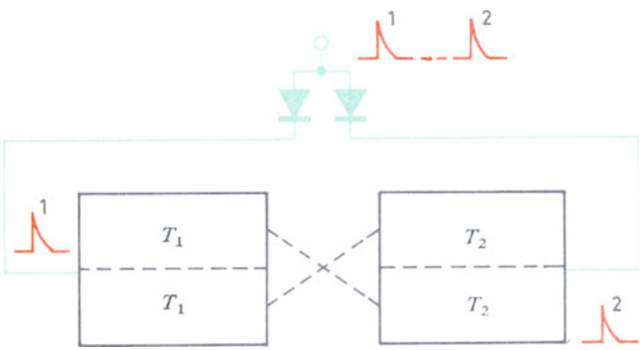


Fig. 534

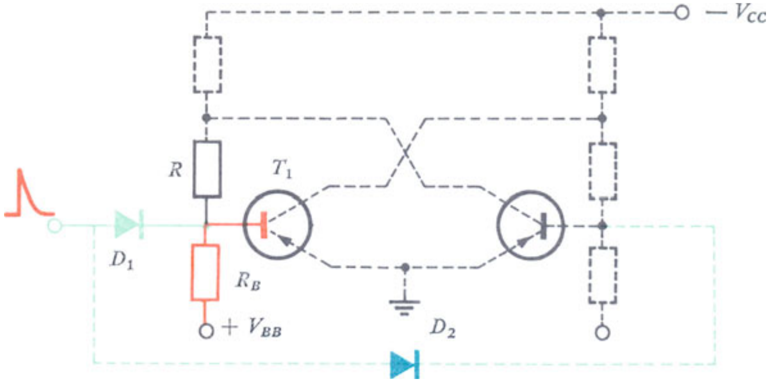


Fig. 535

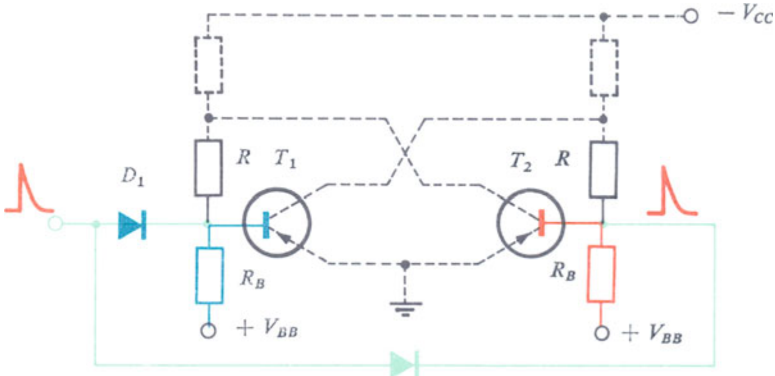


Fig. 536

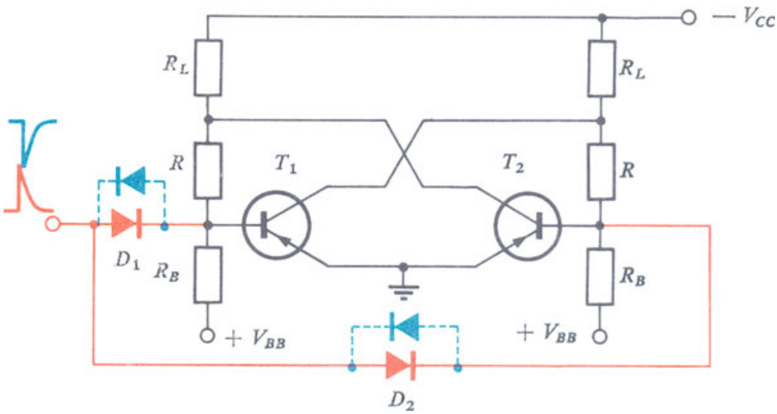


Fig. 537

If we examine the input circuit of transistor T_1 (Fig. 535) more closely we see that it consists, in addition to the base–emitter junction of this transistor, of resistors R_L, R and R_B. Diode D_1 is connected between the base of transistor T_1 and the pulse generator.

T_1 is saturated and the positive pulse supplied by the generator should therefore be applied to its input terminals.

When the anode of diode D_1 is positive with respect to its cathode, the signal, which produces this forward bias, is transmitted to the base of the transistor. The collector current therefore starts to decrease and the system triggers.

In the next stage, that is for a second positive pulse, transistor T_1 is cut-off. The corresponding base–emitter junction is biased in the reverse direction and the base is positive with respect to the emitter. Diode D_1 is then permanently biased in the reverse direction and in these conditions it cannot transmit the trigger pulse.

On the other hand, the saturated state of transistor T_2 makes it suitable for the reception of the positive pulse supplied by the generator. Diode D_2 is not biased in the absence of a pulse, but it becomes biased in the forward direction when a pulse is applied to its terminals (Fig. 536).

Transistor T_2 is now no longer saturated and the multivibrator changes from the stable state.

In the complete diagram of the bistable multivibrator associated with the clamping diode circuit (Fig. 537), we see that:

(a) a train of positive pulses excites successively the bases of the transistor in the saturated state (first pulse at the base of T_1, second pulse at the base of T_2, third pulse at the base of T_1, etc.);

(b) for negative pulses the connection of the diodes will be different, but the transmission of information will take place in the same way (first pulse on the base of T_2, second pulse on the base of T_1, third pulse on the base of T_2, etc.).

In the explanation of the transmission of the pulses, we showed that the presence of diodes in the coupling circuits prevented the same positive pulse being transmitted to the base of two transistors.

The drawback of the simultaneous application of a pulse to the two bases was mentioned earlier; the positive pulse applied to the base of the cut-off transistor can only strengthen this operating state. It is then very difficult for the multivibrator to trigger.

If the amplitude of the positive pulse is greater than the reverse bias voltage of the diode connected in the trigger circuit of the transistor in the cut-off state, the diode becomes biased in the forward direction and the transmitted information drives the transistor more strongly into cut-off.

There are a number of improvements by means of which this risk may be avoided. The most important is to connect the coupling diodes directly to the collectors of the transistors.

Collector drive. The circuit is then as shown in Fig. 538.

The positive pulse supplied by the generator is transmitted to the base of transistor T_1, which is saturated, through capacitor C. The collector current of this transistor decreases and the state of the multivibrator changes.

When the next pulse arrives, the second diode allows the passage of positive information to the base of transistor T_2 causing a new possibility of triggering of the multivibrator. On the other hand, given the conditions of reverse bias of the first diode, transistor T_1 receives no information.

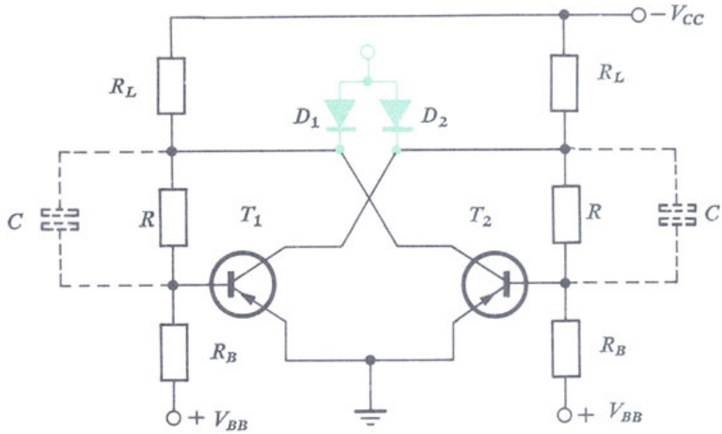


Fig. 538

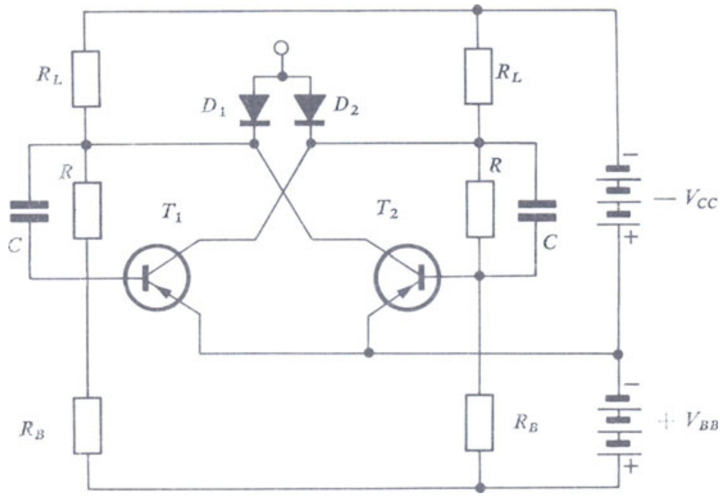


Fig. 539

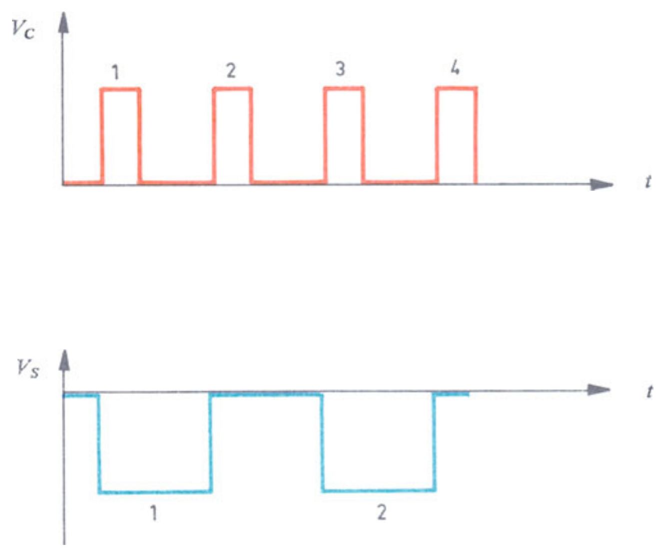


Fig. 540

9.6. Actual circuit

Just as for astable and monostable multivibrators, a diagram of a bistable multivibrator according to the principles already developed will be shown.

Figure 539 gives an example of the production of a bistable multivibrator. It consists of:

(a) two transistors T_1 and T_2;

(b) two load resistors equal to 1 kΩ;

(c) two acceleration circuits ($R = 5{\cdot}6$ kΩ; $C = 470$ pF) acting at the same time as coupling circuits;

(d) two bias resistors equal to 33 kΩ.

The circuit is supplied by a group of batteries giving a voltage of -6 V and biased by another group of batteries giving a positive voltage equal to 6 V.

The trigger pulses are positive.

The principle of operation of a multivibrator of this nature is as follows:

(a) in the absence of a pulse, transistor T_1 is saturated and T_2 cut-off;

(b) when the first positive pulse is received transistor T_1 is driven to cut-off and transistor T_2 to saturation;

(c) in the absence of another pulse the multivibrator remains in this operating state;

(d) the second positive pulse is applied between the base and emitter of transistor T_2, the system triggers and returns to the original conditions.

For every two pulses applied across the input of the stage, a single pulse is received in the output circuit (Figs 540a and b).

10. Switching Circuits Blocking Oscillators

10.1. Astable blocking oscillators

During the study of the principles of switching, we mentioned that the first two functions were concerned with two possibilities:

(a) the generation of rectangular signals;

(b) pulse shaping.

The first function may be performed by a relaxation system. In this connection we should remind ourselves of the astable multivibrator and its principle of operation.

The second function consists in the replacement of a distorted signal by a second pulse with 'as straight a leading edge' as possible (the monostable multivibrator).

The signals are generated by astable blocking oscillators, and the pulses are reshaped by triggered blocking oscillators.

Principle of operation of an astable blocking oscillator

The astable multivibrator consists of two transistors associated with *R* and *C* components. It may be compared to a group of circuits associated with components having a certain time constant.

The astable blocking oscillator is constituted in accordance with the same principle, namely:

(a) a transistor (blue rectangle *B* in Fig. 541);

(b) a transformer (red rectangle *A* in the same figure);

(c) an *RC* circuit (green rectangle *C*).

The three basic components of the astable multivibrator are represented in a synoptic diagram of this nature.

Principle of operation of an astable blocking oscillator

Figure 542 shows the diagram of an astable blocking oscillator with an *RC* circuit in series with the base of the transistor.

This circuit includes, in addition to the transistor, a transformer, the primary of which charges the collector, whilst the secondary excites the input circuit of the transistor.

The supply is from a group of batteries $-V_{CC}$. The bias is obtained by a resistance bridge R_{B1} and R_{B2} connected across the source of supply. The high-value capacitor C_B prevents any fluctuation of the direct current present between the base and emitter of the transistor due to the signal.

The transistor is triggered by the R_E and C_E circuit connected in series with the emitter of the transistor.

The three functions defined by the rectangles *A*, *B* and *C* in Fig. 541 occur again in a circuit of this type.

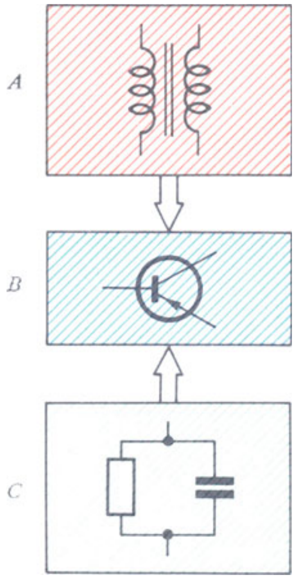


Fig. 541

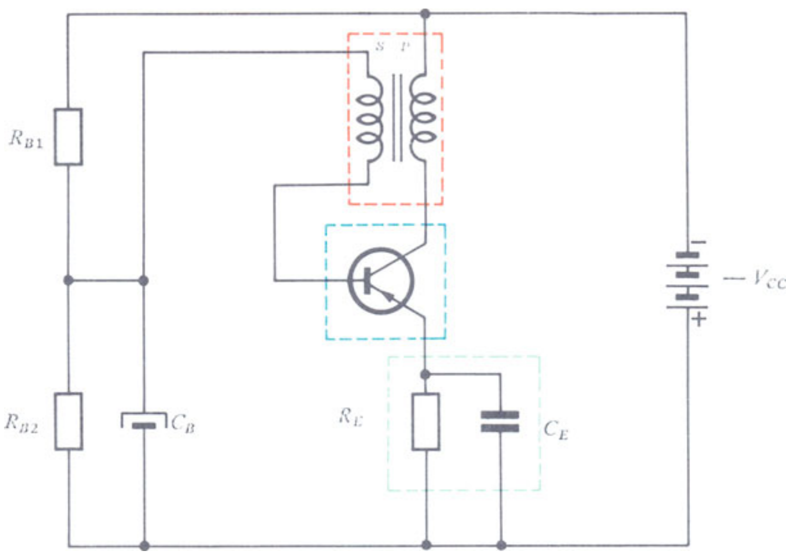


Fig. 542

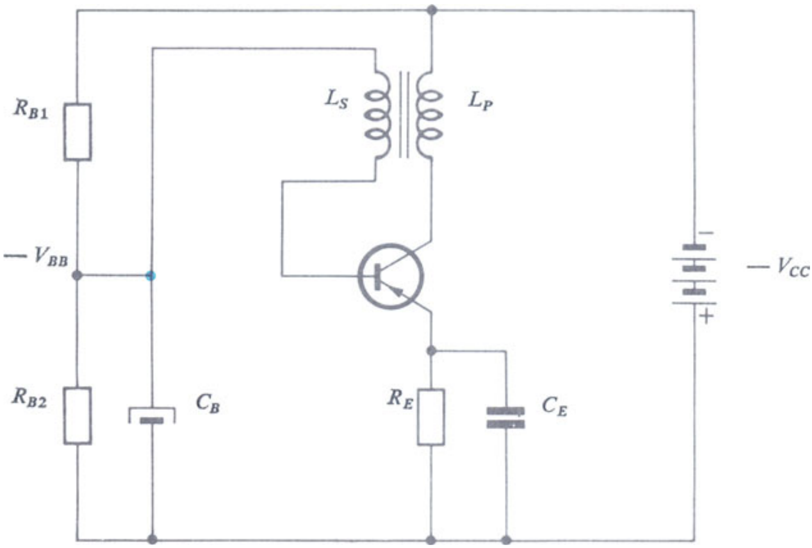


Fig. 543

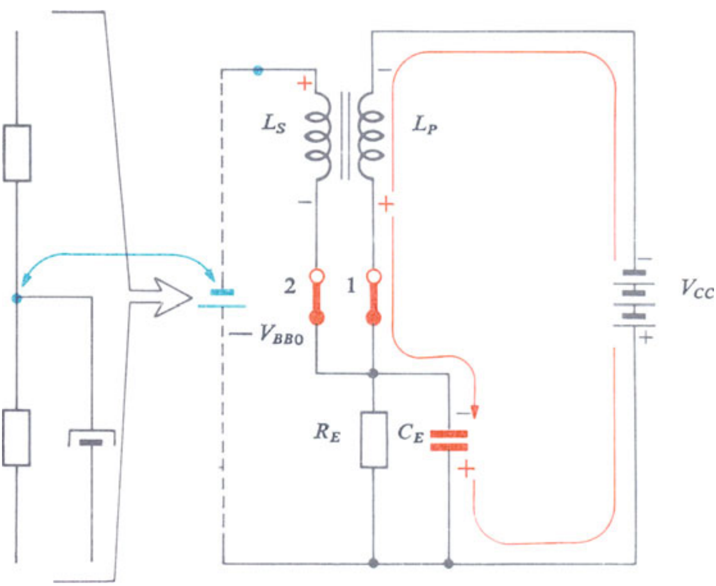


Fig. 544

Brief explanation of the operation

The system is triggered by the R_E C_E circuit. The transistor in fact acts as a switch, in that it can operate in the two limiting states:

(a) cut-off;

(b) saturation.

As an introduction to the preliminary explanations, it will be assumed that the state of the transistor has changed from cut-off to saturation. Its equivalent circuit can be represented simply by two closed switches in Fig. 544 (red switches 1 and 2).

A current appears at once in the collector circuit, and increases until it reaches the saturation value $-I_{CS}$. This current produces two effects:

(a) it causes the appearance of a voltage across the primary (L_P) of the transformer, and this voltage appears in the other direction across the secondary, where it strengthens the driving of the transistor into saturation;

(b) it charges the capacitor C_E with the polarities shown in Fig. 544.

On the left of the figure it will be seen that the biasing bridge R_{B1} and R_{B2} associated with the capacitor C_B is replaced by a single battery supplying a voltage equal to $-V_{BBO}$. It will therefore be easier in what follows to compare the different voltages and to explain the possible triggering of the transistor from one state to the other.

There are in fact two stages to be considered:

(a) the transient stage, corresponding to the period of transistion from cut-off to saturation;

(b) the stable period dependent on the beginning of the entry of the saturation transistor into saturation.

Transient stage. The voltage variation due to the transistor leaving the cut-off state will be considered on the next page. As soon as a current appears in the circuit it produces an e.m.f. in the primary of the transformer. The result is the appearance of a voltage in the other direction across the secondary which hastens the increase of the collector current.

Stable stage. When the transistor becomes saturated, the collector current remains constant and the voltages across the primary and secondary of the transformer disappear, because the values of the base and collector currents are constant. During these two periods (transient and stable stages) the capacitor is charged through the transistor and the primary of the transformer.

The state of saturation is connected quantitatively with the ratio between the output and input currents, namely:

$$I_C/I_B < h_{FE}.$$

In fact the saturation depends primarily on the bias conditions of the two junctions of the transistor. The base–emitter junction should be forward biased, causing a negative voltage between the base and emitter of this transistor.

If we look at Fig. 545 we see that the base–emitter voltage depends:

(a) on the one hand, on the bias voltage $-V_{BBO}$;

(b) on the other hand, on the voltage across the *RC* circuit connected in series with the emitter.

This voltage may be calculated by means of the expression:

$$V_{BE} = -V_{BBO} + V_{(C_E)}.$$

During the initial cut-off period the capacitor C_E has a strong negative charge, causing the base–emitter voltage to be positive. The capacitor then starts to discharge and the base–emitter voltage becomes negative. The transistor conducts and is rapidly driven into saturation.

The process of triggering from cut-off to saturation and the influence of the transformer were explained on the previous page.

Transition from saturation to cut-off. The increase in the charging voltage of the capacitor C_E accounts for the negative voltage between the base and emitter of the transistor:

$$V_{BE} = (-V_{BBO}) + V_{(C_E)}.$$

From the instant when the charging voltage of C_E becomes higher than the bias voltage $-V_{BBO}$, the base is positive with respect to the emitter, and the transistor is no longer saturated.

The collector current begins to decrease and the e.m.f. across the primary of the transformer is in opposition to this variation. The voltage induced in the secondary now drives the transistor more strongly into non-conduction.

Stable stage. The transistor very rapidly enters the cut-off state (the open switch shown in blue in Fig. 546).

Capacitor C_E discharges into resistor R_E and the voltage at its terminals ($V_{(C_E)}$) decreases. The base then becomes less and less positive, and at the instant when this voltage becomes equal to the bias voltage the transistor begins to conduct; the system then triggers.

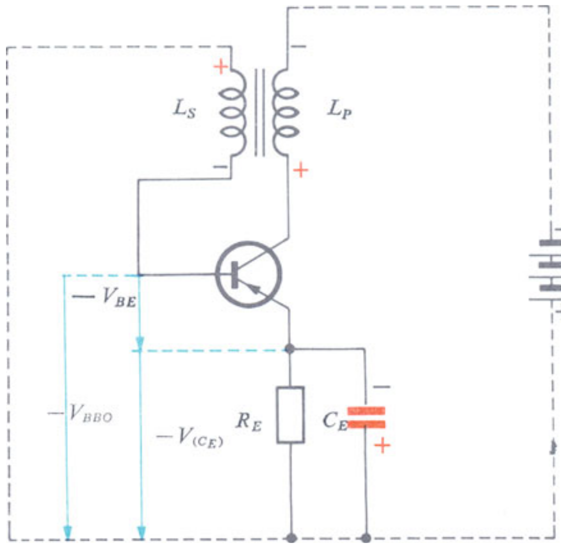


Fig. 545

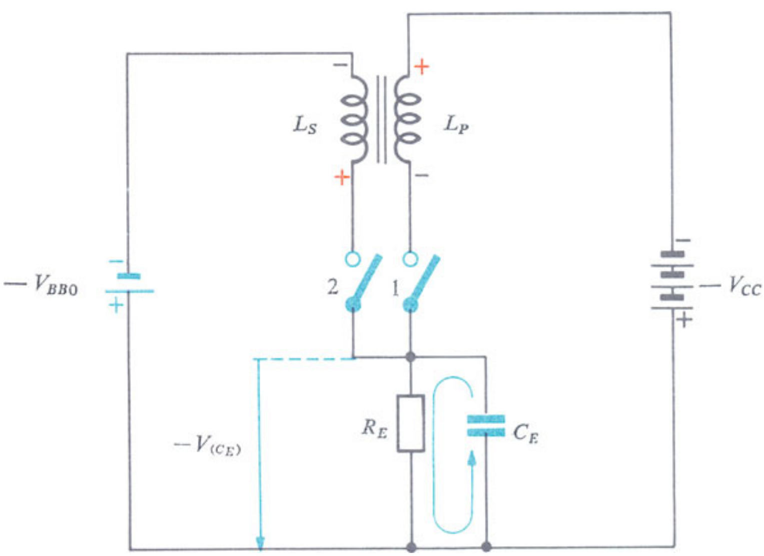


Fig. 546

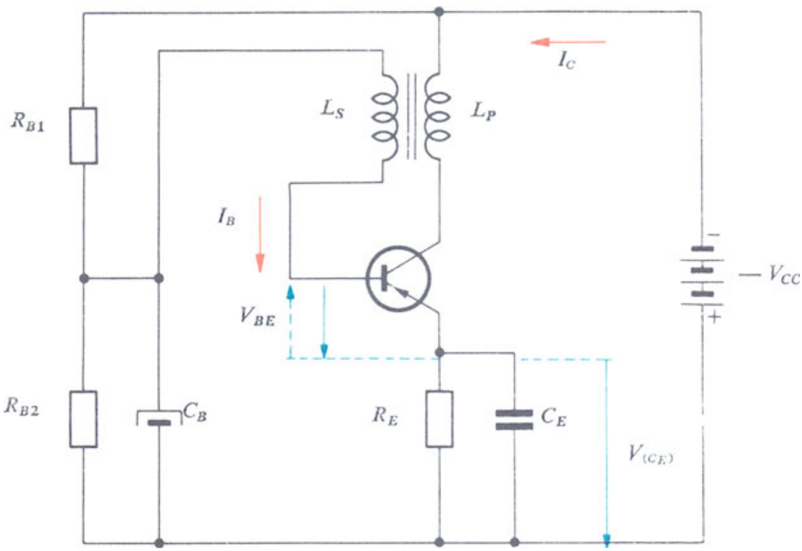


Fig. 547

The quantitative study of an astable blocking oscillator requires the calculation of a certain number of electrical parameters (voltages and currents).

As for all switching circuits, it is possible to distinguish between the period of stable operation and the triggering stages.

Stable periods. The transistor may be in either of two pseudostable states: cut-off or saturation.

The cut-off state requires the simultaneous reverse bias of the two junctions; this condition is assured simply by the quantitative determination of the base–emitter and collector–base voltages of the transistor.

The state of saturation is dependent on the forward bias of the two junctions. It is then more difficult to determine the voltages, and it is better to use the saturation condition defined in the first chapter:

$$I_{CS}/I_{BS} < h_{FE}.$$

Transient periods. Just as in the stable stages, the transistor has two transition periods due to the triggering of the oscillator to the astable blocked condition:

(a) transition from cut-off to saturation;

(b) transition from saturation to cut-off.

A quantitative study of each of these stages would require the use of a large number of equations and of a mathematical treatment quite beyond the scope of this book. It will suffice if we mention the parameters which are affected by the two transient stages and give suggestions for the calculation of the time of triggering.

Transition from cut-off to saturation. When the transistor begins to conduct it is necessary almost at once to respect the condition of saturation. We must therefore determine from instant t_1, corresponding to the triggering, the instantaneous values of the collector and base currents. The ratio of these two currents during the whole of this period should be lower than the current amplification factor of the transistor.

The calculation of the currents requires a complete knowledge of the variations of the voltages induced in the primary and secondary of the transformer.

Transition from saturation to cut-off. We will now proceed to the study of the voltage variations with time (in blue in Fig. 547), at the same time taking into account the influence of the voltages across the primary and secondary of the transformer.

The triggering times are dependent:

(a) on the one hand, on the speed of the transition from cut-off to saturation, that is on the type of transistor and the characteristics of the transformer used;

(b) on the other hand, on the discharge time of capacitor C_E into resistor R_E and on the parameters of the 'transient stage' of the transistor.

Connection of a diode across the primary of the transformer

The capacitor charges inductively during the transient periods; the voltage variations induced in the primary of the transformer may have an amplitude such that the instantaneous value of the collector–emitter voltage becomes greater than the maximum collector–emitter voltage specified by the manufacturer.

The role of the diode (in red in Fig. 548) is therefore to limit these variations and to prevent any danger of the transistor being destroyed.

The influence of the diode on the operation of a circuit of this type was explained in the chapter dealing with the thermal behaviour of transistors.

Regulation of the switching frequency

There are several possible methods of regulating the switching frequency of an astable blocking oscillator; it may be performed by means of:

(a) resistor R_E connected in the emitter of the transistor or by;

(b) the bias bridge.

Adjustment of the emitter resistor. The time constant of the R_EC_E circuit occurs directly in the discharge time of the capacitor through the resistor.

Any variation in resistor R_E has an immediate effect on the discharge time of the capacitor; it therefore modifies the period during which the transistor remains in the cut-off state.

Adjustment of the biasing circuit. If we adjust the base–emitter voltage of the transistor during its period of conduction, simultaneously the threshold of triggering of the astable blocking oscillator is also affected.

In order to understand this effect, all that is necessary is to remember that the transistion from cut-off to saturation is explained by the variations of the voltage across capacitor C_E with respect to the bias voltage of the transistor.

The displacement of the slider of the potentiometer (in red in Fig. 549) causes a variation of the bias voltage of the transistor, resulting in a change in the triggering frequency.

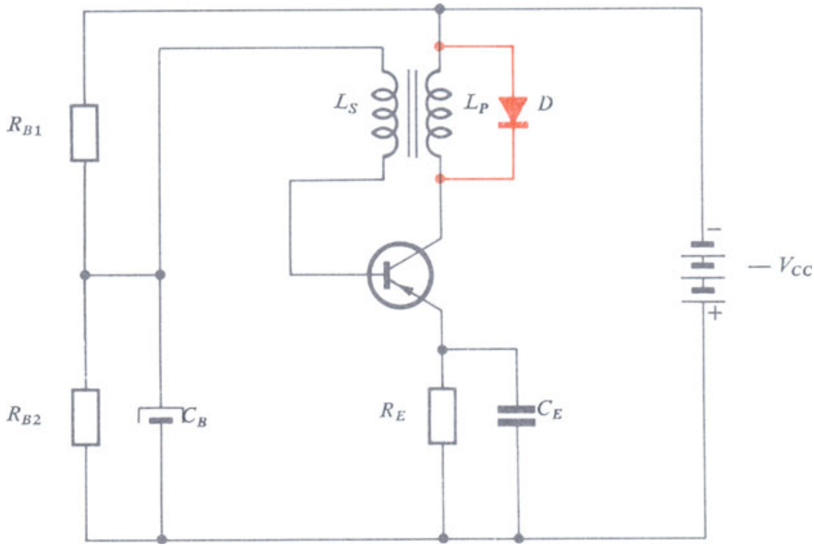


Fig. 548

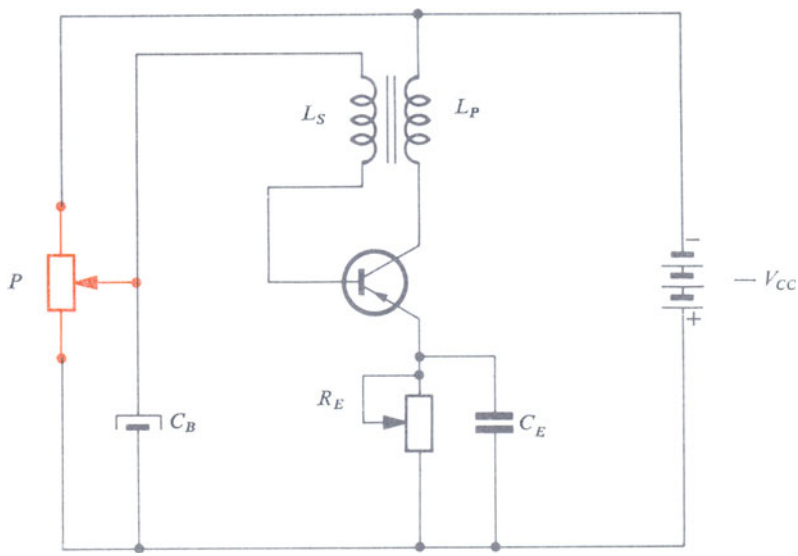


Fig. 549

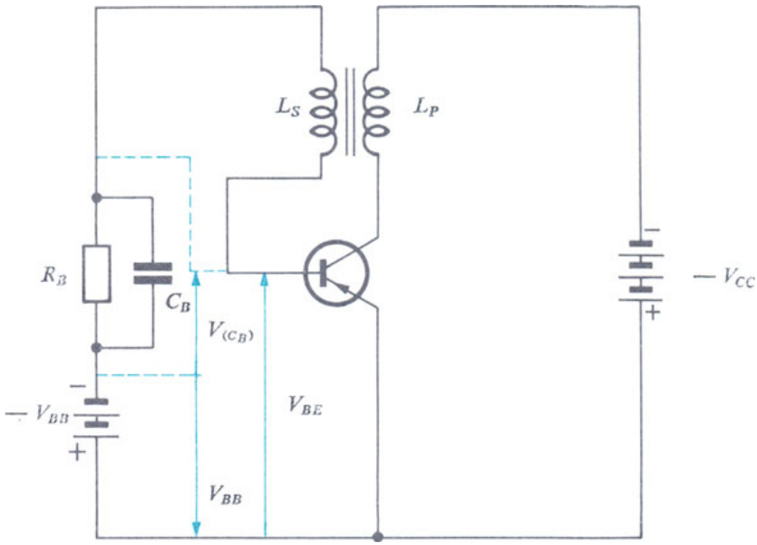


Fig. 550

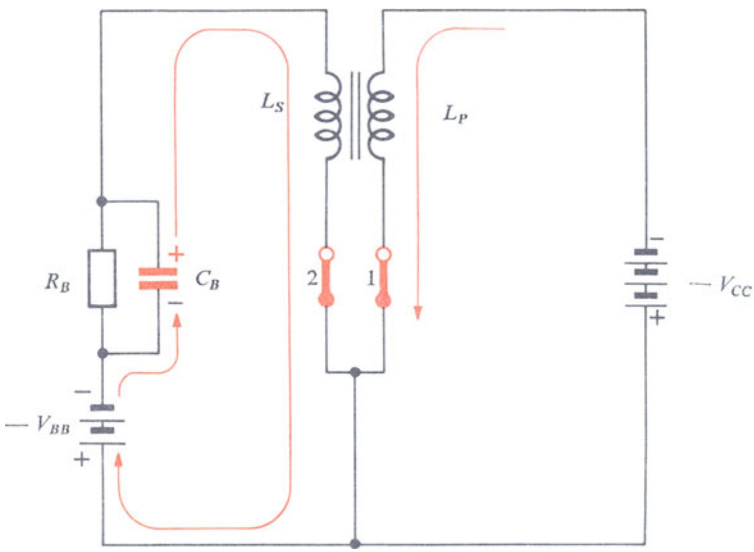


Fig. 551

10.2. Astable blocking oscillator with an *RC* circuit in the base

The second method of producing an astable blocking oscillator is to connect the *RC* circuit in the base of the transistor, as indicated in Fig. 550.

As before, it includes a transistor and a transformer; the polarities of the bias and supply batteries are such that the transistor operates initially in the saturated state.

The expression for the actual voltage applied between the base and emitter of the transistor is given by the relation:

$$V_{BE} = -V_{BB} + V_{(C_B)}.$$

In this expression and in the absence of any operation, the state of initial saturation of the transistor can be illustrated. In fact the capacitor C_B is not charged and the base-emitter voltage of the transistor is equal to, or at least a function of, the battery voltage $-V_{BB}$.

During the period of saturation (Fig. 551) the switches 1 and 2 are closed. Two currents appear in the circuits associated with the transistor:

(a) the collector current flowing in the primary of the transformer;

(b) the base current flowing in the *RC* (R_B, C_B) circuit as well as in the secondary of the transformer.

During this period capacitor C_B charges in accordance with the polarities shown in Fig. 551. The voltage across this capacitor is in opposition to the bias voltage, and in the formula given above for the calculation of the base–emitter voltage, any increase in the voltage across the capacitor produces a decrease in the base–emitter voltage of the transistor.

It is understandable in these conditions that there is a danger that the decrease in the base–emitter voltage may drive the transistor out of saturation and cause a triggering of the stage to the cut-off state.

In Fig. 551 we deliberately omitted the polarities of the voltages appearing across the primary and secondary of the transformer.

In fact, the influence of the transformer is very great during the transition period. It is therefore necessary, in order to determine its importance, to study the first transition period, that is the transition of the transistor from its cut-off state to the saturation state.

The increase in the collector current, corresponding to the instant when the transistor becomes unblocked, produces in the primary of the transformer (L_P) a strong e.m.f. in opposition to the flow of this current (+ sign on the collector side). Because of the direction of the windings, the voltage induced in the secondary is of the opposite sign (− on the base side) and the variation of the base–emitter voltage in this direction speeds up the transition to saturation of the transistor.

When the charging voltage of the transistor becomes very large, the base–emitter voltage of the transistor becomes positive, causing the transistor to change from saturation to cut-off.

Before studying the pseudostable state, it would be interesting to consider briefly the transient stage. The decrease in the collector current produces an e.m.f. in the primary of the transformer in opposition to this variation.

The positive polarity of this voltage now appears on the negative terminal side of the group of supply batteries, and the secondary voltage (+ on the base side) tends to cause the transistor to enter cut-off more rapidly.

During the cut-off state the switches 1 and 2 are open. No current flows in the collector and base circuits. Capacitor C_B discharges through resistor R_B (current shown in blue in Fig. 552).

The voltage across this capacitor decreases at the same time and in the expression:

$$V_{BE} = -V_{BB} + V_{(C_B)}.$$

any decrease in the voltage causes a reduction in the positive voltage applied between the base and emitter. This voltage disappears and the battery produces a new negative voltage across the input circuit of the stage. The transistor begins to conduct again and it very quickly enters into saturation.

The quantitative study of a circuit of this type requires, as before, the definition and use of very complex equations, and it is not our intention to list them in this book. The parameters whose values should be calculated at every instant are the same as those on the previous page.

Regulation of the frequency

In actual circuits it is necessary to be able to vary the triggering frequency of the system. Because of the importance of the discharge of capacitor C_B into resistor R_B on the time of transition from cut-off to saturation, all that is needed in order to regulate the switching frequency is to use a variable resistance R_B.

There are other possible methods of adjusting the switching frequency (mention of the use of a transformer with a saturated core has been omitted deliberately), but it would not be suitable to list them here.

In order to avoid the risk of perforation of the transistor a diode is connected in parallel with the primary of the transformer.

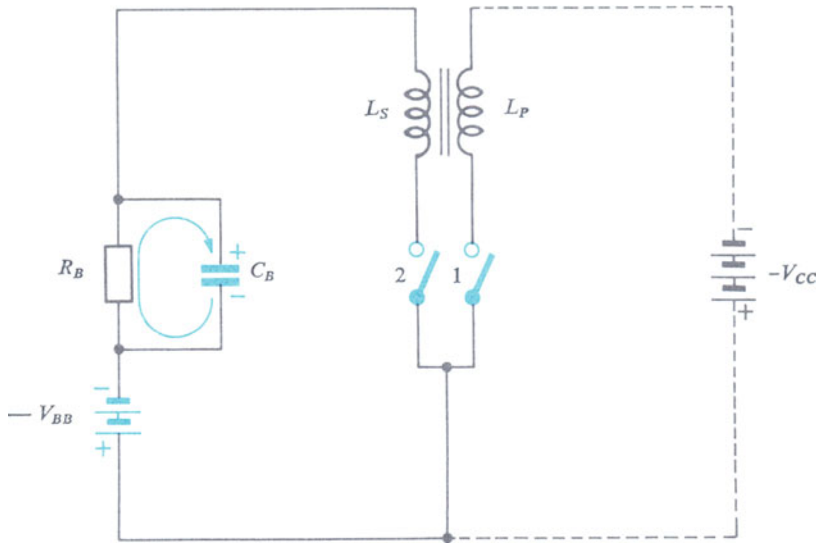


Fig. 552

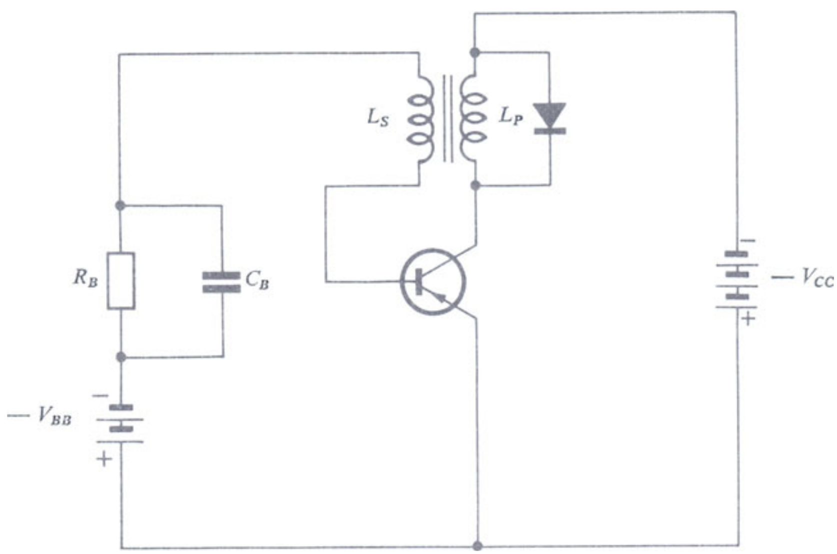


Fig. 553

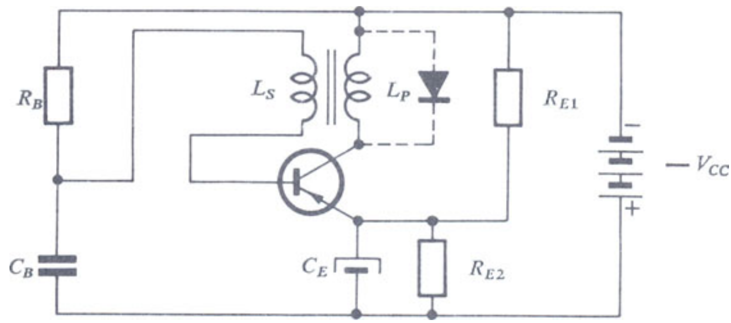


Fig. 554

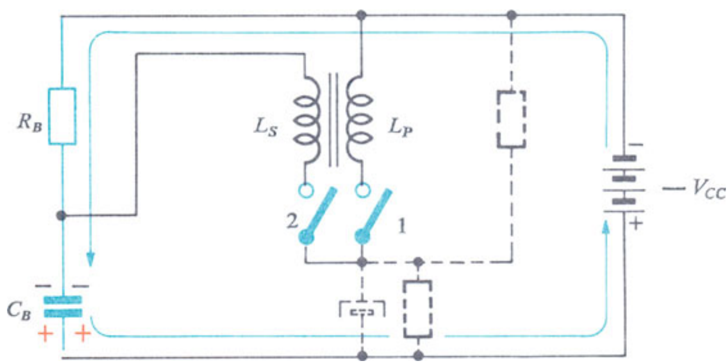


Fig. 555

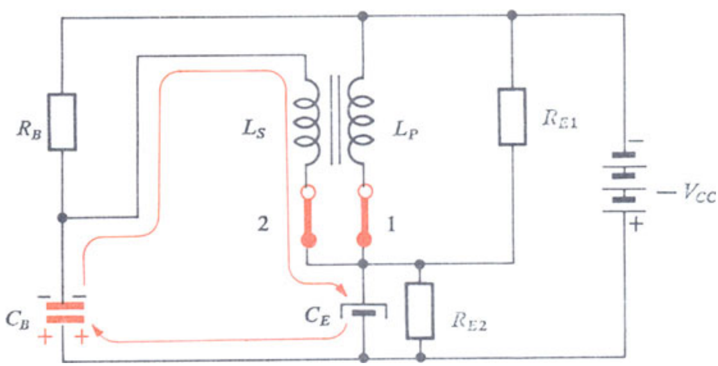


Fig. 556

10.3. Astable blocking oscillator with RC circuit in series

In a third type of astable blocking oscillator the $RC(R_B, C_B)$ circuit is arranged as shown in Fig. 554.

This circuit includes as before a transistor and a transformer. The bias enabling the transistor to be in the conducting state during the initial period is produced by the bridge R_{E1} and R_{E2} connected across the group of supply batteries. The role of the capacitor C_E is to maintain the bias voltage constant.

Period of cut-off

During the cut-off state the switches 1 and 2 are open. As shown in the diagram in Fig. 555 capacitor C_B charges with the polarities $-$ on the base side, and $+$ on the earth side.

The appearance of this voltage neutralizes the permanent voltage across resistor R_{E2}; the base becomes less and less positive and the corresponding voltage disappears. The base then becomes negative and the transistor begins to conduct and very soon reaches saturation.

Period of saturation

The saturation of the transistor is explained by the strong negative voltage appearing across capacitor C_B. The switches 1 and 2 are closed.

It will be seen at once from Fig. 556 that capacitor C_B discharges into a circuit consisting of the switch 2 and resistor R_{E2} in addition to the secondary inductance L_S.

The voltage across this capacitor decreases and the base–emitter voltage becomes less negative. The transistor is no longer saturated and reaches cut-off.

Just as in the two previous circuits, the influence of the transformer on the operation of this stage is of interest. Further, the voltage variations induced in the secondary which strengthen the tendency to triggering are also involved.

Regulation of the frequency

The time constant of the R_B, C_B circuit determines the charging time of the capacitor C_B. In order to regulate the switching frequency of this circuit all that is needed is to make this resistance variable. This can be done either by the use of a potentiometer or by a switching system by means of which resistances of various values can be connected into the circuit.

The presence of the diode across the primary of the transformer is due to the same safety precautions as those applying to the operation of the transistor.

10.4. Astable blocking oscillator

This special circuit, unlike the previous ones, requires the use of a three-winding transformer (L_1, L_2 and L_3 in Fig. 557).

If we consider the various parts of the circuit separately, we find:

(a) the transistor;

(b) the transformer (L_1 and L_2);

(c) the charging and rectifying circuit (on the right of the figure).

If we make the time constant of the CR circuit very long the voltage across the capacitor is practically equal to the peak value of the voltage at the terminals of the L_3 winding of the transformer.

If at a given instant the base–emitter voltage becomes negative, the transistor begins to conduct and immediately becomes saturated. Further, this constant negative voltage is due to the bias circuit R_{B1} and R_{B2} connected at the terminals of the group of supply batteries $-V_{CC}$. The role of the capacitor C_B is to stabilize the bias voltage of this transistor at a constant value.

The principle of operation of a circuit of this nature requires the representation of the different shapes of the voltages and currents shown on the next page.

It may be useful at this stage to give a brief explanation rather than a fuller one.

Any variation in the input circuit (namely, the disappearance of the base current) causes the appearance of strong positive voltage pulses at the terminals of the L_2 winding of the transformer.

This variation is at once transmitted in the reverse direction to the collector of the transistor. The collector current then disappears at the same time as the base current, but a train of oscillations appears in the primary L_1.

The voltage variations following the ringing phenomena produce at a given instant a new negative voltage which tends to trigger the transistor. It then passes from the cut-off to the saturation state.

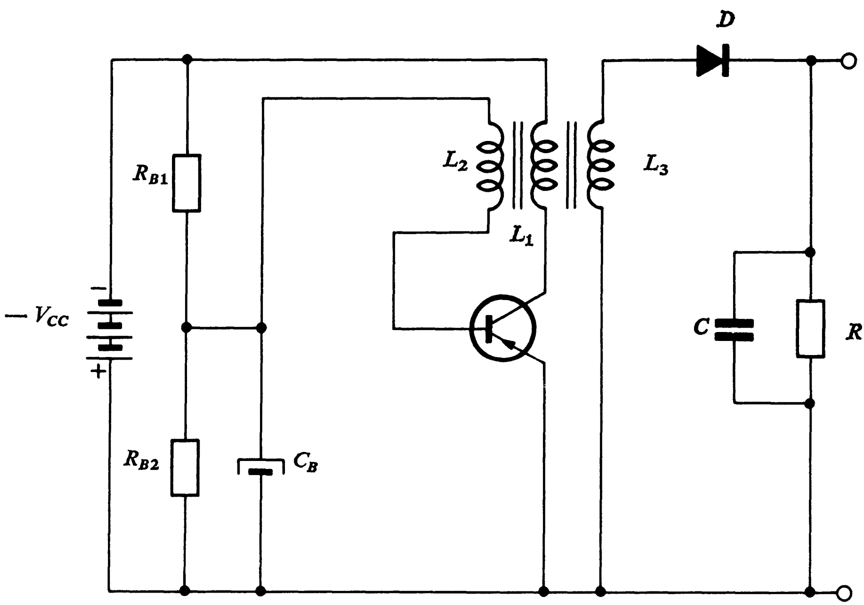


Fig. 557

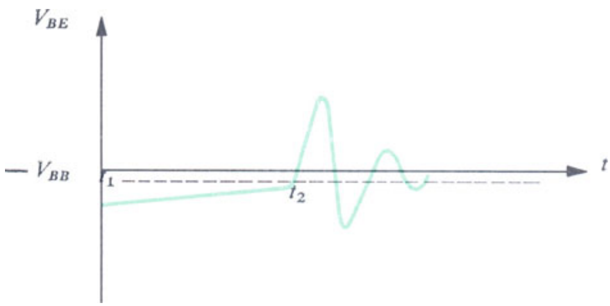


Fig. 558

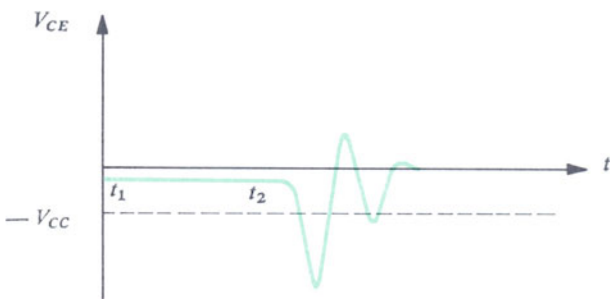


Fig. 559

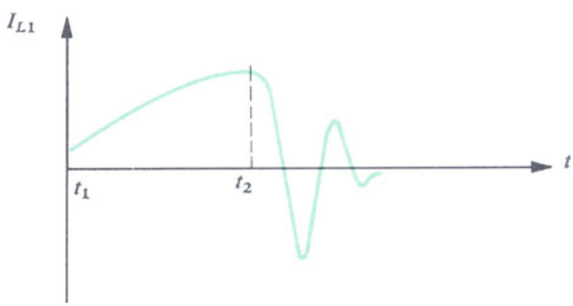


Fig. 560

If there was some doubt, in the previous case, particularly concerning the quantitative study, such doubts are even more justified in this case. The inductive effects causing the transistor to pass from one state to another:

(a) from cut-off to saturation;

(b) from saturation to cut-off;

require the use of equations that are quite beyond the scope of this book.

In order to confirm the explanations given on the previous page, it is interesting to examine the shapes of the voltage and current in the three figures.

Figure 558 shows the variations with time of the base–emitter voltage; Fig. 559 the corresponding variations of the collector–emitter voltage and Fig. 560 the shape of the current in the primary of the transformer.

From instant t_1 to instant t_2 the transistor is saturated and its base–emitter voltage is negative although decreasing.

The collector–emitter voltage of the transistor is permanently near to zero.

The current in the primary of the transformer increases to a maximum at instant t_2.

At this instant the state of the transistor changes from saturation to cut-off.

A train of ringing appears in the three figures; it is explained by the very large variation of the current in the inductance L_1.

If we examine Fig. 558 we see that the base–emitter voltage of the transistor reaches a new negative value; in these conditions it returns to saturation, and so on.

In this type of construction it is advisable to consider the possibility of using two transistors in the same circuit, and in this way of obtaining a symmetric astable blocking oscillator.

10.5. Triggered blocking oscillators

The second function mentioned when discussing the principles of switching consists in the generation of a pulse with as nearly as possible an ideal shape from a pulse with an insufficiently steep leading edge.

This 'reshaping' of pulses may be performed:

(a) either by means of the monostable multivibrator;

(b) or by means of the triggered blocking oscillator.

Principle of operation of the triggered blocking oscillator

The triggered blocking oscillator consists of two main components; the transistor and a transformer.

Unlike the astable blocking oscillator in which the state of permanent operation might well be saturation, the triggered blocking oscillator is always in the state of conduction in the absence of a pulse.

It will be seen in the synoptic diagram of Fig. 561 that the role of the transformer (red rectangle *A*) is therefore to accentuate the tendency of the transistor to trigger. Unlike the astable blocking oscillator, this type of circuit has no *RC* circuit and therefore it cannot trigger unaided from one state to the other.

Basic diagram of a triggered blocking oscillator. The basic diagram of a triggered blocking oscillator shows first of all the two components of the synoptic diagram, namely, the transistor and the transformer.

The transistor is biased by a resistance bridge (R_{B1} and R_{B2}) connected across a group of batteries $+V_{BB}$. The high-value capacitor C_B prevents any fluctuation of the base voltage of the transistor.

In the absence of a signal across this stage the transistor is cut-off.

The base is positive with respect to the emitter, and the collector is negative with respect to the base. A negative pulse applied to the base brings the transistor out of cut-off. The transistor begins to conduct and the variation of the collector current induces voltage variations in the primary and secondary of the transformer which strengthen the triggering effect.

The state of saturation is maintained during the whole duration of the pulse. When this pulse disappears the base–emitter voltage again becomes positive and the transistor is driven once more into cut-off.

Different types of feedback

The astable or triggered blocking oscillators both operate on the principle of feedback, that is the return of a fraction of the energy from the output to the input circuit.

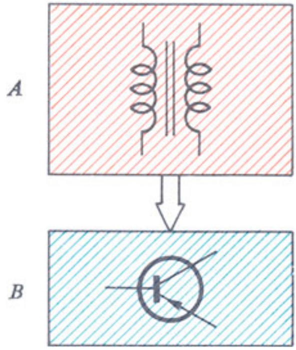


Fig. 561

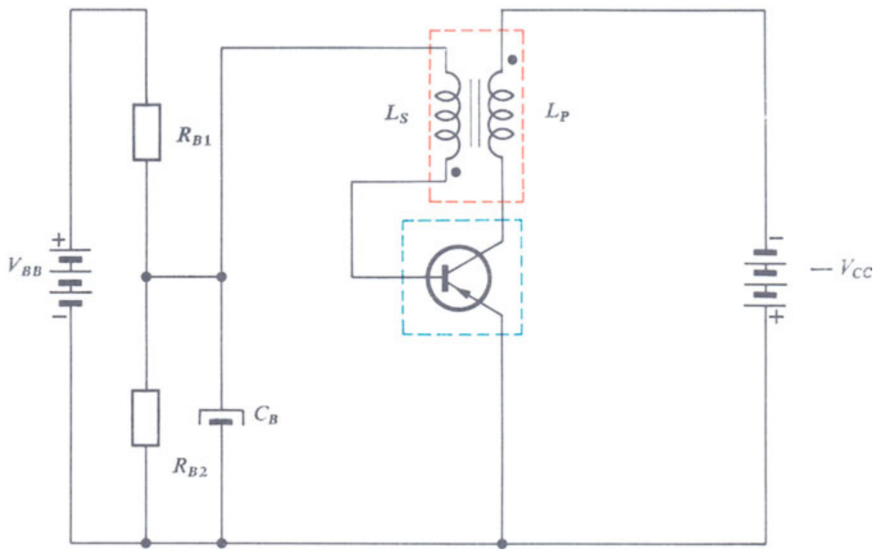


Fig. 562

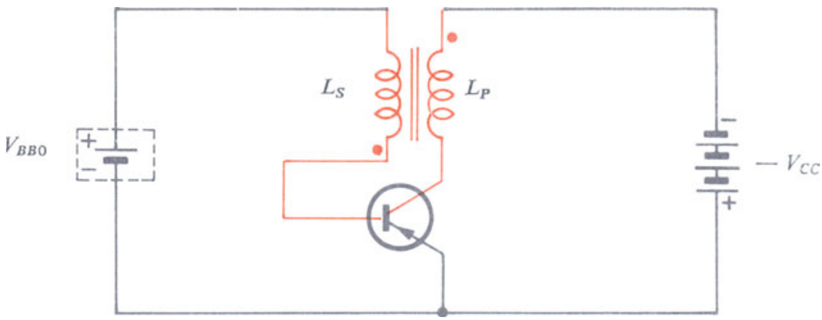


Fig. 563

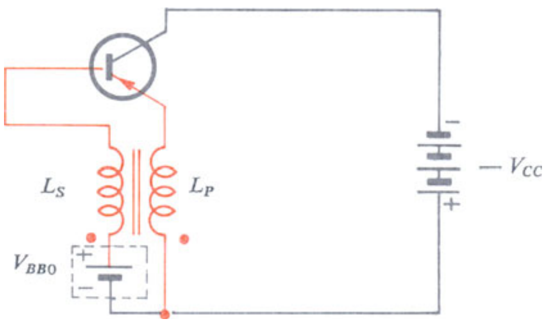


Fig. 564

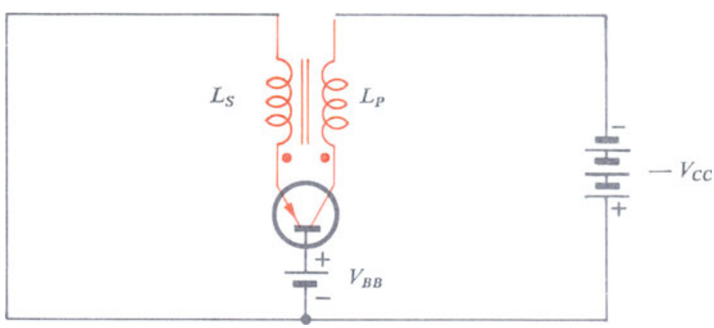


Fig. 565

There are three main types of feedback:

(a) feedback between the collector and base;

(b) feedback between the base and collector;

(c) feedback between the collector and emitter.

Feedback between the collector and base. In this type of feedback the primary is connected to the collector circuit and the secondary is used to excite the transistor.

This type of feedback is illustrated in Fig. 563. The output energy is the energy available at the terminals of the primary of the transformer, and the control energy, during the transient period, is a function of the variations of the voltage and current in the secondary of the transformer.

In addition to these two dynamic components, the circuit must be provided with supply and biasing sources $-V_{CC}$ and V_{BBO} respectively.

Feedback between the emitter and base. Another method of speeding up the transient effects is to connect the transformer between the base and emitter circuits of the transistor. The primary is connected in series with the emitter and the secondary direct to the base of the transistor.

The phase conditions are now different; the signals returned to the input should be in phase with the output signals, and we must therefore use both the transformer windings for this purpose.

Feedback between the collector and emitter. In Fig. 565 the transformer is connected between the collector and emitter of the transistor, which then operates in common base.

The primary of the transformer (L_P) charges the collector of the transistor and the secondary L_S constitutes the driving circuit.

The input and output signals should be in phase, which explains the positions of the two red dots in the figure.

Whatever type of feedback we are considering, the transistor should always, in the absence of an input pulse, be in the cut-off state.

The principle of operation of an oscillator of this nature is always the same; the transformer should strengthen the tendency to variation in either the first transient phase (transition from cut-off to saturation) or in the second (transition from saturation to cut-off). It is therefore very important to respect the phase conditions shown in the three figures.

10.6. Triggered blocking oscillator with feedback between the collector and base

The circuit in Fig. 566 consists primarily of a transistor and transformer. The direction of the windings (L_P and L_S) of the transformer is such that any variation of energy in the primary is immediately transmitted to the secondary in the reverse direction.

The stage is supplied by the group of batteries $-V_{CC}$ and is biased by the group of batteries $+V_{BB}$. The bias bridge R_{B1} and R_{B2} serves to adjust very accurately the direct voltage applied between the base and emitter of the transistor. The purpose of capacitor C_B is to maintain the bias voltage absolutely constant.

Stable state

In the absence of a pulse, the transistor is in the cut-off state. This operating condition can be understood better by reference to the simplified circuit of Fig. 567.

The various components constituting the biasing circuit of the transistor may be connected to a simple battery supplying a voltage $+V_{BBO}$ between the base and collector of the transistor.

This cut-off state constitutes one of the basic differences between the triggered blocking oscillator and the astable blocking oscillator.

The principle of operation of a circuit of this type will be discussed more fully later, but it might be appropriate at this stage to give a brief account of it.

Principle of operation of a triggered blocking oscillator. If a negative voltage pulse is applied at a given instant between the base and emitter of the transistor it brings it into conduction.

The variation of the collector current produces an induced back e.m.f. in the primary of the transformer, producing a voltage across the secondary; this voltage strengthens the state of conduction of the transistor.

Transition from cut-off to saturation occurs very rapidly; it is caused by the presence of a negative voltage pulse supplied by an external generator.

When this pulse disappears the transistor tends to become cut-off again. The collector current decreases and the back e.m.f. appearing at the primary of the transformer produces a voltage in the secondary which strengthens this trigger effect.

The transistor then rapidly reaches cut-off and remains in that state until the arrival of a second negative pulse.

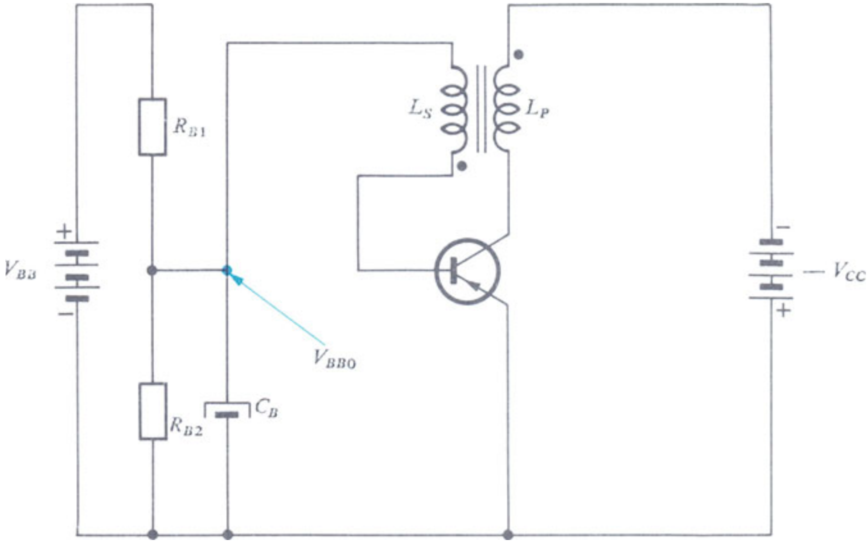


Fig. 566

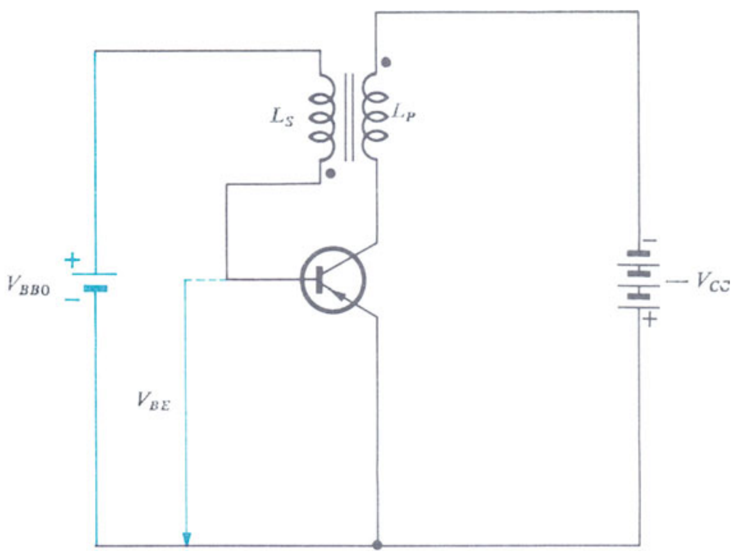


Fig. 567

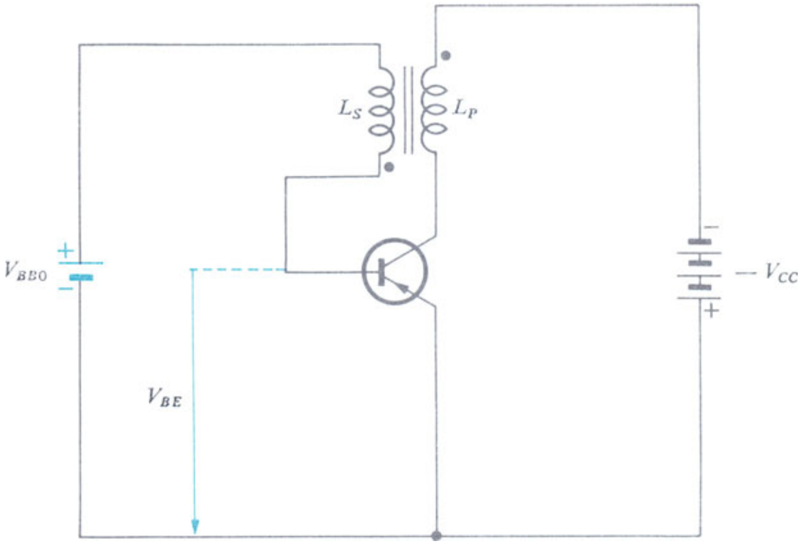


Fig. 568

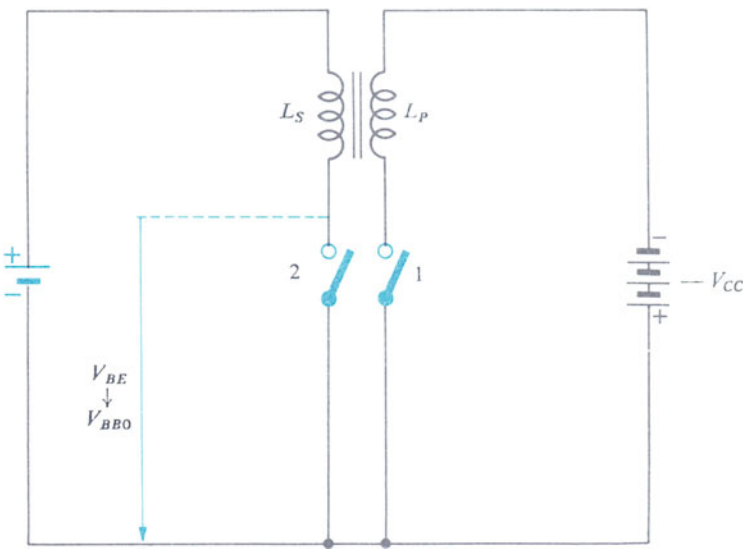


Fig. 569

Cut-off period

Reverting to the stable state discussed above, we see that the transistor is permanently in the cut-off state in the absence of any external information applied to the input of the stage.

In Fig. 568 the collector is very strongly negative with respect to the emitter; the corresponding voltage is near to $-V_{CC}$ because of the absence of current in this circuit.

The base of this transistor is positive with respect to the emitter. The biasing bridge has with advantage been replaced by a battery (V_{BBO}), the positive pole of which is connected to the base of the transistor. In these conditions the base–emitter voltage is equal to the imaginary bias voltage since there is no current flowing in the circuit, that is:

$$V_{BE} = V_{BBO}.$$

The collector and emitter of the transistor are both negative with respect to the base. The cut-off condition is therefore assured.

Simplified equivalent circuit in the cut-off state. The study of the operation of a triggered blocking oscillator requires not only a knowledge of the voltages appearing at the various active points of the stage with respect to earth, but also the calculation of the currents flowing through the various meshes of the circuit.

When the transistor is cut-off, the input and output circuits offer a very high resistance to the current. It is then possible, if we ignore the residual currents, to liken the transistor to a group of two switches in the open position.

In Fig. 569 the input circuit contains, in addition to the imaginary bias battery and the secondary of the transformer, a switch (2) in the open position. The output circuit consists of a group of supply batteries $-V_{CC}$, the primary of the transformer (L_P) and a switch (1) in the open position.

Because of the position of these two switches, no current can flow in the collector and base circuits of the transistor.

The base–emitter voltage is therefore equal to the imaginary biasing voltage, and the collector–emitter voltage is near to the supply voltage. These two equalities are shown in the following expressions:

$$V_{BE} = V_{BBO} \quad \text{and} \quad V_{CE} = -V_{CC}.$$

If no pulse is applied to the input terminals, the triggered blocking oscillator remains in the cut-off state.

Transition from cut-off to saturation

If at a given instant a negative voltage pulse is applied to the input terminals of the stage the transistor begins to conduct and very soon becomes saturated.

Before discussing these transient effects in detail, it would help us to explain the increasing tendency for the circuit to trigger, if we discussed briefly the shape and amplitude of the drive signal.

The transistor is permanently in the cut-off state and this condition may be explained by the presence of a positive voltage $+V_{BBO}$ between the base and emitter. A current can appear in the collector circuit only if the base becomes negative with respect to the emitter; it is therefore essential to apply to the input terminals a voltage pulse greater than the cut-off voltage of the transistor.

In the initial cut-off state the base–emitter junction has a relatively thick barrier. A certain time, designated the delay time of the current in the previous chapter, must elapse before the collector current can increase. Account must always be taken of this delay time, and the drive voltage must be given a value even higher than the permanent positive voltage at the end of this interval.

First transient stage. When the above conditions are respected, the transistor begins to conduct and a current appears in the primary of the transformer. A back e.m.f. with polarities as shown in Fig. 571 tends to oppose the increase in this current.

Because of the phase relation between the primary and secondary of this transformer, the voltage induced in the secondary strengthens the negative voltage of the base and causes the collector current to increase more rapidly.

The role of the transformer appears very clearly in this figure; it strengthens the tendency of the collector current to increase and speeds up the transition of the transistor to saturation.

It would be possible, by taking account of the different electrical parameters of the transistor and of the transformer, to produce a certain number of equations for the calculation of the instantaneous values of the currents and voltages in this stage. However, the relative complexity of these equations and the necessity for the introduction of concepts peculiar to this type of circuit, make it advisable to omit them.

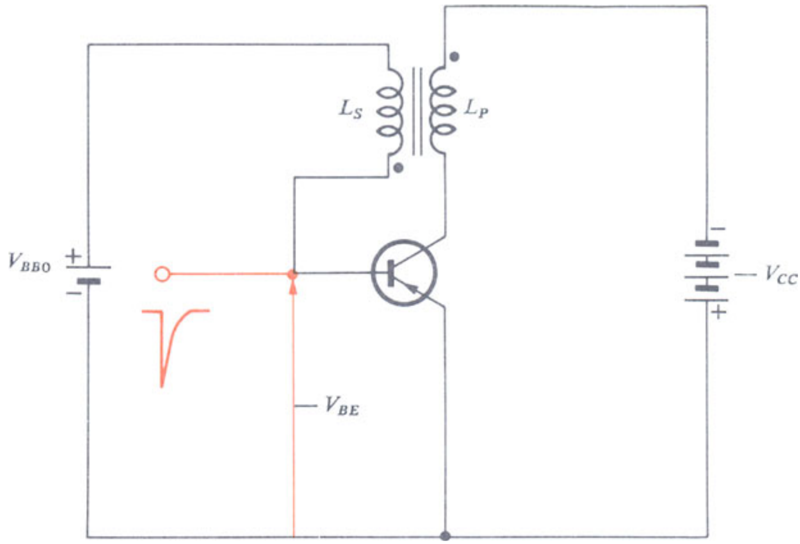


Fig. 570

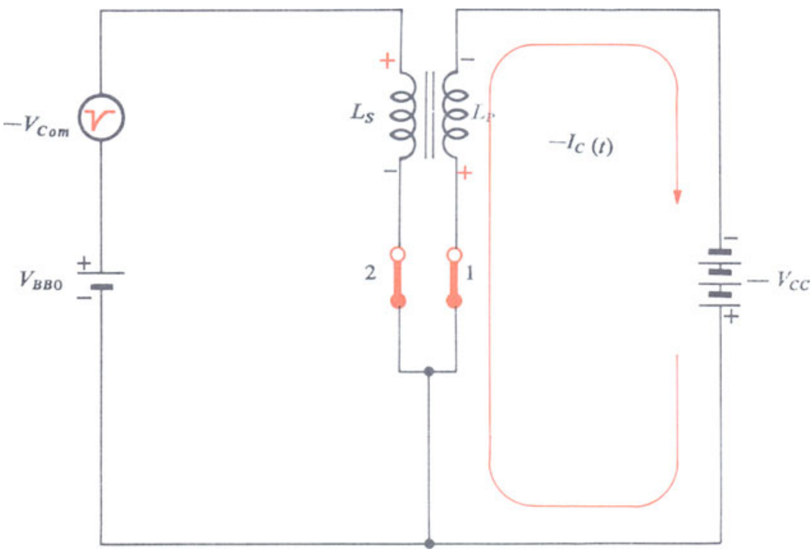


Fig. 571

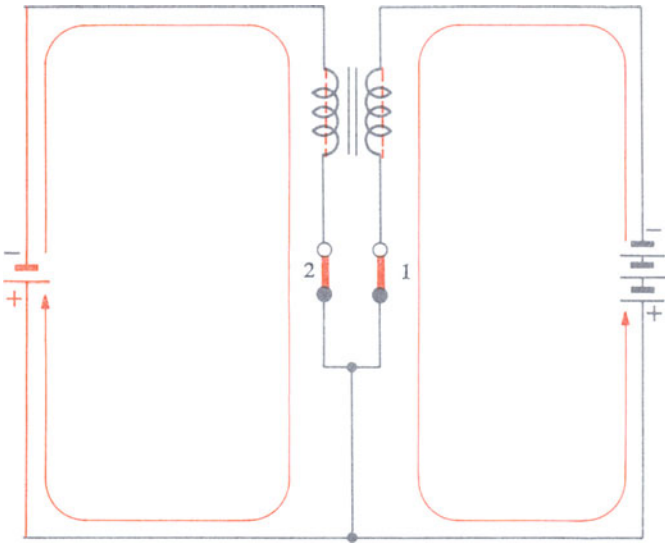


Fig. 572

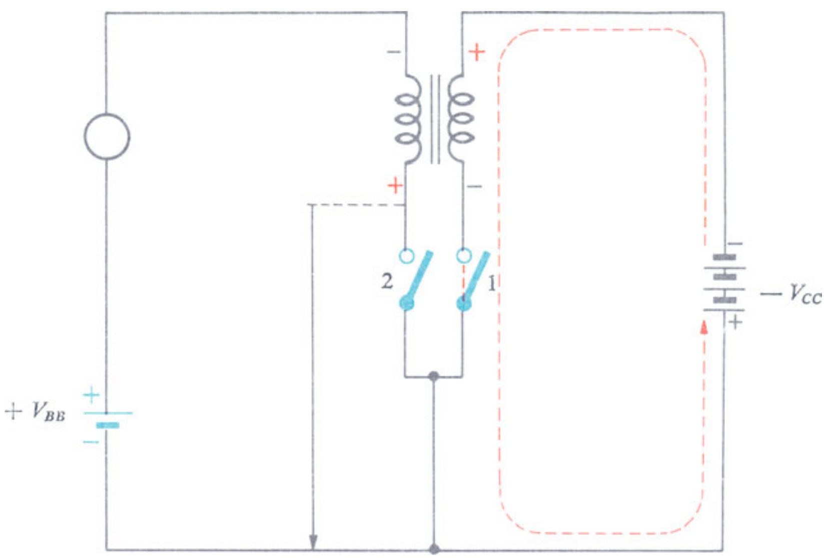


Fig. 573

Period of saturation

When the transistor reaches saturation, switches 1 and 2 are closed (the input and output resistances are very small).

The saturation state may be likened to a stable state of operation. The base and collector currents are constant and the primary and secondary windings may be considered as negligible resistive elements. Figure 572 shows the equivalent network in the saturated state. The collector and base currents may easily be calculated. On the other hand, the collector–emitter and base–emitter voltages are near to zero.

The stage remains in this operating state during the whole duration of the drive pulse.

Second transient stage. When this pulse disappears the positive biasing voltage causes an immediate decrease in the collector current of the transistor (the desaturation time of the transistor has been deliberately ignored).

This tendency of the collector current to disappear produces a back e.m.f. in the primary of the transformer tending to oppose any decrease in this current.

The voltage induced in the secondary is in phase opposition (Fig. 573); this voltage now strengthens the voltage supplied by the biasing battery and accelerates the decrease in the current in the input circuit; it consequently causes the current in the collector circuit to disappear sooner.

The role of the transformer is all-important in this second transient stage just as it was in the first; the transistor is no longer saturated and almost at once reaches cut-off.

Switches 1 and 2 representing respectively the input and output circuits of the transistor are in the open position. No current can flow through the collector and base of the transistor; the triggered blocking oscillator is again in its stable period.

For every voltage pulse applied at the input of this stage there is a corresponding pulse at the output. The regeneration function of the control information is therefore fulfilled; the triggered blocking oscillator is therefore almost identical with the monostable multivibrator.

Influence of a diode across the primary of the transformer

In the circuit of Fig. 574 the transistor is driven by the primary of a transformer. During the stable periods (state of cut-off) or the pseudo-stable state, this transformer offers a very low resistance to the flow of current and its primary is often considered as a simple short circuit.

During the transient periods we have already referred to the back e.m.fs tending to oppose the variations of the collector current. The resulting voltages may be very high, and they must therefore be limited to values lower than the maximum collector–emitter voltage specified by the manufacturer of the transistors.

The reader is referred to the paragraph dealing with the switching of a transistor driven by an inductance; the possibility of an instantaneous value of collector–emitter voltage which might be very much greater than the limiting value $-V_{CEM}$ was mentioned there.

What steps can be taken to avoid the risk of perforation of the transistor?

The use of a diode in parallel with the primary of the transformer is certainly the most effective solution. This diode prevents the appearance of a voltage surge and consequently maintains the collector–emitter voltage at a maximum value that never exceeds the supply voltage $-V_{CC}$.

In all industrial circuits when the transistor is associated with an inductive load, a diode is connected at the terminals of the primary of the transformer.

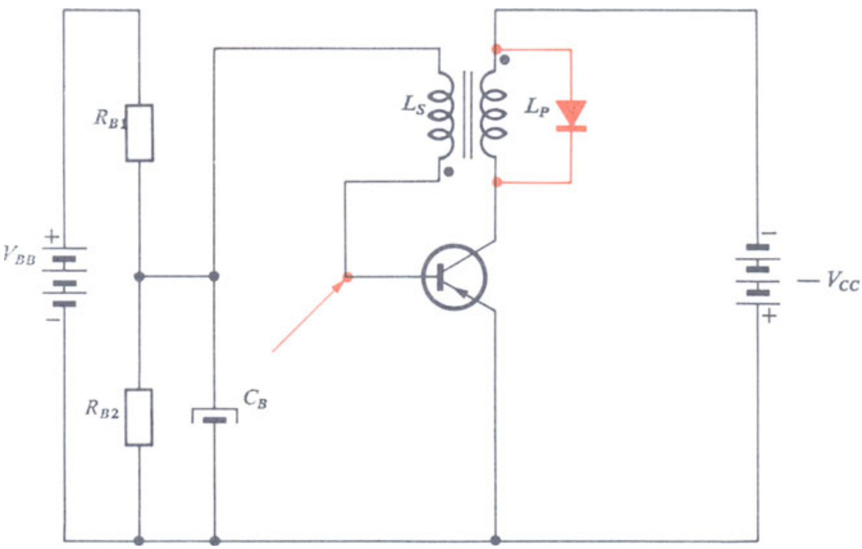


Fig. 574

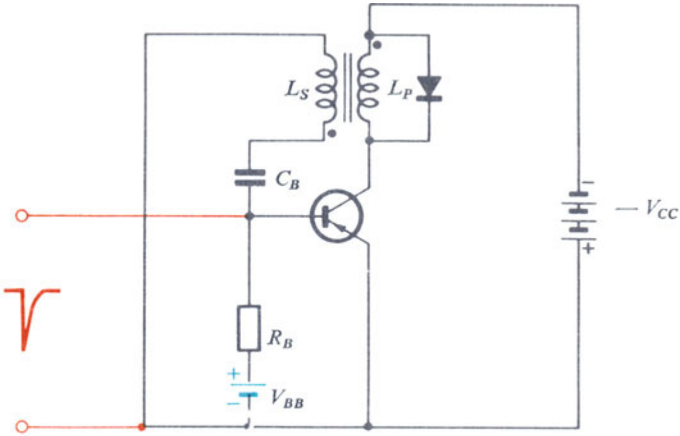


Fig. 575

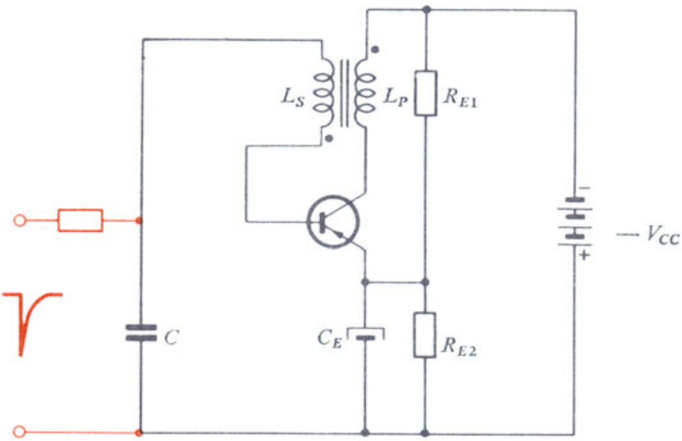


Fig. 576

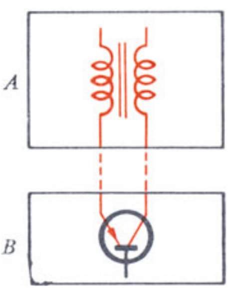


Fig. 577

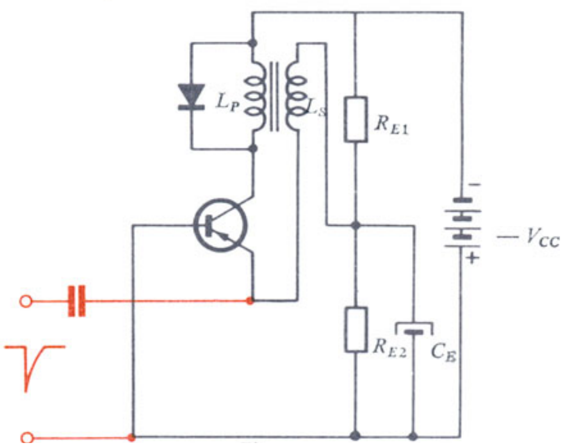


Fig. 578

10.7. Triggered blocking oscillator excited by the base

The circuit shown in Fig. 575 operates in the conditions discussed earlier. It is however different from the basic stage already discussed in that the bias is produced by a simple circuit consisting of a battery ($+V_{BB}$) and a resistor (R_B).

The secondary of the transformer is connected to the base by means of a capacitor C_B.

In the absence of a negative pulse at the input terminals of the stage, the base is positive with respect to the emitter and collector and the transistor is in the cut-off state.

In the simplified equivalent circuit of this arrangement, the transistor appears in the form of two open switches.

As soon as a negative pulse appears, the transistor begins to conduct in accordance with the provisions mentioned above, namely:

(a) a negative voltage pulse larger than the biasing voltage;

(b) a voltage still high enough at the end of the current delay time to allow the appearance of a current in the collector circuit.

The phase relations between the primary and secondary of the transformer are always the same; the voltage induced in the secondary is therefore always in phase opposition to that of the voltage appearing across the primary.

During the first transient phase (transition from cut-off to saturation) the voltage induced in the secondary strengthens the tendency for the appearance of a current in the collector circuit and accelerates the triggering process.

During the second transient stage the primary and secondary voltages are reversed with respect to the first case and the transformer accentuates the triggering from saturation to cut-off.

It would also be possible here to develop a more detailed quantitative study, taking into account especially the action of the capacitor C_B on the operation of the circuit.

In practice it is sufficient to define the triggering times associated with the particulars supplied by the manufacturer of the transistors and to calculate the electrical parameters of the transformer in order that, except for the transient periods, the saturation condition should be maintained during the whole period of conduction of the transistor.

10.8. Triggered blocking oscillator with biasing circuit in the emitter

With signals having a relatively high repetition rate, account must be taken of the desaturation time of the transistor which may then be able to limit the performance of the stage as far as the switching frequency is concerned.

The diode connected to the primary prevents voltage surges between the collector and emitter of the transistor.

A second possibility for the triggered blocking oscillator (with feedback between collector and base) is shown in Fig. 576. It consists in changing the position of the biasing circuit of the base and locating it between the emitter of the transistor and earth.

10.9. Triggered blocking oscillator with feedback between the collector and emitter

The synoptic diagrams in Fig. 577 show the principle of operation of an oscillator of this type.

The primary of the transformer is connected in the collector circuit and the secondary in series with the emitter of the transistor.

In the absence of pulses the transistor is in the cut-off state and the transformer behaves as a purely passive element.

As soon as a positive pulse is applied to the input terminals of the circuit, the transistor begins to conduct and the transformer speeds up the triggering. The same process will occur in the second transient stage.

Figure 578 shows an actual circuit operating with this type of feedback.

The primary of the transformer is connected in the collector circuit, and the secondary between the emitter and the biasing circuit. This latter circuit consists of a resistance bridge connected directly to the terminals of a group of supply batteries. The role of capacitor C_E is to maintain the emitter–base voltage constant.

In the absence of input pulses the transistor is cut-off and no current flows through the input and output circuits.

If a positive pulse is applied between the emitter and earth (that is between the emitter and base of the transistor) the transistor begins to conduct and very soon enters saturation. This state of saturation is maintained during the whole duration of the pulse.

In order that triggering may take place it is, as before, essential that the amplitude of the drive pulse should be greater than that of the biasing voltage of the transistor.

The advantage of this type of circuit is that it is very suitable for the reshaping of the positive drive pulses.

The diode connected across the primary of the transformer prevents any voltage surges between the collector and base of the transistor.

There are many different types of triggered blocking oscillators, but in every case the principle of operation remains the same. In addition to certain components (resistors, capacitors, diodes and auxiliary transistors), they always include a transistor and a capacitor connected in accordance with the three basic feedback principles mentioned at the beginning of this chapter.