RRE JOURNAL

NUMBER 45 OCTOBER 1960

R. R. E. JOURNAL

CONTENTS

Issued by the Royal Radar Establishment, Ministry of Aviation

No. 45 October 1960

Wha is enemie to science and cunning,
Bot Ignorants, that understandis nocht?
Whilk is sa Nobill, sa precious, and sa ding,
That it may not with eirdlie thing be bocht.
Weill wer that man over all uther, that mocht
All his lyfe dayis in perfite studie wair
To get science; for him neidis na mair.

Robert Henryson

EDITORIAL NOTE

This issue of the R.R.E. Journal concludes the reprint of the papers issued at a series of Lectures on "The Fundamental Principles of Transistor Circuits" given in the R.R.E. Physics Department during March and April, 1959. To preserve these Lectures in a suitable form and at the same time to introduce them to a wider audience seemed a task that the R.R.E. Journal was well fitted to perform, and accordingly the Editor prevailed on Mr. S.W. Noble and his colleagues to allow the present reprinting to take place. No attempt has been made to bring the Lectures up to date, and they must be read as expressing the state of the subject in March 1959. Similarly, no effort has been made to alter the lecture style into that of a formal paper.

Lectures 1 to 3 of "The Fundamental Principles of Transistor Circuits" have appeared in the last issue of the R.R.E. Journal, together with the first three lectures of Dr. A.F. Gibson's earlier course on "The Physics of Transistors: the titles and authors of these lectures appear on the next page.

H.G.

The Physics of Transistors

Lecture 1.	Introduction	A.F. Gibson
Lecture 2.	Basic Ideas	A.F. Gibson
Lecture 3.	The Theory of P/N Junction Diodes	A.F. Gibson
List of Symbols		

The Fundamental Principles of Transistor Circuits

Lecture 1.	The Elementary Theory of Transistors*	S.W. Noble
Lecture 2.	Transistor A.C. Amplifiers	R.C. Bowes
Lecture 3.	Tuning Circuits and Waveform Generators	R.C. Bowes

*Delivered in two parts

The above have appeared in the R.R.E. Journal No. 44 for April 1960.

LECTURE 4. TRANSISTOR SINE-WAVE OSCILLATORS

by P.J. Baxandall

1. INTRODUCTION

Most of the material in Sections 3, 4, 6, 7, 8 and 9 below is covered in a paper* presented by the author, in May 1959, at the I.E.E. International Convention on Transistors and Associated Semiconductor Devices. In particular, Sections 7 and 8 below give a considerably shortened treatment compared with that in the I.E.E. paper.

However, in preparing the draft for this R.R.E. Journal article, the opportunity has been taken of adding a small amount of material that has not appeared before. The main additions are Sections 4.6 and 11.3, plus some new material in Section 7.4.

2. GENERAL COMMENTS ON TRANSISTOR OSCILLATORS

L - C oscillators are often regarded as a combination of an amplifier, a resonant circuit and a feedback circuit. The latter usually involves the equivalent of a transformer, which may be formed by providing a tapping on the resonant circuit or by means of a separate feedback winding. At high frequencies additional elements may be included to compensate for unwanted phase changes in the amplifier.

With a given transistor, it is possible to obtain oscillation at frequencies approaching that at which the transistor, with optimum matching of input and output impedances, gives unity power gain, and it should be noted that this frequency may be appreciably higher than the alpha cut-off frequency. At such high frequencies, class 'A' operation is essential, since any form of non-linear operation inevitably gives reduced transistor power gain; and the efficiency inevitably is low, since nearly

*BAXANDALL, P.J.: "Transistor Sine-Wave LC Oscillators", Proc. I.E.E., Vol. 106, Part B Supplement No. 16, pp. 748-758 and Discussion on page 791.

all the transistor output power must be fed back to the input.

A useful discussion of the above aspects appears in Section 14 of Reference 1.

In the vast majority of applications of transistors, the working frequency is fortunately much lower than the maximum oscillation frequency of the transistor, and this allows far more scope for the design of ingenious circuits to meet different performance requirements as well as possible.

When designing an oscillator, there are various attributes that may need to be considered, amongst which are :-

(a) Power output
(b) Efficiency
(c) Distortion
(d) Frequency stability
(e) Range of frequency variation required
(f) Ambient temperature range to be coped with
(g) Simplicity and cost

The choice of circuit depends, of course, on which factors are considered to be of the greatest importance, since it is impracticable to design a cheap and simple oscillator that is nearly 100% efficient, has less than 0.1% distortion, is stable to 1 in 10^4 over a 30°C temperature range, is adjustable over a wide band of frequencies and is able to work into loads of widely varying impedance! Nevertheless, it is remarkable the extent to which, using transistors, a number of apparently conflicting requirements can be reconciled.

After thinking about and experimenting with transistor oscillators a good deal during the past year or two, the lecturer has become convinced that most oscillator application requirements can be met much more satisfactorily and simply using transistors than they can using valves.

Apart from the obvious advantage of small size, the main virtues of transistor oscillators may be summarized as follows:-

(a) Relatively high power-conversion efficiency even at power levels of only a few milliwatts. Three factors contribute to this:-

(i) There is no heater to waste power. (In a valve oscillator with an output of, say, 100 mW, the heater will usually consume at least a watt, whereas a transistor oscillator with the same output may well have a total power consumption of less than 130 mW.)

(ii) Transistors, unlike valves, continue to operate satisfactorily even at very low values of h.t. current and voltage. If, for example, an output of 1 mW is required, a valve oscillator might consume, say, 1 mA at 100 volts, i.e. 100 mW. A transistor oscillator with an output of 1 mW, on the other hand, might consume 1 mA at 2 volts, or even less. (It is reported in Section 14, page 12, of Reference 1 that a transistor oscillator has been successfully operated at an h.t. voltage of 0.18 volt and an h.t. current of 1.6 µA, i.e. an h.t. power of just under 0.3 µW. The term "h.t." seems of questionable suitability under these conditions!)

(iii) Transistors bottom to a much lower percentage of the h.t. voltage than do valves, so that in class 'C' oscillators, and to an even greater extent in the novel circuits described in Sections 6, 7 and 8, the percentage of the h.t. power that is dissipated at the collector of a transistor may often be made much smaller than the percentage dissipated at the anode in a valve class 'C' oscillator.

(b) Because of the very low total power dissipation in an oscillator such as the local oscillator of a transistor radio receiver, and, in most circumstances, the correspondingly low dissipation in the other associated circuits, the temperature rise after switching on is very small indeed, perhaps less than 1°C, so that, compared with valve apparatus, the problem of avoiding frequency drift during warm-up is virtually non-existent. This was strikingly illustrated in an experimental V.H.F. F.M. receiver, using transistors throughout, designed by R.C. Bowes. The local oscillator of this receiver operates at about 90 Mc/s, using the circuit of Fig. 14, 10 (a) of Reference 1. It was intended to provide A.F.C., using a reverse-biassed junction diode as a variable capacitance, because it was feared that a transistor oscillator at this, to us at the time, unprecedentedly high frequency might not be adequately stable. It was found in practice, however, firstly that there was no difficulty at all in making the circuit oscillate, and secondly that, for all practical purposes, there was no drift; within a few seconds of switching on, the frequency was within a few kc/s of the right value. The A.F.C. circuit was therefore dispensed with as such, manual adjustment of the bias on the diode providing instead a convenient fine-tuning control of non-microphonic nature.

(c) The fact that a transistor oscillator can readily be battery operated, with long battery life, is very convenient in signal generators, bridge sources etc., since the oscillator and its battery can be placed inside one small screening box, completely obviating the problem of filtering supply leads met with valve oscillators.

Since it is practicable to use transistors as switches even at very low power levels, it has been found that the instances where a class 'A' oscillator is the best choice are much less frequent than with valves. A good class 'A' oscillator requires the additional complication of some form of amplitude control device, which may be either a thermistor or an A.G.C. circuit operated by rectifying the oscillator output. A switching circuit such as that described in Section 9 is simpler and cheaper, and, with a high-Q coil, such as may readily be made with a Ferroxcube pot core, can give under 0.1% harmonic distortion. (Examples of class 'A' oscillators are given, however, in Sections 11 and 12 of these notes).

Conventional class 'C' oscillators are discussed in Section 4, where it is concluded that, despite the excellent switching properties of transistors, it is not possible to achieve, in reliable practical circuits, quite the values of efficiency that might at first sight be expected. The main factors that mitigate against obtaining very high efficiencies are:-

(a) The necessity in most cases of providing some degree of stabilization against effects caused by variation of I_{co} with temperature. The usual method is to obtain most of the bias by a resistor in the emitter circuit, as in Fig. 4(b), instead of by a resistor in the base circuit as in Fig. 4(a), but much more power is lost using emitter bias.

(b) The necessity of avoiding squegging prevents one from using as large a value of bias capacitor as would otherwise be desirable, which in turn prevents one obtaining the small angles of collector-current flow that are necessary for the highest efficiency.

The fact that, were it not for these difficulties, overall efficiencies of about 90% could be obtained, whereas with valves such efficiencies are out of the question, due mainly to heater power, seemed to the lecturer to offer a great incentive to try to find ways of overcoming the weaknesses of conventional class 'C' oscillators. This investigation has led to some useful push-pull circuits which give much better performance, and are easier to design, than ordinary class 'C' oscillators. That described in Section 7 is thought to be the most suitable for most purposes.

However, in many oscillator applications, the performance requirements are not critical, and very simple single-transistor circuits may then be used, such as the class 'C' oscillators of Fig. 4. An alternative, and sometimes even cheaper, solution, less widely adopted but sometimes very suitable, is a class 'B' oscillator as described in Section 5.

In Section 3 is discussed an important general principle relating to simple transistor L - C oscillator circuits, which does not seem to be as widely appreciated as it should be. It is quite usual to refer to these simple circuits as common-base oscillators, common-collector oscillators or common-emitter oscillators and to analyse them in terms of the corresponding transistor equivalent circuit. However, since all these circuits are shown to be essentially equivalent to a combination of a tuned auto-transformer and a transistor, and since input and output circuits are effectively combined via the transformer, it is purely a matter of convenience which terminal is called the common one in analysing the circuit.

Realization of these equivalences would seem to simplify considerably the problem of choosing and designing a suitable circuit, the choice being largely a matter of which point it is convenient to have earthed in the practical application.

3. THE EQUIVALENCE OF VARIOUS WELL-KNOWN OSCILLATOR CIRCUITS

Fig. 1 shows several simple transistor oscillator circuits, in which the means adopted for securing stable d.c. operating conditions and control of the amplitude of oscillation have been omitted, since they are not relevant to the present discussion.

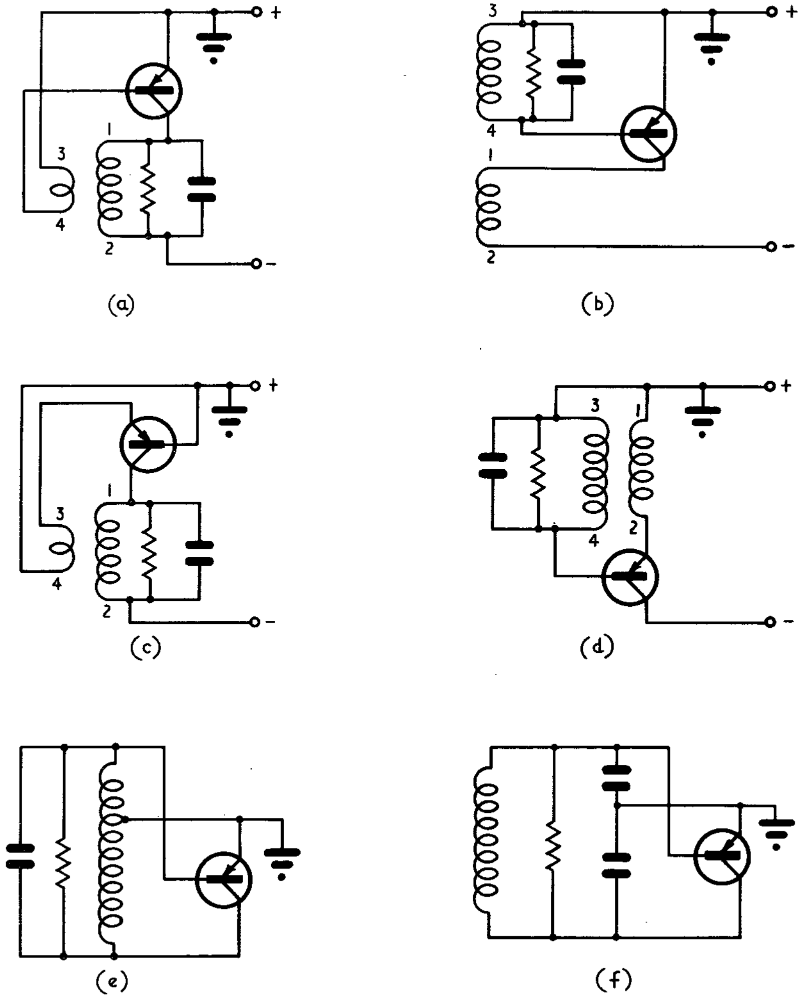


Figure 1 Simple Oscillator Circuits

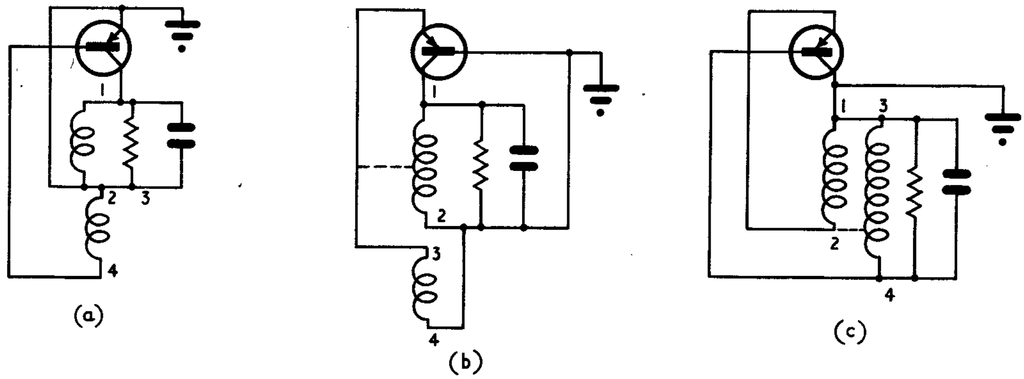


Figure 2 Redrawn Oscillator Circuits

That these circuits are really all equivalent to one another, assuming the coils to be tightly coupled magnetically, may be seen by appropriately redrawing some of them. Since the d.c. supply leads are at the same a.c. potential, they may be joined together in a circuit diagram concerned only with a.c. conditions. Thus Fig. 1(a) may be drawn as in Fig. 2(a) when the two windings are seen to constitute effectively an auto-transformer with the tuning capacitor and load across the top section. A redrawn version of Fig. 1(b) is the same except that the load and tuning capacitor come across the lower part of the auto-transformer. On redrawing Fig. 1(c), Fig. 2(b) is obtained, and here the arrangement of windings shown (same direction of turns from terminal "1" through to terminal "4") is equivalent to connecting the emitter to a tapping between "1" and "2" as shown dotted, and with this modification the circuit is seen to be equivalent to Fig. 2(a) but with the tuning capacitor and load across the total auto-transformer winding; equivalent capacitor and load values may, of course, be determined for connection across only one section of the auto-transformer. It may further be observed that Fig. 2(b), with dotted modification, is the same as the Hartley circuit of Fig. 1(e), except for the position of the earth connection, the latter being of no consequence for analytical purposes.

Redrawing Fig. 1(d) gives Fig. 2(c), and the coil between "1" and "2" in the latter may be removed and replaced by an equivalent tapping on the coil between "3" and "4" as shown dotted; the circuit is then seen to be equivalent to the others except for the position of the earth connection.

With regard to Fig. 1(f), the question is really to what extent a capacitive tapping on a tuned circuit is equivalent to a tapping on the coil, and the exact equivalence of the two circuits shown in Figure 3 is considered to provide the simplest basis for the comparison.

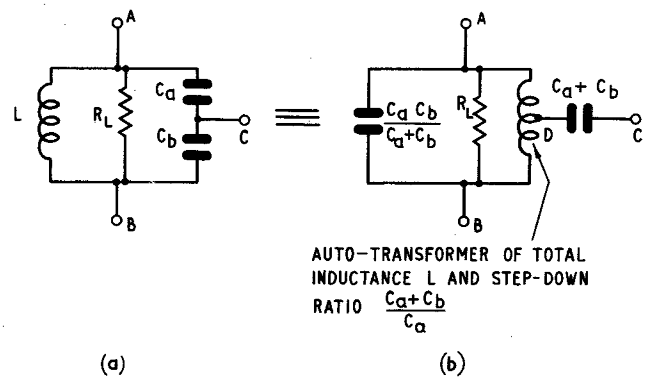


Figure 3 Equivalent Circuits for "Capacitance Tap"

(The method may be extended to circuits of the Gouriet or Clapp type[(2),(3)], having three series capacitors across the coil). In most practical circumstances the reactance of $C_a + C_b$ in Fig. 3(b) is small enough for the voltage across it to be negligible, and there will then be no significant difference in behaviour between circuits such as Fig. 1(e) and (f), provided that the tapping ratio and total inductance are made the same in each.

One of the things that becomes immediately evident from the above is that a circuit such as Fig. 1(d), which might be described as an emitter-follower oscillator, has no theoretical advantages due to (apparently) employing the transistor as an emitter-follower. That confusion of thought has been caused by matters such as this may be seen from a letter to the Editor of "Wireless Engineer"[(4)] from four of us at R.R.E., where it was shown that certain single-valve cathode-follower R - C oscillator circuits, which had been claimed to have marked advantages due to the low-distortion properties of a cathode follower, were in fact exactly equivalent to well-known circuits of conventional type, and could not, therefore, possess any particular virtues.

4. CLASS C OSCILLATORS

4.1 Base-current-biassed circuits

Since a transistor, in striking contrast to a valve, passes substantially no current at zero bias voltage, means must be provided to bias it on initially if self-starting of oscillation is to be obtained when the d.c. supply is switched on. For efficient class C operation, however, a large bias in the opposite direction to the above is necessary, and this bias may be produced by rectification, in a manner comparable with the usual arrangement in a valve oscillator, by including a resistor and capacitor in the base circuit. The simplest arrangement that satisfies both these requirements is shown in Fig. 4(a), and is capable of overall power-conversion efficiencies well above 70%.

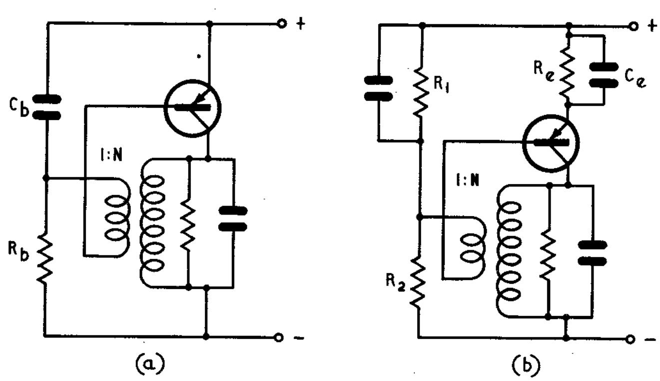
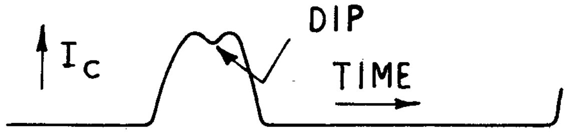


Figure 4 Class 'C' Oscillators

The turns ratio, N, is typically 2 or perhaps a little more. R_b should normally be adjusted to give appreciable bottoming, since this very effectively stabilizes the output voltage amplitude; in many cases it is reasonable to reduce R_b to about half the value at which bottoming begins. The onset of bottoming is quite clearly defined, as R_b is reduced, and is indicated by the appearance of a dip in the collector-current pulse:-



If R_b is reduced to a value <u>much</u> lower than that at which the dip appears, the dip becomes enormous and the direction of current flow may reverse during the central part of the pulse. This condition should be avoided, since it gives a large increase in distortion together with reduced efficiency and increased collector dissipation.

The circuit of Fig. 4(a), in common with others in which the bias is produced by the flow of base current, has the undesirable feature that the degree of bottoming is greatly affected by the value of α' of the particular transistor used, and in practice it is usually necessary to adjust R_b to suit the transistor sample.

A further undesirable feature of the Fig. 4(a) circuit, especially in low-power versions, is that the degree of bottoming is considerably affected by variations in I_{co}. An increase in I_{co} is equivalent to raising the current in R_b, and results in an increase in the amplitude of the base-current pulses and hence in the degree of bottoming. Thus the circuit is not well suited to situations where wide variations in ambient temperature are expected, and variations in I_{co} and α' during the life of the transistor should also be reckoned with. Nevertheless, the circuit has the virtues of simplicity and cheapness combined with high efficiency, and provided it is run with a fairly high degree of bottoming, the disadvantages just mentioned are not always serious.

4.2 Emitter-current-biassed circuits

A circuit that is usually much preferable to the above is shown in Fig. 4(b). Depending on the component values used, this circuit may or may not operate under class C conditions, but such conditions are assumed here. The bias required for class C operation is produced by the emitter current flowing in R_e, R_1 and R_2 preferably providing only the small amount of negative bias that is necessary to ensure reliable self starting - typically 0.3 volt with a normal germanium transistor. The resistance of the base bias network should preferably be low enough to ensure that only insignificant voltage changes occur across this network when base current flows during oscillation.

The effects of changes in α' and I_{co} are very much less in this circuit than in the previous one, since the base current now adjusts itself automatically until the emitter current is sufficient to give a voltage across R_e nearly equal to the peak negative value of the voltage applied to the base.

The main disadvantage of the circuit, compared with Fig. 4(a), is that a substantial fraction of the h.t. power is dissipated in R_e, giving reduced overall efficiency. If the collector dissipation is to be kept to a minimum for a given power output, the turns ratio N should be made small, so that a small angle of flow is obtained, but this reduction of transistor dissipation is achieved at the expense of an increased power loss in R_e. There is thus a value of N that gives maximum overall efficiency in any given situation. With N = 2, an overall efficiency of 60% can be achieved in practice.

4.3 Combined emitter-current and base-current biassed circuits

A compromise between the two circuits of Fig. 4 is possible by making R_1 and R_2 large enough to develop, by the flow of base current, a considerable fraction of the total bias voltage. A higher overall efficiency than can be obtained with only emitter bias is then possible, but with increased dependence of performance on α' and I_{co}.

4.4 The avoidance of squegging

A further factor which complicates the design of the above circuits is the necessity of avoiding squegging. The squegging mechanism is discussed in some detail in Reference (5), and its application to transistor oscillators is considered in the author's I.E.E. paper referred to in the Introduction to this article. Here the conclusions only will be given :-

(a) If pronounced bottoming is present, it exerts a very potent stabilizing influence on the squegging mechanism, and it is then possible to employ a much larger value of bias capacitor than when there is little or no bottoming. It is not normally to be recommended, however, that this condition of operation be adopted; a curious mode can occur in which only alternate collector-current pulses have a bottoming dip, and the output voltage then contains a small amount of sub-harmonic.

(b) If bottoming is not present, so that the above stabilizing influence is inoperative, then an approximate criterion, applicable to the Fig. 4(b) circuit, is that squegging will not occur if $R_e C_e$ is less than T_d, where T_d is the decrement time constant of the tuned circuit, given by:-

$$T_d = \frac{Q}{\pi f_o} \qquad \ldots(1)$$

in which Q refers to the loaded Q of the tuned circuit and f_o is the frequency of oscillation.

No such circuit on which the lecturer has experimented has failed to squegg if R_eC_e has been made equal to $2T_d$, nor has it failed to give continuous oscillation if R_eC_e has been made equal to $\frac{1}{2}T_d$.

(c) The same criterion as above may be applied to class C oscillators in which the bias is produced by a base resistor and capacitor, provided that the resistor is returned to approximately emitter potential. In the circuit of Fig. 4(a), however, the bias resistor R_b is returned not to emitter potential, but to the negative side of the supply, and the above criterion is then not directly applicable. The fundamental factors involved in determining whether or not squegging will occur are, however, not the time constants as such, but the initial rates at which the bias voltage and the amplitude of the oscillatory voltage fed back to the base, respectively, decay if transistor current ceases. It may then be shown that the approximate criterion for the Fig. 4(a) circuit is that squegging will not occur provided:-

$$R_b C_b < (1 + N) T_d \qquad \ldots(2)$$

Assuming that the collector voltage swings up almost to emitter potential, the value of R_b as shown in Fig. 4(a) is, it is easily shown, approximately 1 + N times the value of resistance that would give the same base current if connected straight across C_b. Thus the Fig. 4(a) circuit needs 1 + N times as much resistance, and equation (2) above shows that it can stand just 1 + N times more with the same margin against squegging if C_b is the same in both circuits. Thus, with a given value of C_b, the Fig. 4(a) circuit is no more prone to squegging than a circuit operating under the same conditions but with a smaller value of R_b returned to approximately emitter potential, and this has been verified to be the case experimentally. The Fig. 4(a) circuit has the advantage of using the minimum possible number of components for a class C oscillator, and its only disadvantage compared with the arrangement using a smaller value of R_b is that the power dissipated in R_b is greater, but since this power is often less than 5% of the output power, it is not usually a very important consideration.

4.5 Practical design procedure

To conclude this discussion of class 'C' oscillators, a few comments on practical design procedure may be useful. It is, of course, possible to go into the design of class 'C' oscillators in great analytical detail, but even when this is done, fairly drastic simplifying assumptions are nearly always made, so that a small amount of experimental adjustment of values is often necessary in order to obtain the wanted performance after building the circuit.

The author's present philosophy is to use one of the newly-developed circuits, such as those described in Sections 7 and 9, when a fairly stringent performance specification must be reliably met, and to use a simple class 'C' or class 'B' oscillator when it is desired to use only one transistor and to make the coil windings as simple as possible, or, alternatively, when the frequency is too high for satisfactory operation of circuits described in Sections 7 and 9.

A very satisfactory design procedure for most class 'C' oscillator problems has been found to be:-

(a) Choose the circuit of Fig. 4(a) if the cheapest arrangement is thought adequate.

(b) Choose circuit of Fig. 4(b) if a performance less dependent on α' and temperature is needed, making R_1 and R_2 low enough for the base current (which is approx $I_{ht} \div \alpha'_{dc}$) to cause only a small fraction of a volt change in the voltage across R_1.

(c) Make the turns ratio N equal to 2, as a measure of standardised design; the advantages to be gained from departing from this value in particular cases are usually so small that it hardly seems worth the trouble involved.

(d) Decide what peak voltage will exist across the collector winding. This may sometimes be determined by the peak collector voltage rating of the transistor, or, alternatively, the available supply line voltage V_{dc} may be the determining factor. In the circuit of Fig. 4(a), the peak collector voltage will be nearly V_{dc}, i.e. this is the peak a.c. voltage across the collector winding; when the collector swings negative, the peak voltage between collector and emitter is obtained by adding this peak a.c. voltage to the supply voltage itself, so that the transistor must withstand a maximum voltage of nearly 2 V_{dc}. In the circuit of Fig. 4(b), with N = 2, the peak collector-winding voltage will be about two thirds of V_{dc} with about one third of V_{dc} appearing across R_e; the peak voltage across the transistor when it is off is then about 4/3 times the supply voltage.

(e) Knowing the output power required, the shunt load resistance referred to the collector winding may then be determined, since the voltage across this winding is known from (d).

(f) The next quantities to be determined are the shunt L and C values referred to the collector winding, and before these can be evaluated it is necessary to decide what loaded-Q value to employ. For high efficiency, most of the transistor output power should go to the load and only a little to the tuned-circuit losses. The loaded Q must therefore be made much less than the unloaded Q of the tuned circuit itself; but the lower the loaded Q is made the greater the harmonic distortion becomes, so that a suitable compromise is called for. When efficiency is important, a loaded Q of 10 is often used, and if the unloaded Q is 100 or more, which can be readily achieved over a wide frequency range with modern Ferroxcube pot cores(6) or dust pot cores(7), then less than 10% of the output power will be dissipated in the tuned circuit losses. The second-harmonic distortion will be in the region of

5%. Having decided on an appropriate loaded Q, the L and C values referred to the collector winding are then calculated; e.g. if the loaded Q is to be 10, then it is near enough to say that X_L and X_C should be one tenth of the load resistance referred to the collector winding, at the operating frequency. The necessary number of turns for the collector winding is then determined from the core data, and the feedback winding requires half this number when a pot core is used; with a slug core or air core, the coupling may be much less than 100% and more feedback turns may be needed to compensate for this.

(g) The output-winding turns (if a separate output winding is used) are then calculated so as to make the actual load shunt resistance look like the required value when referred to the collector winding.

(h) A separate tuned winding may have to be used if the capacitor value required to tune the collector winding is inconveniently large; this will almost certainly be necessary if variable tuning is required. To avoid the appearance of snigs on the output waveform due to the pulse nature of the collector current, it is safest to tune the collector winding or the output winding whenever practicable. This trouble may be minimized, when a separate tuned winding is employed, by choosing a pot core that will give the required Q with the minimum turns, and by winding the windings one inside the other, rather than side-by-side. It is preferable to put the collector winding on one side of the tuned winding and the output winding on the other; the problem is really the same as that which is involved in the design of an output transformer with a tertiary winding for negative feedback.

(j) All that remains is to determine the bias R and C values. With the circuit of Fig. 4(a), assuming it is only just on the verge of bottoming, the current in R_b will be approximately equal to the supply current divided by α'_{dc}, and the supply current can be estimated from a knowledge of the output power and an assumed efficiency figure, say 75%. The voltage across R_b is approximately $1.5\ V_{ht}$ with $N = 2$, so that a value may be deduced for R_b at which it would be predicted that the circuit would be on the verge of bottoming. From equations (1) and (2) the value of C_b at which squegging would be on the point of occurring may be estimated. By reducing R_b to half the above value, a satisfactory degree of bottoming should be obtained, i.e. enough to stabilize the output voltage reliably but not enough to have much adverse effect on efficiency or distortion. With R_b thus reduced, squegging should be safely avoided, but an extra margin may be introduced if desired by also halving the value of C_b; the circuit should then be non-squegging even if the load resistance is made considerably lower than the intended value.

With the Fig. 4(b) circuit, again assuming initially that it is only just on the verge of bottoming, R_e must be such as to develop about one third of the supply voltage when the working current flows through it. As before, the supply current may be estimated from the power output and an assumed overall efficiency, about 55% being a likely value for the latter in this circuit. Having determined R_e, the critical value of C_e, from a squegging point of view, may be obtained from the relationship $R_e C_e = T_d$ in which T_d is as given by equation (1). The actual value of R_e for the first trial of the oscillator should be rather less than the above, say 70% of it, to give reliable bottoming and hence stabilization of output voltage. A somewhat lower value of C_e than that calculated above, say half the calculated value, is recommended; too excessive a reduction will, however, give an unnecessarily large increase in the angle of current flow and the efficiency will suffer. (The increase in the angle of flow is because, with small C_e values, an appreciable a.c. voltage exists across C_e and causes the base-to-emitter voltage to reach the value at which collector current starts, sooner than would be the case if the emitter voltage were d.c. only.)

(k) When the oscillator has been built, it is desirable to check, with the correct load applied, that the circuit is operating with a suitable degree of bottoming. This may be done by inserting a suitably-small resistance value in series with the collector circuit and thus observing the current waveform. (When working above 100 kc/s or so, care should be taken to connect the oscilloscope input cable straight across the ends of the series resistor, so as to avoid introducing significant inductance - a few inches of wiring

can have a reactance comparable with the resistance value used, and can result in quite misleading waveforms. This problem tends to be more acute when investigating transistor circuits than when investigating valve circuits, because the current-monitoring resistor value usually has to be smaller in transistor circuits to avoid significantly affecting the operation of the circuit. Another important point is that the oscilloscope cable, say UR70, should be terminated in a 70 ohm resistor at the oscilloscope end, and preferably fed via a 70 ohm resistor at the input end also, if there is sufficient gain to permit this - without the 70 ohm resistors serious ringing may appear on the waveform due to cable reflections and can be very misleading.)
As an alternative to the above method, the effect on the sine-wave voltage output amplitude of varying the bias resistor may be investigated. When the bias resistor value is raised above the value at which bottoming ceases, any further increase in its value will cause the output voltage to fall off relatively rapidly. About half this critical value of bias resistor should be adopted for the circuit of Fig. 4(a), and rather more than half for that of Fig. 4(b). If squegging ensues as soon as bottoming ceases, a lower value of bias capacitor is necessary.

(l) If the output voltage is not quite what is required, a slight adjustment of the number of turns on the output winding may be necessary.

(m) It is normally reasonable to make the current in R_2 in Fig. 4(b) about 5% of the main supply current, with about 0.3 volt drop across R_1 at zero base current. The capacitor across R_1 should have a reactance not greater than about a tenth of R_1.

The above design procedure will be found quite straightforward when the operating frequency is a great deal less than the alpha-cut-off frequency of the transistor. With a loaded Q of 10, an angle of flow of about 75° will be obtained. Even at frequencies as low as 0.1 f_α, however, hole storage will have a considerable influence on the circuit behaviour, and will result in an increase in the angle of flow with a consequent reduction in efficiency. As the frequency is increased, the base current pulses become more and more markedly bidirectional, and even though the mean base current remains much as at low frequencies, the peak values become much larger. Thus, in the Fig. 4(a) circuit, the a.c. voltage across C_b becomes much larger than in a low-frequency oscillator, and the effect of this is to increase the angle of flow and reduce the efficiency. This effect is less evident in the Fig. 4(b) circuit, since here it is the emitter current pulses that charge the bias capacitor, and these remain substantially unidirectional up to quite high frequencies.

The criterion already given for squegging is still usable under the higher-frequency conditions referred to above, with sufficient accuracy to be a useful guide, but one must be prepared to do rather more experimenting with circuits which operate at a moderately large percentage of f_α.

4.6 Low-Level Class 'C' Operation without bottoming

In the above discussion of class 'C' oscillators it is assumed that an appreciable power output is required and that efficiency and output voltage regulation are matters of importance.

In some applications, however, the power output required is quite small and efficiency as such is of no great concern. An example of such circumstances occurs in the local oscillator of a radio receiver, and a circuit much like that of Fig. 4(b) is often used, but with rather different operating conditions from those discussed previously.

Suppose that R_1 and R_2 are chosen to supply a volt or more of negative bias at low impedance, and that the peak a.c. voltage fed back into the base circuit is several times less than this fixed bias. Then, to a reasonable approximation, the mean emitter current is determined only by the d.c. base bias and the value of the emitter resistor. Provided, however, that the peak a.c. voltage fed back to the base circuit is several times kT/q ($\doteq$ 25 mV), and provided C_e is large enough, the transistor current will flow in fairly short pulses, whose mean value is determined in the manner described above; consequently the fundamental component of these pulses is also fairly closely determined.

By suitable adjustment of the dynamic resistance of the tuned circuit, the required output voltage may be obtained. This output voltage, even though bottoming is not allowed to occur, can be quite satisfactorily independent of changes in transistor parameters, but not, of course, independent of changes in loading. For operation in this manner, the turns ratio N is typically 10.

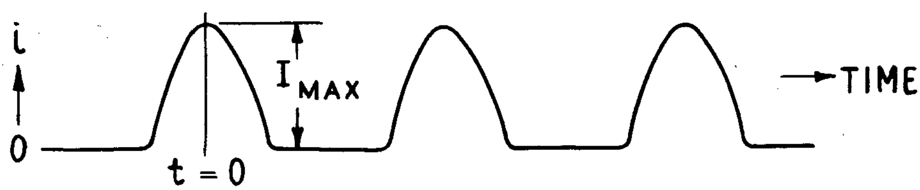
It has been observed that the random amplitude-modulation noise level on the output of a simple class 'C' oscillator is much less when bottoming occurs than when there is no bottoming. In a particular 1 Mc/s oscillator made by the author, the noise dropped very rapidly by about 20 dB as soon as a trace of bottoming was allowed to occur.

5. CLASS B OSCILLATORS

As with class 'C' oscillators, only single-ended circuits will here be considered, since in cases where push-pull is regarded as worth while, the new circuits described in Part 2 of these notes are thought to be a better choice.

The cheapest class 'B' oscillator circuit is as in Fig. 4(a) but with C_b replaced by a resistor. This resistor should have a value suitable to limit the peak base current to just over the value required for bottoming, and the lower resistor should be chosen so as to give about 0.3 volt d.c. across the top resistor in the absence of base current, thus ensuring self starting.

Assuming the peak a.c. collector voltage to be equal to the supply voltage, the value of the load resistance referred to the collector winding may be calculated to give the power output required. The peak value of the fundamental component of the collector current may also be determined from these quantities. What the peak instantaneous value of the actual collector current waveform is must then be determined. The waveform is:-



and the Fourier series for it is:-

$$i = \frac{I_{max}}{\pi}\left\{ 1 + \frac{\pi}{2}\cos\omega t + \frac{2}{3}\cos 2\omega t - \frac{2}{15}\cos 4\omega t + \frac{2}{35}\cos 6\omega t \; \ldots\ldots\ldots\ldots \right. \qquad \ldots(3)$$

Thus the fundamental component has a peak value of $\frac{1}{2} I_{max}$, so that in the oscillator the peak collector current is of twice the peak value of the sine-wave current flowing in the collector load resistance.

Provided the frequency is very low, i.e. below $f_{\alpha'}$, the peak base current is simply $I_{max} \div \alpha'$. A standardized value of 2 may very reasonably be adopted for N, as with class 'C' oscillators, and hence the approximate base circuit resistance necessary may be determined. This circuit suffers from dependence on I_{co} and α' as in the class 'C' counterpart, and its behaviour is rather complex at higher frequencies. For these reasons a circuit that is much preferable in most circumstances is as in Fig. 4(b) but without C_e. If R_1 and R_2 (whose sole purpose is to provide about 0.3 volt of bias to ensure self starting) are reasonably low, the capacitor across R_1 may also be omitted, resulting in a very cheap circuit that is, nevertheless, very easy to design and very satisfactory in performance provided an overall efficiency of about 45% is acceptable - this latter figure assumes N = 2. In this circuit the peak collector current is given approximately by the peak fed-back voltage divided by R_e, and is largely independent of transistor parameters. The peak a.c. collector voltage is approximately two thirds of the supply voltage if N = 2.

It is an advantage of these class 'B' circuits that there is no possibility of squegging.

With a loaded Q of 10, the theoretical value for second-harmonic distortion is approximately 2.8%; there is ideally no third harmonic, nor 5th, 7th etc., as may be seen from equation (3).

6. IMPROVED HIGH-EFFICIENCY OSCILLATORS - GENERAL APPROACH

For high-efficiency operation in any transistor oscillator, it is axiomatic that when current flows through a transistor the voltage across it should be as small as possible, so as to minimise the power dissipation in the transistor.

In the class 'C' oscillator, since the voltage across the transistor varies approximately sinusoidally, it is necessary to switch the transistor on during only the very short time intervals for which the transistor voltage is nearly zero*, and this inevitably leads to difficult design compromises as described in Section 4.

In general, what is required for improved performance is to add some suitable non-power-dissipating (i.e. reactive) elements between the transistor(s) and the tuned circuit, so as to allow the transistor voltage to remain small during the whole of a long conduction time and yet to permit the tuned circuit meanwhile to execute a sinusoidal oscillation.

One idea[(8)], which occurred independently to the author after reading reference (9), involves the addition of one or more low-loss parallel tuned circuits, tuned to harmonic frequencies, between the transistors and the main tuned circuit. This scheme did not seem to offer as complete and simple a solution as the schemes described in Sections 7 and 8 below, and, in view of the practical success of the latter when using transistors, it was abandoned.

In both the new push-pull oscillators described below, the transistors are used as switches, and each transistor is on for half an oscillation period, i.e. the angle of flow of collector current in each transistor is 180o. In this sense the circuits may be described as class 'B' oscillators, but they differ from conventional class 'B' systems in that the voltage across each transistor is substantially zero throughout the whole of the conduction time of that transistor, so that much higher

---

* If the loaded Q of the tuned circuit is too low, the transistor voltage will inevitably not be small for the first part of the conduction time, which necessarily results in reduced efficiency. The new circuits described do not require high values of loaded Q (i.e. low decrement) for high efficiency.

efficiencies can be obtained than those normally associated with class 'B' operation. These high efficiencies can be obtained, moreover, with a performance which is independent of variations in temperature and transistor current gain to an extent which is not achievable with conventional class 'C' oscillators, and the circuits are easier to design.

The same principles may be applied, of course, to high-efficiency power amplifiers, and it is suggested that the term "class 'D'" might be appropriate to describe the mode of operation.

A disadvantage of these class 'D' circuits is that the maximum frequency for satisfactory operation is lower, for a given type of transistor, than when class 'C' is used. It is hoped in due course largely to overcome this disadvantage.

7. HIGH-EFFICIENCY CURRENT-SWITCHING OSCILLATOR (Pat.Appn.No. 8865/59

7.1 Principle of operation

Fig. 5 shows the high-efficiency current-switching oscillator in its simplest form.

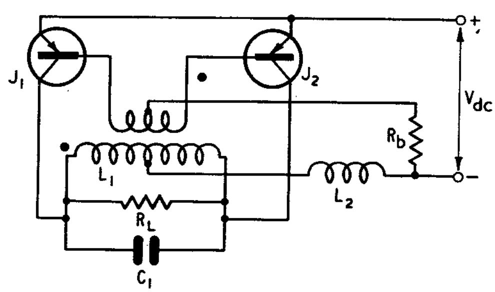


Figure 5 Basic Class 'D' Current-Switching Oscillator Circuit

When the left-hand transistor J_1 is on, the left-hand end of the tuned circuit is clamped almost to the positive supply potential and the right-hand end of the tuned circuit executes a negative-going half sine wave of oscillation. Similarly, when J_2 is on, the left-hand end of the tuned circuit executes a negative-going half sine wave of oscillation. The voltage waveform at the centre-tap of the tuned circuit is therefore a succession of half sine waves as shown in Fig. 6.

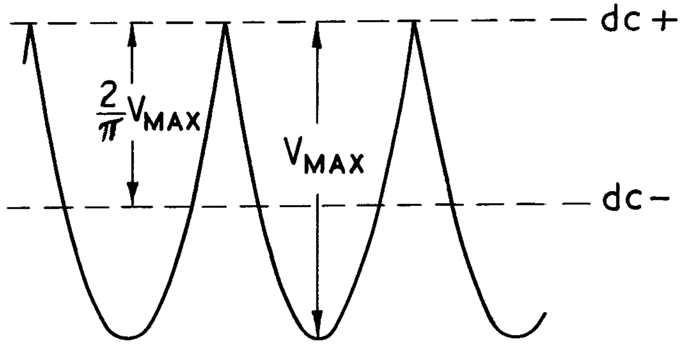


Figure 6 Waveform on Collector Winding Centre Tap

This waveform has a mean value of $2/\pi$ of its peak value, and the potential of the left-hand terminal of L_2, with respect to the positive supply line, also has this same mean value, being directly connected to the tuning coil centre tap. The right-hand terminal of L_2, however, has its d.c. potential fixed by being connected to the negative supply line. What happens is that the amplitude of oscillation automatically adjusts itself until the mean potential at the negative tuned circuit centre tap is very nearly equal to that at the supply terminal, the small difference, divided by the d.c. resistance of L_2, determining the feed current. L_2 would, ideally, be of almost infinite inductance, so that a constant current would be maintained and switched alternately through J_1 and J_2. In practice a suitable choice is to make L_2 about five times L_1, which, with a loaded Q of 10, will limit the peak-to-peak current fluctuation in L_2 to about a quarter of the mean supply current.

The base current is, to a first order, V_{dc}/R_b, and R_b is simply chosen to make this current comfortably sufficient to bottom the transistors at the collector current demanded.

At the moment when conduction is being switched over from one side to the other, the voltage across the tuned circuit is swinging through zero, so that very little voltage can exist across either transistor during this switchover operation, but the full supply current is flowing. The oscillator has therefore been termed a high-efficiency current-switching oscillator.

A turns ratio of about 10 : 1 between the whole collector winding and the whole base winding is normally suitable. If this ratio is made too small, the power dissipated in R_b is considerably increased.

Some convenient design formulae, taken from the author's I.E.E. paper referred to in the Introduction to this article are:-

$$(\text{r.m.s. voltage between collectors}) = \frac{\pi}{\sqrt{2}} V_{dc} \qquad (4)$$

(The actual voltage obtained will be a few percent less than this, since the collector-to-emitter voltage drop when bottomed, and also the d.c. voltage drop in L_2, have been neglected.)

$$(V_{ce})_{max} = \pi V_{dc} \qquad (5)$$

where $(V_{ce})_{max}$ is the peak collector-to-emitter voltage applied to each transistor when it is in the non-conducting state. The d.c. supply voltage must be made low enough to avoid exceeding the collector voltage rating of the transistors.

$$(\text{Third-harmonic distortion}) = \frac{100}{8\,Q}\ \% \qquad (6)$$

7.2 Effect of hole storage

An important point in the operation of this circuit is that whereas the mean base current is fairly closely determined by R_b and V_{dc}, it is possible, due to hole-storage effects, for much larger instantaneous base currents to flow during the short time that current is being switched over from one transistor to the other. Thus, when J_1 is being switched off, a large reverse base current of transient nature may flow into its base from the base of J_2, thus rapidly clearing the minority carrier hole concentration in J_1 and injecting holes into transistor J_2, which therefore bottoms quickly. This ability to supply large transient currents in the emitter-base circuit ensures that current is switched from one collector to the other cleanly and rapidly even at quite high working frequencies.

7.3 Modification to reduce power loss in R_b

Ideally (assuming all reactances to be lossless and the transistors to be ideal switches), the only place where unwanted power dissipation occurs is in R_b, but the magnitude of this power loss may be less than 5% of the output power. The power dissipated may, however, be considerably reduced, if desired, by inserting a choke in the feedback winding centre-tap lead and taking the other end through a diode to the positive terminal of the supply. The diode conducts when the circuit is in its normal operating state and passes the base current. The base current is kept nearly constant by the choke and its value may be adjusted suitably by putting a small amount of resistance in series with the diode or by adjusting the feedback winding turns. Quite a high value of resistance between the feedback winding centre tap and the negative terminal of the supply will then provide enough base current to secure self starting, with only a very small loss of power in this resistor; the diode is, of course, non-conductive during this starting up process.

7.4 Practical Designs

A practical oscillator based on the above principles, designed for energising a measuring bridge at about 100 kc/s, is shown in Fig. 7.

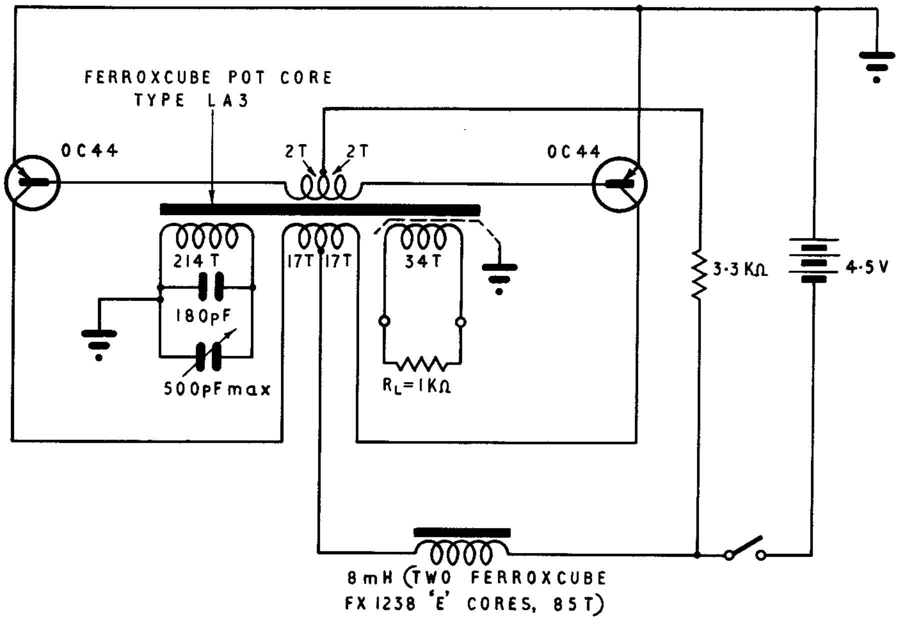


Figure 7 Practical Circuit

The following results were obtained with this oscillator:-

Power input: 120 mW Power output: 95 mW

Efficiency: 79%

The total power loss is thus 25 mW, and an approximate estimate of the distribution of this loss is:-

L_1: 11 mW, L_2: 1 mW, R_b: 8 mW, Transistors: 5 mW.

In this particular model the transistors are called upon to develop much less output power than that which they are capable of providing in this type of circuit. The oscillator described in Section 8.4 is a better illustration of the potentialities, with regard to output power, of circuits in which bottoming occurs throughout the whole of a 180° angle of flow.

Dr. A.C. Moore has employed this circuit successfully at 250 c/s to drive an air-cored a.c. electromagnet, consisting of two 4 ins. dia. 200-turn Helmholtz coils, these coils forming the main tuning inductance of the oscillator. The Q of this tuned circuit is about 6, and the tuning capacitance 25μF. Using two OC72's with a heat sink, a peak current of about 1 amp is obtained in the magnet. A point which has come to light as a result of this work is that, on switching off the supply, the choke L_2 develops a large voltage transient tending to make the collectors highly negative and the bases highly positive. In a low-frequency oscillator there may be sufficient energy available from L_2 to damage the transistors. The cure is either to reduce the supply voltage slowly by a series pot, or to open a switch in series with the choke first and then interrupt the supply to R_b afterwards.

A 400 c/s amplifier working on these principles, and driven by an independent multivibrator, has also been made. This is now in use at Harwell. An output of 5 watts is obtained at an overall efficiency of 85%. The loaded Q is only about 3, giving approximately 4% third harmonic distortion. This low value of loaded Q, which would be quite impracticable in a class 'C' amplifier, makes it possible to keep the tuned circuit power losses low without requiring an excessively high value of unloaded Q. The unloaded Q is about 20 in this design, and an inductor weighing several pounds is required even for this value.

An interesting feature is the extreme linearity of the relationship between output voltage and d.c. supply voltage to the power amplifier. This linearity continues down to supply voltages of well under 100 mV, making linear modulation to a depth of over 99% practicable.

Experience with this design has made it very evident that at much higher power levels, where the large inductors required would have much higher values of unloaded Q, efficiencies well above 90% would be obtainable.

More recently a 3.5 watt tape erase/bias oscillator operating at 50 kc/s has been built, using two OC24's in the circuit of Fig. 5. This gives less than 0.1% second-harmonic distortion and has an overall efficiency of nearly 80%.

7.5 Squegging

Squegging can occur in this type of oscillator, giving an output waveform consisting of bursts of oscillation and quiescent periods alternately. A quantitative analysis has not yet been done, but the mechanism appears to be as follows:-

(a) During the quiescent part of the squegg waveform both transistors are bottomed and current is building up linearly in L_2, ultimately becoming so large that the available base current is insufficient to hold the transistors bottomed any longer.

(b) Oscillation then starts to build up very rapidly, due to the large value of the choke current, which is now switched alternately to the two ends of the tuned circuit. If L_2 is large enough, so that there is sufficient stored energy available from it, the oscillation amplitude may reach a value considerably in excess of that given by formula (4), resulting in a substantial reverse voltage across L_2.

(c) With a reverse voltage across L_2, the mean current in L_2 falls, ultimately reaching zero and then flowing from right to left. Once reverse current is flowing in L_2, energy is being taken out of the tuned circuit instead of being put in, causing the oscillation to be damped out and a state finally reached in which there is no oscillation, the transistors are bottomed and current in L_2 is building up once again in the normal direction. The sequence of events then repeats itself.

It is found in practice, as would be expected from the above, that squegging is most likely to occur when L_2 is made unnecessarily large and/or R_b is made unnecessarily small. With L_2 chosen according to the recommendation in the paragraph following Fig. 6, there seems to be a large margin of safety against squegging, and continuous oscillation may be obtained even when the load is removed.

8. HIGH-EFFICIENCY VOLTAGE SWITCHING OSCILLATOR (Pat. Appn. No. 8865/59

8.1 Principle of operation

The basic circuit of the voltage-switching oscillator is shown in Fig. 8; the bias arrangements necessary to ensure self-starting are, however, omitted for clarity and will be described later.

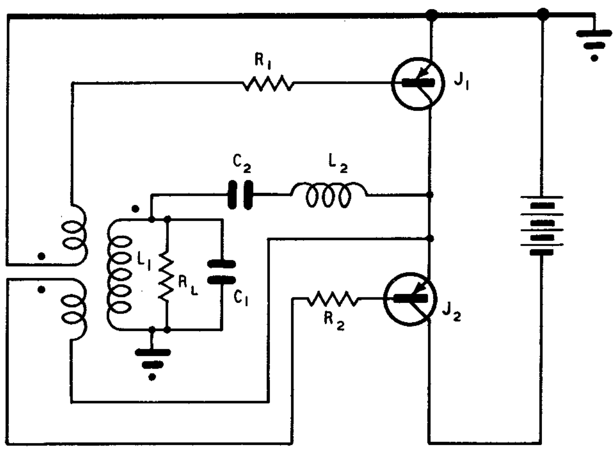


Figure 8 Basic Class 'D' Voltage - Switching Oscillator

The transistors act as switches and the waveform of the voltage on the common lead between the two transistors is approximately a square wave. Since the series tuned circuit L_2C_2 has ideally zero impedance at the fundamental frequency of this square wave, the fundamental component appears unattenuated across the parallel tuned circuit L_1C_1 to which the load R_L is connected.

Thus the peak voltage appearing across R_L is ideally $4/\pi$ times the amplitude of the square wave, and the latter is half the supply voltage if the transistors are assumed to be ideal switches.

At harmonic frequencies the series tuned circuit has a high impedance, so that only a small current at third and other odd-harmonic frequencies flows to the tuned circuit L_1C_1, i.e. the current in L_2C_2 is approximately sinusoidal. This current is supplied in turn by the two transistors, so that the collector current of each transistor alternates between approximately a half sine wave and zero during successive half cycles.

The resistors in the base circuits are made low enough to ensure that, when collector current is flowing, the base current is always adequate to keep the transistors bottomed, and since the base-current waveform also consists approximately of half sine waves, it keeps in step with the demand for collector current.

Thus it is evident that in this oscillator, at the instants when conduction is switched over from one transistor to the other, the transistor current is nearly zero but large voltage changes occur across the transistors, and the oscillator has therefore been termed a high-efficiency voltage-switching oscillator. These conditions are the converse of those applying to the oscillator described in Section 7, in which the full current is flowing at the switching instants but both transistor voltages are nearly zero. In both cases very little power dissipation occurs in the transistors during the switchover operation - an important advantage of these circuits over some others.

The detailed mechanism whereby conduction is transferred from one transistor to the other is discussed at length in the author's I.E.E. paper referred to in the Introduction to this article, and only brief comments will be made here.

If the waveform of the voltage on the lead joining the two transistors is observed on a C.R.O., it will be found that, whereas the waveform is approximately a square wave, there may be an overshoot or ringing during the transitions. These effects are influenced by hole-storage in the transistors, by whether or not L_2C_2 is tuned to precisely the same frequency as L_1C_1, and by the value of the effective bias voltage in the base circuit in relation to the h.t. voltage (in a practical circuit such as Fig. 9).

The effects so far referred to, unless present in a very exaggerated form, do not exert any great influence on the overall performance of the circuit. A condition which should be avoided, however, is that in which one or both transistors fail to remain bottomed during some part of their conduction time, such failure showing up as a dip in the top or bottom of the voltage square wave on the lead joining the two transistors. If this effect occurs anywhere near the middle of the conduction time, it results in greatly increased transistor dissipation, and the base resistors should be lowered in value to provide the increased base current necessary to maintain bottoming. In choosing the initial values for these resistors, due allowance should be made for the considerably reduced value of α' that is exhibited when a transistor is operated at high peak values of collector current.

8.2 Arrangements to ensure self-starting

In order to ensure that oscillation will begin to build up when the supply is switched on, it is necessary to arrange that both transistors are initially biassed into the conducting state. A system which achieves this result with very little extra power dissipation is shown in the practical circuit of Fig. 9, and involves the use of two diodes. Before oscillation has built up, current in R_3 and R_4 flows to the bases of the two transistors and the diodes are reverse biassed. There is d.c. negative feedback, since a change in the voltage on the lead joining the two transistors will increase one base current and decrease the other, and an equilibrium state for starting in which neither transistor is bottomed is therefore ensured.

When oscillation has built up, however, the mean base currents greatly exceed the currents in R_3 and R_4 and most of the current flows in the diodes. The capacitors across the diodes keep the diode voltages approximately constant once oscillation has built up, and permit the flow of reverse base current, due to hole storage, when the transistors are being switched off.

It is possible to replace the diodes by resistors, but R_3 and R_4 must then be made of much lower value and the power loss is then considerably greater.

8.3 Choice of tuned circuit values

Referring to the basic circuit of Fig. 8, the tuned circuit L_1C_1 has a loaded Q of approximately $R_L/\omega_0 L_1$. At the resonant frequency, $\omega_o/2\pi$, of the tuned circuits, the series resistance of the series tuned circuit is effectively R_L, so that, in this limited sense, the series tuned circuit may be said to have a loaded Q of $\omega_o L_2/R_L$. If these Q's are made equal, it may be shown that the peak magnetic energy will be the same in the two inductors, and if the inductors themselves have the same unloaded Q's, assumed much higher than the loaded values, then the same power dissipation will occur in each. These conditions are thought to be the most suitable choice for practical designs, the two inductors being made of the same physical size.

Normally a suitable value of loaded Q to adopt for both tuned circuits is 3, which keeps the tuned circuit losses very small, but since there are two tuned circuits contributing to reduction of harmonic distortion, these low losses are not accompanied by undesirably high distortion.

The third-harmonic distortion in the output voltage, assuming equal loaded Q values for the tuned circuits, is given approximately by:-

$$(\text{Third-harmonic distortion}) = \frac{3}{64\,Q^2} \times 100\% \qquad (7)$$

Thus, with the recommended Q of 3, the third-harmonic distortion is approximately 0.5%, and higher order odd harmonics (there are, ideally, no even harmonics) have a magnitude approximately inversely proportional to the cube of their order, and are therefore normally quite negligible.

8.4 A practical design

The experimental 45 kc/s oscillator shown in Fig. 9, which uses a pair of OC44's, has been in continuous operation for about 3500 hours at an output power of nearly 1.5 watts and an overall efficiency of about 85%.

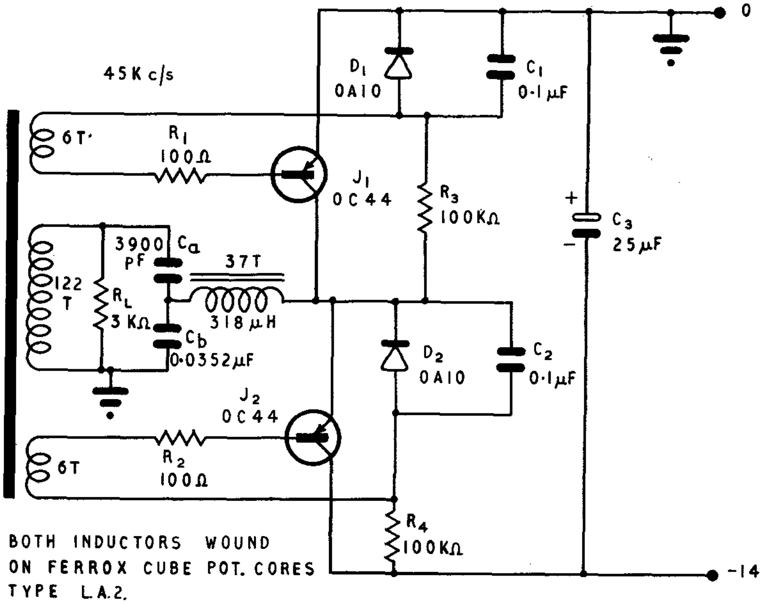


Figure 9 Practical Circuit

Despite the fact that the peak collector current is about 350 mA, a value far beyond the maker's rating, no change in any of the waveforms has been observed over this period. Even though the current is so high, the mean power dissipated by the transistors is much less than the maker's rating.

The measured third-harmonic distortion is 0.42%, and the measured second-harmonic distortion about 0.1%.

It will be seen that the arrangement of the tuned circuits in Fig. 9 is not the same as in the basic circuit of Fig. 8. On calculating values for the basic circuit it was found that L_1 came out inconveniently small and C_1 inconveniently large. The obvious solution would have been to replace L_1C_1 by a circuit of more practical L/C ratio and connect the L_2C_2 circuit to a tapping on the coil. It is possible to avoid the necessity for such a tapping, however, by exploiting the equivalence of the two basic tuned circuit configurations shown in Fig.3., and if the right ratio for $(C_a + C_b)/C_a$ is chosen, it is practicable to utilize $(C_a + C_b)$ as the tuning capacitance for the series tuned circuit, thus saving a capacitor. This technique has been adopted in the circuit of Fig. 9.

If the loaded Q-factors for the series and parallel tuned circuits are made equal, as recommended in Section 8.3, then the ratio of C_b to C_a that allows $C_a + C_b$ to be the only tuning capacitance for the series tuned circuit may be shown to be:-

$$\frac{C_b}{C_a} = Q^2 \qquad (8)$$

Thus, if Q = 3, as adopted in the Fig. 9 circuit, then $C_b = 9C_a$ and the effective transformer ratio, $(C_a + C_b)/C_a$, is 10. Thus the load resistance to be connected across the parallel tuned circuit is 100 times the resistance presented to the transistors at fundamental frequency.

8.5 Conclusions

Considering the oscillators described in Section 7 and the present Section, the former would seem better suited to widespread application, since it uses considerably fewer components, is capable of operating satisfactorily at higher frequencies, and, having only one tuned circuit, the frequency is more readily made adjustable. Nevertheless the present voltage-switching oscillator has the following advantages, which may suit it to particular applications:-

(a) The maximum voltage appearing across each transistor is approximately equal to the supply voltage, whereas in the oscillator of Section 7 it is πV_{dc}. Consequently the present oscillator can operate from π times the supply voltage of the previous design, and, for a given power output, will require only about $1/\pi$ of the supply current. For reasons which are gone into in the I.E.E. Paper, it follows that the present oscillator is capable of higher collector efficiency.

(b) For the same power output and efficiency, the present oscillator gives considerably less harmonic distortion at full-load; but if the load resistance is raised, the distortion remains approximately constant, whereas in the previous oscillator the distortion falls off as the load resistance is raised.

(c) The present oscillator does not suffer damage if the output is short-circuited, whereas in the previous one the transistors would immediately be damaged.

9. LONG-TAILED-PAIR OSCILLATOR

9.1 The basic circuit

In many low-power oscillator applications, high efficiency is not in itself of great importance, whereas simplicity, and a performance which is readily predictable and relatively independent of variations in transistor parameters, represent highly desirable characteristics. A simple circuit which has been found very useful for such applications is shown in Fig. 10 (a).

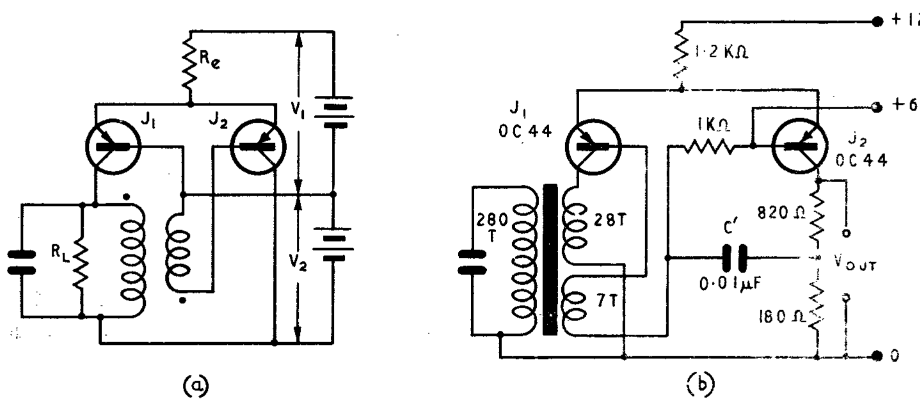


Figure 10 Long-Tailed-Pair Oscillators

In this circuit, the current flowing in R_e is switched alternately through J_1 and J_2 by the feedback voltage applied between the bases; the magnitude of this voltage is quite uncritical and one volt peak is sufficient to give satisfactorily quick switching action. When transistor J_1 is on, the collector current is given, to a close approximation, by V_1/R_e, so that the tuned circuit is fed with a current square wave of peak-to-peak magnitude V_1/R_e and a mark-to-space ratio of unity. The peak value of the fundamental component of this square wave is $2V_1/\pi R_e$, and the peak value, V_{max}, of the approximately sine-wave voltage at the collector is therefore given by:-

$$V_{max} = \frac{2V_1 R_L}{\pi R_e} \tag{9}$$

where R_L is taken to include the tuned circuit losses.

It is desirable to make V_{max} considerably less than V_2, so as to avoid any risk of J_2 bottoming. V_1 must at least be greater than the peak value of the fedback voltage, but by suitably choosing R_e it may be made as much larger than this as desired, to suit any available positive supply voltage.

Performance in reasonable agreement with (9), say within $\pm$ 20%, is obtained up to a frequency in the region of a quarter of the alpha-cut-off frequency of the transistors.

9.2 Harmonic distortion

The harmonic distortion is as given by equation (6) on page 28. 0.1% third-harmonic distortion requires a Q value of 125, which is often readily obtainable in practice.

9.3 Provision of square-wave output

The circuit may also be used to provide a square wave of current or voltage, controlled accurately in frequency by a high-Q tuned circuit, and this fact is exploited in the beat-frequency oscillator described in Section 10.

A voltage square-wave may be obtained by inserting a resistor in the collector lead of J_2. With the circuit otherwise as in Fig. 10(a), the square wave would have a rounded top, due to the current variation that results from the variation in base voltage while J_2 is conducting; the higher V_1 and R_e are made, the less the extent of this rounding. The effect may be eliminated, however, if the base of J_2 instead of that of J_1 is taken to the junction of the two batteries. With this modification, it is the wave-form of the current fed to the tuned circuit that is a distorted square wave, and it may then be necessary to reduce the dynamic resistance of the tuned circuit to prevent the voltage across J_1 reaching zero on peak positive excursions of the collector voltage, which coincide with peak negative excursions of the emitter voltage.

9.4 Modification to give trigger action

A much improved square-wave output may be obtained if the circuit is made to function in the same manner as a multivibrator during the changeover action only, and this may be achieved by the modification shown in Fig. 10(b). This circuit was designed for a recent application and generates a 12 kc/s square wave with rise and fall times of a fraction of a microsecond.

To obtain a nice flat top to the square wave, C' should be made very large. There is, however, a major snag to doing this, which is that on switching on the supply, the circuit oscillates simply as an emitter-coupled multivibrator, at some quite low frequency, instead of in the wanted mode. By momentarily connecting the emitter of J_2 to the negative supply line, the circuit may be got into the wanted mode; but the necessity for doing this is unlikely to be acceptable in most applications! (When operating in the wrong mode, the tuned-circuit waveform consists merely of a succession of small-amplitude exponentially-decaying transient oscillations.)

For completely reliable self-starting in the correct mode, the oscillation frequency of the circuit considered as a multivibrator should be higher than the tuned-circuit frequency. On switching on, the circuit then starts up as a multivibrator, going periodically through a state in which both transistors are conducting, and therefore in which there is positive feedback suitable for building up the sine-wave oscillation. Once the latter has reached sufficient amplitude, it takes charge of the frequency and the free-running multivibrator action ceases.

If C' is reduced too far, however, a burst of several cycles of multivibrator action may occur each time the sine-wave voltage is near zero.

In practice, however, there is a wide range of values for C' throughout which the desired behaviour is obtained, and in many applications the slightly curved top to the square wave is of no concern.

10. BEAT-FREQUENCY OSCILLATOR (Pat. Appn. No. 31841/59)

10.1 General

This beat-frequency oscillator is a good illustration of an application in which the use of transistors has made it possible to obtain a very high-grade performance much more easily and economically than would be the case using valves. The main advantages associated with the use of transistors are:-

(a) The total power consumption being in the region of 200 mW only, the problem of reducing the warm-up frequency drift to an adequately low value is greatly eased.

(b) By using a switching type of frequency-changer or phase-sensitive-rectifier, with the transistors operated in the inverted mode [10], [11], very little distortion is introduced by the frequency-changer and, in addition, the coupling between the high-frequency oscillators introduced via the frequency-changer can be made so low that buffer stages become unnecessary*.

(c) By employing high-frequency oscillators of the switching type, as described in Section 9, in combination with the above-mentioned switching type of frequency changer, the random noise output from the B.F.O. is made very small, being more than 70 dB below the signal output; and since the instrument is battery operated, there is no hum output.

---

*In this latter respect the balanced switching circuit employed is much preferable to that used by Mayo[12], [13], but the basic philosophy is otherwise much the same.

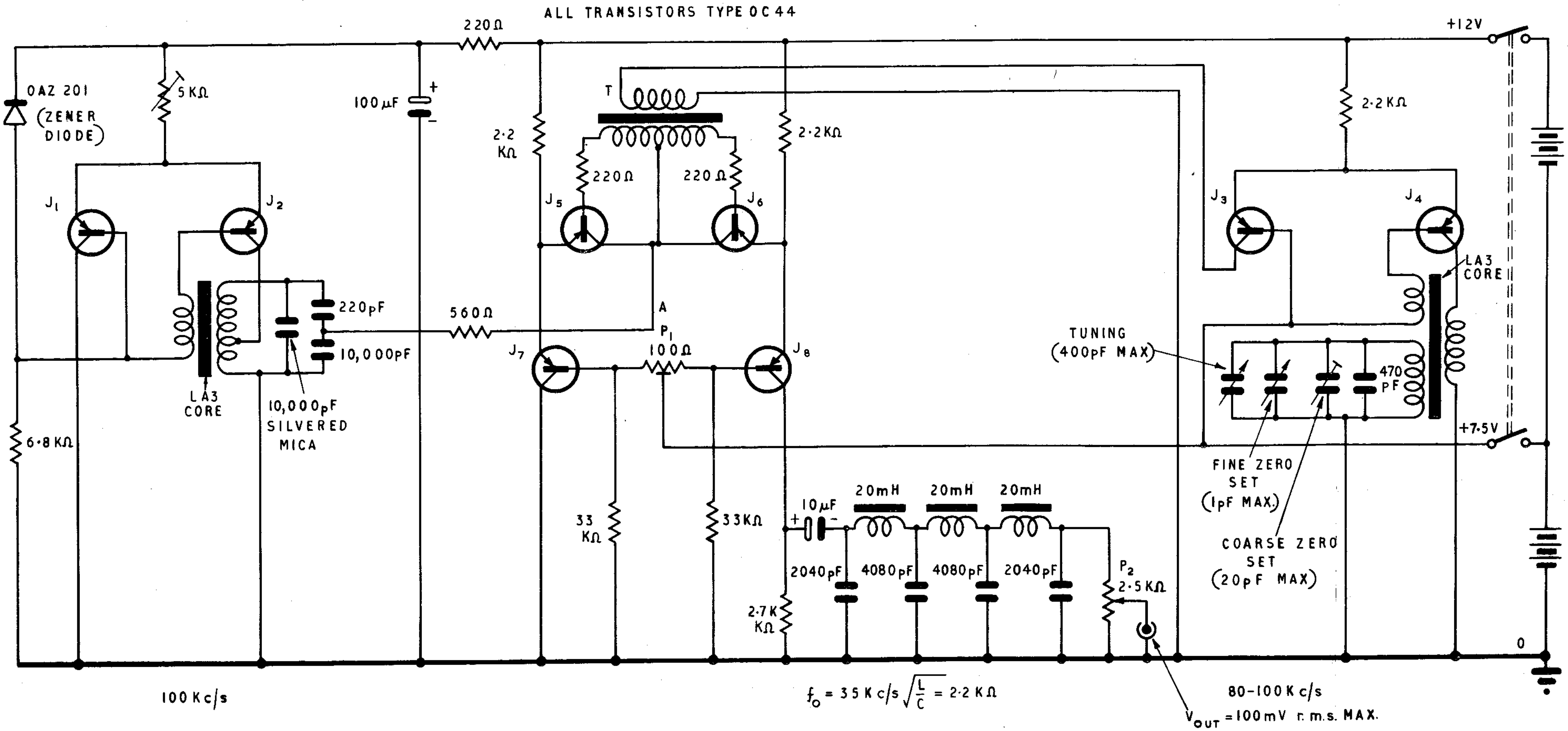


FIG. 11. BEAT-FREQUENCY OSCILLATOR.

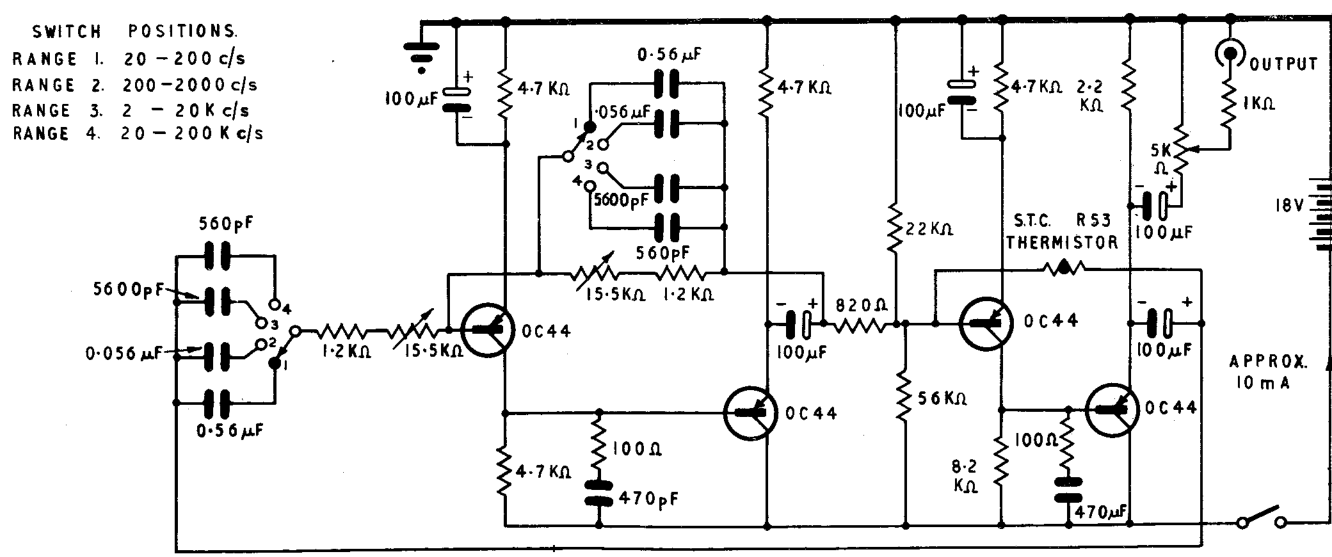


FIG 12 R-C OSCILLATOR

10.2 Description of circuit

Referring to the circuit diagram of Fig. 11, the variable-frequency oscillator, employing transistors J_3 and J_4, supplies a square wave of current via the transformer T to the switching transistors J_5 and J_6, which causes them to switch a sine-wave current, from the fixed-frequency oscillator, alternately to the emitters of J_7 and J_8 respectively. J_7 and J_8 operate as common-base amplifiers, with low input impedance, so that very little a.c. voltage is produced at their emitters, but nearly all the a.c. input current appears in one or other of their collector circuits. The current in J_8's collector circuit is used to provide the wanted beat-frequency output, via a low-pass filter whose function is to remove the high-frequency components.

Thus, as far as current reaching J_8 collector is concerned, the sine-wave input current is effectively multiplied by a square-wave function going alternately, at the variable oscillator frequency, between 0 and 1. This is equivalent to multiplying by the sum of two functions, i.e. a constant term of value $\frac{1}{2}$ plus an a.c. square wave going between $+\frac{1}{2}$ and $-\frac{1}{2}$. The a.c. square wave may further be resolved into a fundamental plus odd harmonics. The wanted beat frequency is due to the multiplication of the sine-wave input (f_1) by the fundamental component of the square wave (f_2), and has a frequency of ($f_1 - f_2$). If the sine-wave input actually contains some distortion, which is inevitably the case, then the effects are as follows:-

(a) If the sine-wave input contains x% of third harmonic ($3f_1$), then an output at ($3f_1 - 3f_2$) will occur, due to the multiplication of the $3f_1$ distortion input by the third harmonic ($3f_2$) of the square wave. However, since the third harmonic of a square wave is only of one third the amplitude of the fundamental, the percentage third-harmonic distortion in the beat-frequency output will be only x/3. Similarly the percentage of fifth-harmonic in the output will be only one fifth the percentage present in the fixed-frequency oscillator output, etc.

(b) Any even-harmonic distortion in the sine-wave input would ideally product no distortion at all in the beat-frequency output, since the square wave should contain no $2f_2$ term to give rise to a $(2f_1 - 2f_2)$ output term.

With a Q of 200 in the fixed-frequency oscillator tuned circuit, the third-harmonic distortion in the oscillator output is approximately 0.06%, giving rise to 0.02% third-harmonic distortion in the beat-frequency output - a figure which is verified approximately by direct measurement of the output distortion.

The measured second-harmonic distortion in the beat-frequency output from the Fig. 11 circuit is approximately 0.04%, measured at 1000 c/s, and most of this distortion is generated in the common-base stage, J_8.

At low beat frequencies there is another mechanism that can cause second-harmonic distortion in the beat-frequency output, i.e. periodic slight pulling of the frequency of one primary oscillator due to unwanted coupling with the other primary oscillator. Mayo[14] has deduced that the second-harmonic distortion in the beat-frequency output due to this pulling mechanism is given by:-

$$\text{(Second-harmonic distortion)} = \frac{f_o}{2\ f} \times 100\% \qquad (10)$$

where f is the beat frequency at which the distortion is determined and f_o is the beat-frequency setting at which the primary oscillators just lock together.

It is difficult to determine f_o experimentally with any great accuracy, due to slight frequency instability of the oscillators, but it is certainly less than 0.1 c/s. This would give, according to (10), 0.125% second-harmonic distortion at 40 c/s. The measured distortion at 40 c/s is well under 0.1%.

P_1 enables the d.c. **voltage** between the bases of J_7 and J_8 to be so adjusted that point "A" remains at the same potential no matter whether it is J_5 or J_6 that is in a conducting state. If P_1 is incorrectly set, a square wave at the frequency of the variable oscillator appears at the point "A", causing a current to be fed through to the fixed-frequency oscillator, which increases the tendency of the two oscillators to lock together, and hence increases the second-harmonic distortion at low beat frequencies. With P_1 correctly set and the fixed-frequency oscillator inoperative, the waveform at "A" consists of a series os small pulses, accompanied by slight ringing, having an amplitude of about 20 mV and a p.r.f. of twice the variable-oscillator frequency. A small amount of fundamental component may be observable on careful inspection.

With a switching type of frequency changer, the amplitude of the beat-frequency output is, ideally, independent of the amplitude of the square-wave switching-current input and is dependent only on the amplitude of the sine-wave input. Since the latter comes from the fixed-frequency oscillator, the beat-frequency output amplitude should be independent of frequency. In practice the only significant cause of output variation with frequency is the lack of perfectly flat response in the low-pass filter over the wanted frequency band. With the circuit as shown in Fig. 11, the output remains within ± 0.25 dB over the working frequency range of 20 c/s to 20 kc/s. In the final version of this B.F.O., however, an output emitter follower and calibrated 600 ohm attenuator have been added, and a small amount of response-correction has been incorporated, resulting in an output that remains within ± 0.1 dB over the working frequency range.

It will be seen that the supply voltage for the base circuit of the fixed-frequency oscillator is stabilized by means of a Zener diode - the transistor counterpart of a neon stabilizer - thus making the beat-frequency output amplitude rather less dependent on battery voltage.

In its present form, the instrument exhibits a drift of beat frequency, under ordinary laboratory conditions, of less than 2 c/s in an hour.

The problem of covering the frequency range 20 c/s to 20 kc/s on a single dial with a usable scale shape has been solved by interposing a variable-velocity-ratio link motion between the capacitor shaft and the dial shaft, the capacitor itself being of the type used in ordinary broadcast receivers. The link motion is simple and cheap, spring-loaded conical bearings of no great precision being used; there is no noticeable backlash.

This instrument was exhibited at the International Transistor Exhibition in London, and is being manufactured by Messrs. Roband Electronics Ltd.

11. THERMISTOR-CONTROLLED R - C OSCILLATOR (Pat. Appn. No. 44213/59)

11.1 General

In some of the well-known valve R - C oscillator circuits, the high input impedance at the grid of a valve is a necessary attribute for the correct functioning of the circuit. Since transistors do not possess this high-input-impedance feature, the direct adaptation of existing valve circuits to use transistors is not always satisfactory.

11.2 Circuit Description

An unusual approach to the problem has been adopted in a circuit devised by the author and shown in Fig. 12. In this circuit, the R - C phase-shift elements are connected in the input and feedback arms of a two-stage negative-feedback amplifier of the virtual-earth type. A second, and similar, two-stage amplifier has a thermistor in its negative-feedback arm and the amplifier is connected in cascade with the first one. By connecting the output of the second amplifier back to the input of the first, oscillation is obtained and the thermistor adjusts the loop gain to unity at a certain stable amplitude of oscillation; ideally the conditions are then that the first feedback amplifier gives a gain of $\frac{1}{2}$ and the second a gain of 2.

The oscillator gives an output of approximately 1 volt r.m.s. at less than 0.05% total harmonic distortion, and the output voltage varies by less than ± 0.1 dB throughout the working frequency range of 20 c/s to 200 kc/s. (The distortion figure just mentioned is not applicable at very low or very high frequencies; below about 100 c/s the thermistor temperature is able to follow the fluctuations in heating power during each cycle to a significant extent, and this gives rise to an increase in the third-harmonic distortion - the measured third-harmonic distortion at 40 c/s is 0.08% approximately, compared with 0.02% at 1000 c/s. Some increase in distortion with rise of frequency would also be expected to appear in the 20 to 200 kc/s band, since there is a considerable fall off in forward gain in each feedback amplifier in this frequency range.)

The unusually small variation in output voltage with frequency results from the fact that, in this circuit, the oscillator output voltage is also, to a close approximation, the total voltage across the thermistor; and since the thermistor is operated approximately at the maximum of its current-voltage characteristic - see Fig. 13 - , small errors in the tracking of the two-gang frequency-determining pot, or in the matching of the associated capacitors, have almost no effect on the output voltage.

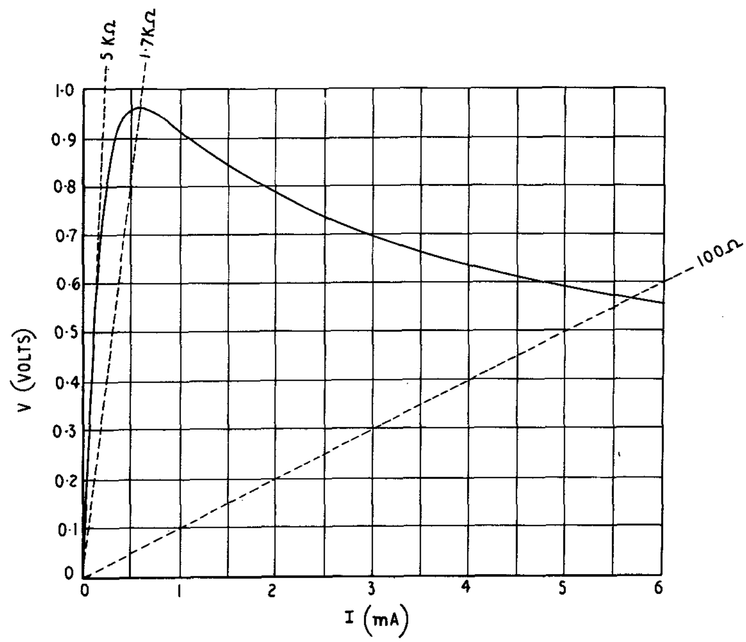


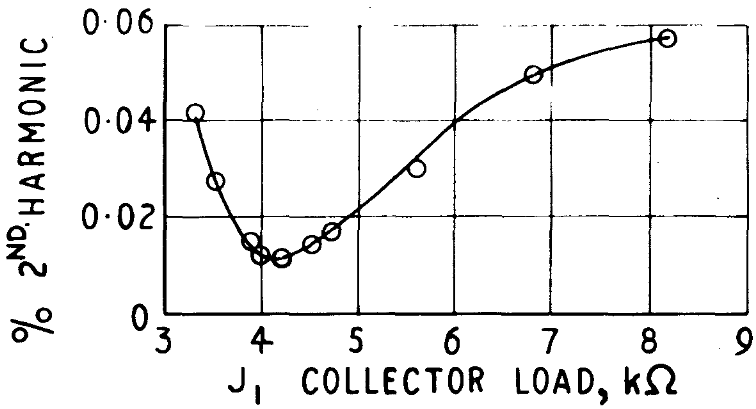
Figure 13 Thermistor Characteristic

The variation in output voltage with supply voltage is also remarkably small - a 50% change in supply voltage giving less than 0.1 dB change in output voltage.

It should be pointed out that the termistor used is one of a new range of low-power thermistors introduced by S.T.C. specifically for use in transistor circuits. These new Type R thermistors, when operated at the maximum voltage point, dissipate just over half a milliwatt only, and develop a temperature rise of about 30°C. The older Type A thermistors, widely used in thermistor-controlled valve oscillators, require over ten times as much power to reach comparable operating conditions.

The circuits associated with the 100μF capacitors in Fig. 12 introduce some attenuation and phase shift at the low-frequency end of the range (e.g. at 20 c/s), but since these effects occur inside the forward path of the feedback amplifiers, their influence on the fedback gain and phase of the amplifiers is extremely small and does not greatly affect the frequency calibration on the lowest-frequency range.

There is an interesting point relating to the choice of value for the collector load resistor of transistor J_1. The distortion introduced by the transistors themselves is mainly second-harmonic distortion, and the first transistor in each two-stage amplifier introduces much more distortion than the second (emitter-follower) stage. With respect to the final output waveform, the distortion introduced by J_1 is in antiphase with that introduced by J_3, so that if these distortions can be made of approximately equal magnitude, they can be made to almost cancel out. The sort of effect produced is sketched below, and whereas the exact curve obtained varies to some extent with the setting of the frequency dial, a collector load of 4.7 kΩ keeps the distortion under 0.02% at all settings.



An oscillator based on the Fig. 12 circuit is now in production by Messrs. Solartron Ltd.

11.3 Amplitude Bounce Effects

An undesirable feature of this oscillator is that the output amplitude bounces about a good deal while the frequency dial is being rotated.

At the time the lecture was given it was believed that this trouble was inherent in thermistor-controlled oscillators using wire-wound potentiometers for frequency control - and it seems that this view is very generally held. More recently, however, it has become clear that such bounce effects can be largely avoided if the right kind of phase shift network is used.

If, in the Fig. 12 circuit, one half of the ganged potentiometer suddenly changes in resistance by one winding turn, without the other half changing, then there will be a sudden small change in gain. Being a class 'A' oscillator, this change in gain will cause the oscillation amplitude to begin to rise (if a gain increase occurred) or begin to fall (if a drop in gain occured). Later, the thermistor acts to correct this gain change, but there is an initial bounce in amplitude before this occurs. The effect is particularly pronounced at high frequencies, since more cycles of oscillation can then occur before the thermistor has had time to readjust the gain.

There are some RC circuits, however, in which a change in value of one of the R's causes no change in gain but only a change in phase. An example is the well-known phase-shifting circuit in which a series combination of C and R is connected across a push-pull input source. The voltage between the junction of C and R and earth will then be independent of both the value of R and the frequency, these quantities affecting only the phase of this voltage. An RC oscillator may be made using two such networks, each giving 90° phase shift, and will be found to be almost free from any tendency to bounce.

There are other circuits possessing the above desirable characteristic in various degrees. One of these developed by J.H. Wood and the author, has enabled a very satisfactory thermistor-controlled oscillator to be made using only three transistors and a single linear wire-wound potentiometer for tuning. A 10 : 1 frequency sweep is obtainable with a good scale shape. The distortion is less than 0.1%. A Patent has been applied for.

12. CRYSTAL-CONTROLLED OSCILLATORS

12.1 Low-Level class A oscillator

The 80 kc/s oscillator shown in Fig. 14(a), designed by S.W. Noble, operates under class A conditions at the series resonant frequency of the crystal, the mean power dissipated in the crystal being about 15 μW only.

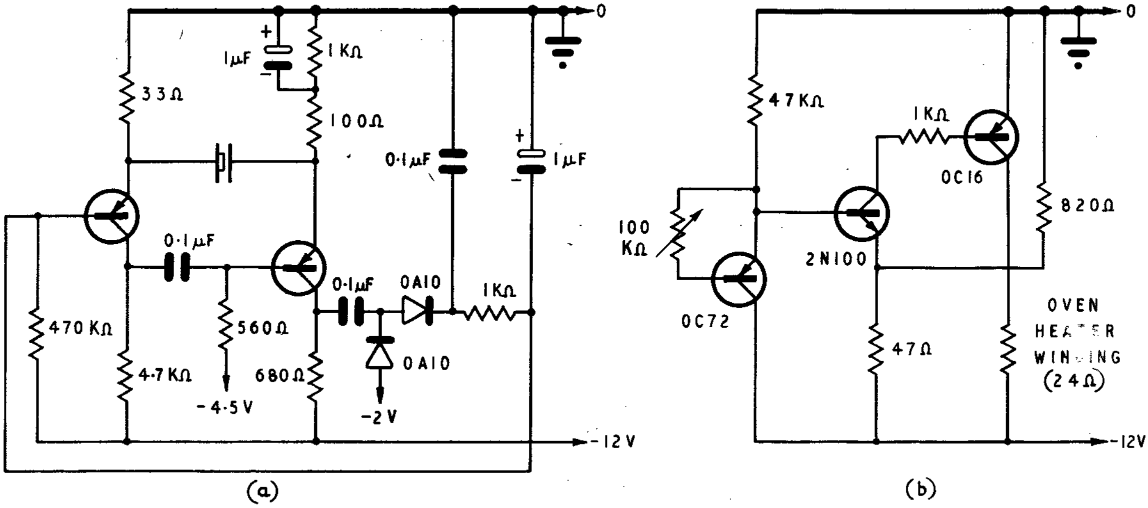


Figure 14

(a) 80 kc/s Crystal Oscillator (b) Crystal-Oven Circuit

It will be seen that the collector of the first transistor is R - C coupled to the base of the second transistor, the latter functioning as an emitter follower and giving positive feedback, via the crystal, back to the emitter of the first transistor. The loop gain round this positive-feedback path is dependent on how much of the crystal current goes into the 33 ohm emitter resistor and how much goes into the transistor.

The small-signal impedance seen looking into the emitter of a transistor, with a negligibly-low impedance in the base circuit, is $r_e + r_b(1 - \alpha)$, of which the first term, r_e, will be much larger than the second at working currents of the order of 1 mA, especially with a high-frequency transistor. The value of r_e is given by :-

$$r_e = \frac{kT}{q\,I_e} = \frac{25}{I_e} \qquad (11)$$

where I_e is the d.c. emitter current in mA and r_e is in ohms.

Thus, by changing the d.c. emitter current, the division of a.c. crystal current between the 33 ohm emitter resistor and the transistor may be varied, and hence the loop gain may be controlled.

The d.c. emitter current is actually varied automatically, so as to give just unity loop gain, by a backed off diode rectifier in the collector circuit of the second transistor; this rectifier starts to bias the transistor base positively as soon as the signal amplitude at the collector of the second transistor exceeds a certain value.

It should be emphasised that the transistor parameter, r_e, that is being used to control the gain is one that can be predicted very closely; at a given emitter current it varies hardly at all from one transistor sample to another. This parameter is, in fact, much more accurately predictable than any of the normal parameters of a thermionic valve. It is worth bearing this fact in mind for exploitation on suitable occasions.

The crystal is a G.E.C. Type JCF/195 unit, and has an effective series resistance, at the operating level adopted, of approximately 265 ohms. The Q is over 200, 000, which makes the control system quite sluggish in operation, though this does not matter.

This oscillator was actually built for a demonstration model quartz crystal clock, employing some of G.H. Perry's ferrite core and transistor circuits for the frequency division. To improve the timekeeping of the clock, a simple temperature-controlled oven for the crystal was devised by S.J. Widdows, and its circuit is shown in Fig. 14(b).

An OC72 transistor is used as the temperature-sensitive element. Its d.c. operating conditions are determined in what would be, in more normal circumstances, the worst possible way - i.e. a large resistance is placed in the base circuit and the transistor is biassed on due to the flow of i_{co} in this resistance. Thus the collector current is made highly temperature dependent, and by adjusting the magnitude of the base resistance, the complete circuit may be made to control at a chosen temperature. Notice that the use of one NPN transistor in the d.c. amplifier has enabled a very cheap and simple circuit to be employed.

This system has worked very well and holds the oven temperature at $35^{\circ}C \pm \frac{1}{2}^{\circ}C$ for months on end. No very thorough determinations of the timekeeping of the clock have been made, but its rate is certainly constant to better than $\pm$ 0.1 second per day.

12.2 Switching type crystal oscillator

The circuit shown in Fig. 15 is based on a design due to N.F. Moody[(15)], though the version published actually used NPN transistors.

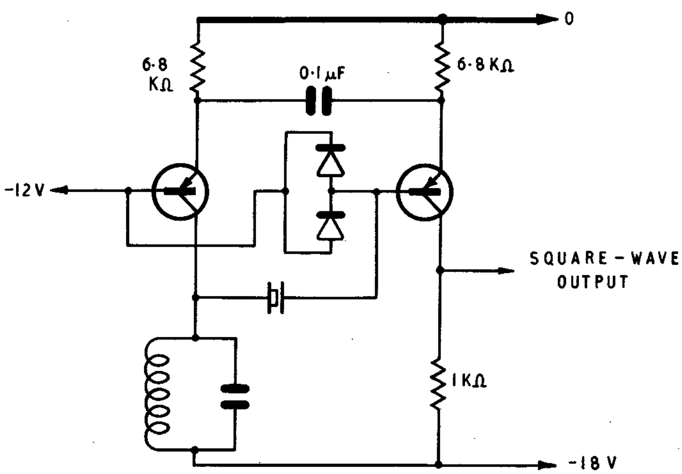


Figure 15 Crystal Oscillator (N.F. Moody)

The crystal operates at its series resonant frequency and the crystal current is approximately sinusoidal. Most of this crystal current flows into one or other of the two junction diodes, and the diodes limit the peak-to-peak voltage excursion on the right-hand transistor base to something in the region of 1 volt*. This base voltage is sufficient to switch conduction quite quickly from one transistor to the other, so that the

---

* The use of two junction diodes connected directly in parallel in this manner, to perform the function of a voltage limiter with built-in delay voltage, is a very satisfactory system with modern junction diodes, which pass almost no forward current until a few tenths of a volt of forward voltage is applied. One very potent application is in protecting meters from overload.

collector-current waveform of the right-hand transistor is an approximate square wave of well defined amplitude. The output may be taken as a square-wave voltage by inserting a resistor as shown, or alternatively one or more tuned circuits may be inserted, tuned to the fundamental and/or odd harmonic frequencies. In the original application, a 1 Mc/s crystal (having an effective series resistance at resonance of about 200 ohms) was used, and outputs were obtained at 1 Mc/s and 3 Mc/s.

At 1 Mc/s the diodes would need choosing with some care, as excessive hole-storage would seriously impair the performance; Mullard OA5 or G.E.C. EW78 or EW781 would probably be satisfactory. (The original design specified a type of Texas diode not normally available here). OC44's of high f_α, or preferably Philco SB100's, should be suitable for the transistors at 1 Mc/s.

At first sight it might appear reasonable to join the two emitters directly together and use a single tail resistor. It was found, however, that this arrangement did not give reliable self starting. This is hardly surprising, since the impedance in the base circuit of the right-hand transistor at low frequencies, when neither diode is appreciably conducting, is extremely high, and a small amount of i_{co} would be sufficient to hold the base at approximately the negative limit of its possible excursion, thus getting the circuit into a state in which the right-hand transistor is permanently on and the left-hand one is permanently off. The arrangement shown, however, avoids the possibility of this occurring.

In some circumstances an adequately good performance might be obtainable if the two diodes were replaced by a resistor, in which case a single tail resistor could safely be used.

REFERENCES

(1) HUNTER, L.P.: "Handbook of Semiconductor Electronics", (McGraw-Hill, Inc., 1956).

(2) GOURIET, G.G.: "High-Stability Oscillator", Wireless Engineer, 1950, 27, p.105.

(3) CLAPP, J.K.: "An Inductance-Capacitance Oscillator of Unusual Frequency Stability", Proc. I.R.E., 1948, 36, p.356.

(4) BAXANDALL, P.J., GIRLING, F.E.J., GOOD, E.F., and NOBLE, S.W.: "RC Cathode-Follower Oscillators", Wireless Engineer, 1956, 33, p.129.

(5) EDSON, W.A.: "Vacuum Tube Oscillators" (John Wiley and Sons, Inc., 1953), p.227.

(6) MULLARD Ltd.: "Mullard Ferroxcube", (Book of design data issued by Mullard Ltd.)

(7) SALFORD ELECTRICAL INSTRUMENTS Ltd.: "Gecalloy Radio Cores", (Booklet of design data issued by S.E.I.Ltd., Times Mill, Heywood, Lancs.)

(8) TYLER, V.J.: "A New High-Efficiency High-Power Amplifier", Marconi Review, 1958, 21, p.96.

(9) HEAD, J.W., and MAYO, C.G.: "Combined Limiter and Discriminator", Radio and Electronic Engineer, 1958, 35, p.85.

(10) CHAPLIN, G.B.B., and OWENS, A.R.: "Some Transistor Input Stages for High-Gain D.C. Amplifiers" Proceedings I.E.E. Part B, May 1958 (Vol. 105, p.249)

(11) AVERY, D.G., and BOWES, R.C.: "A Design for a Multi-Channel Infra-Red Spectrometer Using Transistor Electronics", Journal of Scientific Instruments, 1958, 35, p.214.

(12) MAYO, C.G., and BEADLE, D.G.: "Equipment for Acoustic Measurements", Electronic Engineering, 1951, 23, p.368.

(13) MAYO, C.G.: "Beat-Frequency Tone Source", Wireless Engineer, 1952, 29, p.148.

(14) MAYO, C.G.: "Distortion in Beat-Frequency Sources", Wireless Engineer, 1952, 29, p.92.

(15) LAKE and TENNUCI: "A Doppler Radar Modulator Using Transistors", Report No. 5088-1 of Defense Research Telecommunications Establishment, Electronics Lab., Ottawa.

THE FUNDAMENTAL PRINCIPLES OF TRANSISTOR CIRCUITS

LECTURE 5. THE TRANSISTOR AS A SWITCH

by S.W. Noble

1. INTRODUCTION

For the discussion of the small signal properties of a transistor in the first lecture, a 'one dimensional' symmetrical structure was considered in the interests of mathematical tractability. Although alloyed junction transistors, especially if designed primarily for the best possible performance as small signal linear amplifiers, are usually made geometrically unsymmetrical with the collector area much greater than the emitter area, it is nevertheless possible with fair accuracy to discuss their small signal properties in terms of equivalent symmetrical 'one dimensional' structures. That this is possible arises from the fact that, for linear amplification purposes, one has no interest whatsoever in the performance of the transistor under 'inverted' operation conditions, that is to say, with the normal collector electrode used as an emitter and the normal emitter used as a collector.

Generally speaking, in applications of a transistor as a switching element, it is necessary to consider the properties of the device in rather more general terms, since at different times one and the same device may be required to operate with:-

(a) Normal emitter junction forward biassed and normal collector junction reverse biassed.

(b) Normal emitter junction reverse biassed and normal collector junction forward biassed.

(c) Both junctions forward biassed.

(d) Both junctions reverse biassed.

Regime (a) above is often termed the normal active state, regime (b) the inverted active state, regime (c) the 'on' state and regime (d) the 'off' state of the transistor.

For purposes of analysis, it would be rather unrealistic to discuss switching behaviour in terms of an idealized 'one dimensional' symmetrical transistor and a more generalized geometry as indicated in Fig. 1 will be assumed. No attempt will be made to relate the important circuit parameters quantitatively to the geometry of the device, but certain relationships between some of these parameters will be deduced, such relationships being independent of the assumed geometry. For simplicity, the following assumptions will be made:-

(1) That the emitter and collector regions are of similar uniform P-material and that voltage drops within these regions are negligible.

(2) That the base region is of uniform N-material.

(3) That the junction injection efficiencies are approximately unity.

(4) That fields in the base region are insufficiently great to influence significantly the transport of holes between the emitter and collector boundaries.

(5) That effects of depletion layer width modulation (Early effect) may be neglected.

Of the above assumptions, (3) and (4) are likely to be seriously violated at high current densities [see lecture 1, section (4.2.4)] and hence some modifications to the quantitative results obtained from the analysis may be expected to be necessary under these conditions. Nevertheless the essential qualitative basis on which the theory of switching behaviour is founded can still be expected to hold good even in circumstances where exact performance calculations are impracticable.

The static properties, that is the relationships between currents and voltages under steady state conditions, of a transistor switch will first be discussed, after which the transient behaviour of the device during the finite periods of time required to effect transitions between the off and on states will be dealt with. The discussion will in no sense be exhaustive, and no more than an attempt to present the essential elementary theory will be made.

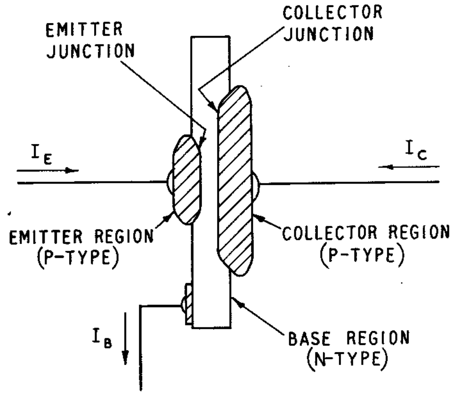


Figure 1

2. THE RELATIONSHIPS BETWEEN THE EMITTER AND COLLECTOR CURRENTS AND THE JUNCTION VOLTAGES.

Referring to Fig. 1, it should be noted that if no electric fields exist within any regions of the transistor other than the depletion layers at the two junctions, all the elements of a junction will be biassed to the same extent. In consequence, no matter what be the shapes of the emitter-to-base and collector-to-base junctions, the minority carrier densities in the base region at all points in the neighbourhood of a junction are equal, and of a value dependent only on the junction voltage (lecture 1, equation 22). Let us denote the carrier density at the emitter junction by p_I, the carrier density at the collector junction by p_2, and the equilibrium carrier density in the base region by p_e.

With no voltage applied to either junction the carrier density everywhere in the base will be p_e. If now the collector junction voltage be held at zero but a small external voltage be applied to the emitter junction, we shall have $p_1 = p_e + (\Delta p)_1$, $p_2 = p_e$, where $(\Delta p)_1$ is the change in carrier density at the emitter junction caused by the external voltage. The emitter and collector currents will each be linearly related to $(\Delta p)_1$ and we can, in fact write:-

$$I_E = k_{11}(\Delta p)_1 \qquad (1)$$

$$I_C = k_{21}(\Delta p)_1 \qquad (2)$$

where k_{11} and k_{21} are constants. At this point it should be noted that with the positive directions of I_E and I_C being taken as towards the transistor (Fig. 1), in contrast with the convention adopted previously (see Fig. 5(a) of lecture 1), k_{11} will be positive and k_{21} negative. Also the ratio $(-k_{21}/k_{11})$ will be the normal α of the transistor, which in future will be denoted by α_N. The fact that the collector voltage has been maintained at zero ensures that no collector cut-off current, I_{Co}, can flow under the above conditions. Now consider the situation with the transistor operating under inverted conditions, that is with the emitter on short-circuit load and with a small external voltage applied to the collector. The emitter is now acting as a collector and vice versa. We shall now have $p_1 = p_e$, $p_2 = p_e + (\Delta p)_2$ and we can write:-

$$I_E = k_{12}(\Delta p)_2 \qquad (3)$$

$$I_C = k_{22}(\Delta p)_2 \qquad (4)$$

Here k_{12} is negative and k_{22} is positive, the ratio $(-k_{12}/k_{22})$ being the inverse α of the transistor, denoted by α_I. [We may note that the sort of geometry indicated in Fig. 1, which is favourable for a high value of α_N, tends to cause a rather low value of α_I, since a considerable proportion of the holes injected from the outer regions of the collector have little chance of reaching the emitter before recombining either at the outer surface of the base region (the main hole 'sink' at low current densities) or within the bulk volume of the base material itself (the main hole 'sink' at high current densities). For a medium frequency small alloyed transistor, typical values of α_N and α_I are about 0.98 and 0.8, respectively, as pointed out in R.C. Bowes' second lecture, section (2.1). On the other hand for a specially designed transistor with symmetrical geometry such as the Mazda XS101, a value of 0.95 for both α_N and α_I may be regarded as reasonable.]

We can now consider the situation in which both junctions are biassed simultaneously such that $p_1 = p_e + (\Delta p)_1$ and $p_2 = p_e + (\Delta p)_2$. On account of the linearity of the equations for I_E and I_C in terms of $(\Delta p)_1$ and $(\Delta p)_2$, we can apply the principle of superposition and determine I_E by adding the right-hand sides of (1) and (3) and likewise I_C

by adding the right-hand sides of (2) and (4). We thus get:-

$$I_E = k_{11}(\Delta p)_1 + k_{12}(\Delta p)_2 \qquad (5)$$

$$I_C = k_{21}(\Delta p)_1 + k_{22}(\Delta p)_2 \qquad (6)$$

An important analogy can now be observed. Consider a linear, passive electrical network of any degree of complexity having two pairs of external terminals. A voltage V_1 is placed across one pair and the current I_2 flowing in a short-circuit load connected across the second pair is measured. The ratio I_2/V_1 can be expressed as a <u>transfer admittance</u> Y_{21}. Next apply a voltage V_2 to the <u>second</u> pair of terminals and measure the current I_1 flowing in a short-circuit load across the <u>first</u> pair. The ratio I_1/V_2 can be denoted by Y_{12}. The <u>reciprocity principle</u>, which is proved in standard works on the theory of linear networks states that $Y_{21} = Y_{12}$. For such a network, the current-voltage relationships so far as the external terminals are concerned can be written:-

$$I_1 = Y_{11}V_1 + Y_{12}V_2 \qquad (7)$$

$$I_2 = Y_{21}V_1 + Y_{22}V_2 \qquad (8)$$

Equations (5) and (6) can be regarded as characterizing a certain type of linear, passive 'network' which is analogous to the linear, passive electrical network characterized by equations (7) and (8). The 'network' to which (5) and (6) relate must be regarded as the whole effective base region of the transistor, bounded by the two junctions and the external surface. The quantities in equations (5) and (6) analogous to V_1 and V_2 in equations (7) and (8) are the incremental carrier densities $(\Delta p)_1$ and $(\Delta p)_2$ in the junction neighbourhoods. The four 'k' coefficients are analogous to the respective four 'Y' admittance parameters, and we note from the reciprocity relationship that $k_{12} = k_{21}$. It should be noted that the conditions for the base 'network' to be passive and linear are not violated by variations in surface and volume recombination factors in different parts of the base region. If, however, electric fields exist in this region, the conditions for the 'network' to be both linear and passive are violated, and the reciprocity relationship can no longer be presumed to hold. We shall show below that, as a result of the reciprocity relationship, only three of the four quantities α_N, α_I, I_{Eo}, I_{Co} are independent.

We next require to express I_E and I_C in terms of the junction voltages. We first re-write equations (5) and (6) in the form:-

$$I_E = k_{11}(p_1 - p_e) + k_{12}(p_2 - p_e) \tag{9}$$

$$I_C = k_{21}(p_1 - p_e) + k_{22}(p_2 - p_e) \tag{10}$$

Then, if V_E and V_C be the respective junction voltages, we have (lecture 1, equation 22):-

$$p_1 = p_e \exp \frac{qV_E}{kT}, \quad p_2 = p_e \exp \frac{qV_C}{kT}$$

Equations (9) and (10) now become :-

$$I_E = a_{11} \left[\exp \frac{qV_E}{kT} - 1 \right] + a_{12} \left[\exp \frac{qV_C}{kT} - 1 \right] \tag{11}$$

$$I_C = a_{21} \left[\exp \frac{qV_E}{kT} - 1 \right] + a_{22} \left[\exp \frac{qV_C}{kT} - 1 \right] \tag{12}$$

where a_{11}, a_{12}, a_{21}, a_{22} = $p_e k_{11}$, $p_e k_{12}$, $p_e k_{21}$, $p_e k_{22}$ respectively, and where, from the reciprocity relationship, $a_{12} = a_{21}$.

3. RELATIONSHIPS BETWEEN THE 'a' COEFFICIENTS AND α_N, α_I, I_{Eo}, I_{Co}.

We shall assume that I_E and I_C can be expressed in terms of α_N, α_I, I_{Eo} and I_{Co} by the pair of equations:-

$$I_E = I_{Eo} - \alpha_I I_C \left[\text{for } V_E \text{ highly negative} \right] \tag{13}$$

$$I_C = I_{Co} - \alpha_N I_E \left[\text{for } V_C \text{ highly negative} \right] \qquad (14)$$

In the above equations, the cut-off current I_{Eo} is defined as the current which flows to the emitter under conditions of heavy reverse voltage bias and with the collector junction on open circuit (not the current which flows when both junctions are reverse biassed). The cut-off current I_{Co} is defined in a similar manner.

Since the four 'a' coefficients in (11) and (12) are proportional to the corresponding 'k' coefficients in (9) and (10), we have, from the previous section:-

$$\alpha_N = - \frac{a_{21}}{a_{11}} \qquad (15)$$

$$\alpha_I = - \frac{a_{12}}{a_{22}} \qquad (16)$$

The procedure for determining I_{Eo} and I_{Co} is a little more complicated. If we make V_E highly negative and V_C zero, we have, from (11) and (12), $I_E = - a_{11}$, $I_C = - a_{21}$. It is important to note here that we could equally well write $I_C = - \alpha_N I_E$ and that we must regard the application of a reverse bias to a junction as producing a reverse current at this junction which will in turn lead to a proportional forward increment of current at the other junction if its bias is held constant - α_N and α_I must be regarded as applicable to either direction of current. We next reduce I_C to zero by placing a negative bias on the collector of sufficient magnitude to cause an incremental change of current a_{21}. This change in collector current produces a corresponding change in emitter current of magnitude $- \alpha_I a_{21}$, which, from (16), is equal to $a_{12}a_{21}/a_{22}$. The total emitter current is now, by definition, equal to I_{Eo}, and we have:-

$$I_{Eo} = - a_{11} + \frac{a_{12}a_{21}}{a_{22}} \qquad (17)$$

Similarly, $$I_{Co} = - a_{22} + \frac{a_{21}a_{12}}{a_{11}} \qquad (18)$$

Although the 'a' coefficients in equations (11) and (12) can be regarded as fundamental parameters characterizing a transistor, the quantities α_Nm α_I, I_{Eo}, I_{Co} are more easily measured directly, and it is often convenient to express the 'a' coefficients in terms of these latter quantities. This is very easily done.

Substituting (16) in (17)

$$I_{Eo} = - a_{11} - a_{21}\alpha_I$$

But, from (15), $- a_{21} = a_{11}\alpha_N$ and hence $I_{Eo} = - a_{11}(1 - \alpha_N\alpha_I)$

Thus, $$a_{11} = - \frac{I_{Eo}}{1 - \alpha_N\alpha_I} \qquad (19)$$

Similarly, $$a_{22} = - \frac{I_{Co}}{1 - \alpha_N\alpha_I} \qquad (20)$$

From (20) and (16) $$a_{12} = \frac{\alpha_I I_{Co}}{1 - \alpha_N\alpha_I} \qquad (21)$$

Similarly, $$a_{21} = \frac{\alpha_N I_{Eo}}{1 - \alpha_N\alpha_I} \qquad (22)$$

From (21) and (22) and the equality of a_{12} and a_{21} we derive the exceedingly important relationship:

$$\alpha_I I_{Co} = \alpha_N I_{Eo} \qquad (23)$$

4. EVALUATION OF THE 'SIMULTANEOUS' SATURATION CURRENTS, $I_{E\ sim}$, $I_{C\ sim}$

In almost all applications of a transistor as a switch, the emitter and collector terminals are the 'contacts' and the switching signal is applied between the base and one or other of these electrodes, usually either from a voltage generator in series with a resistance (so that, when the transistor is 'on', it is being in effect current controlled) or from a current generator with arrangements to limit the voltage excursion of the base in the 'off' state. When the transistor is 'off' both junctions are biassed in the reverse direction and drawing currents $I_{E\ sim}$ and $I_{C\ sim}$ respectively. Under <u>normal</u> operation of the transistor, that is with the switching signal applied between base and the normal emitter, $I_{C\ sim}$ represents an unwanted leakage current flowing to the external circuit. Under inverted operation, however, with the switching signal applied between base and normal collector, $I_{E\ sim}$ is the unwanted leakage current. We shall now evaluate these currents.

If V_E and V_C are both highly negative, equations (11) and (12) become:

$$I_E = I_{E\ sim} = - a_{11} - a_{12}\ .$$

$$I_C = I_{C\ sim} = - a_{21} - a_{22}\ .$$

Substituting from (19), (20), (21) and (22) we find:

$$I_{E\ sim} = \frac{I_{Eo} - \alpha_I I_{Co}}{1 - \alpha_N \alpha_I}$$

$$I_{C\ sim} = \frac{I_{Co} - \alpha_N I_{Eo}}{1 - \alpha_N \alpha_I}$$

From (23), we can finally write:

$$I_{E\ sim} = \frac{I_{Eo}(1 - \alpha_N)}{1 - \alpha_N \alpha_I} \qquad (24)$$

$$I_{C\ sim} = \frac{I_{Co}(1 - \alpha_I)}{1 - \alpha_N \alpha_I} \qquad (25)$$

The practical importance of equations (24) and (25) have been discussed by R.C. Bowes (Timing Circuits and Waveform Generators, p.4), and it is to be observed that operation of a transistor switch under <u>inverted</u>

conditions can often effect an enormous reduction in the flow of leakage current to the external circuit. From (23), (24) and (25) we find:

$$\frac{\text{Leakage current under inverted operation}}{\text{Leakage current under normal operation}} = \frac{I_{E\ sim}}{I_{C\ sim}} = \frac{\alpha_I(1 - \alpha_N)}{\alpha_N(1 - \alpha_I)}$$

Even though α_N and α_I are both close to unity, the ratio of $(1 - \alpha_N)$ to $(1 - \alpha_I)$ may be a very small quantity in a typical unsymmetrical transistor.

[Physically, it may seem a little surprising at first sight that, in a transistor in which I_{Eo} and I_{Co} are nearly equal, $I_{E\ sim}$ should be an order smaller than $I_{C\ sim}$. It must, however, be remembered that in a typical transistor of good quality, both I_{Eo} and I_{Co} are limited more by the total generation rate of minority carriers within the volume of the base region rather than by the respective junction areas since, provided that the effective value of the diffusion length L_p is sufficiently great, most of these thermally generated carriers will succeed in diffusing to and being collected at whichever junction is biassed negatively. However, if <u>both</u> junctions are reverse biassed, the one with the greater collecting area tends, since the supply of carriers is limited, to 'rob' the other junction.]

5. THE PERFORMANCE OF A TRANSISTOR SWITCH IN THE 'ON' STATE

In comparison with thermionic devices, which exhibit very low leakage current effects in the '<u>off</u>' state, a transistor switch in the '<u>on</u>' state has a remarkably good performance and can handle large currents with a very low voltage drop. Its bi-directional properties also enormously enhance its potentialities in the switching field, and lead to the possibility of remarkably simple and economical circuit techniques. We shall now evaluate the voltage-current relations in the 'on' state.

The 'a' coefficients have already been expressed in terms of α_N, α_I, I_{Eo} and I_{Co} by equations (19) to (22), and these expressions may be substituted in equations (11) and (12) and the resulting equations solved for the purpose of expressing the two exponential quantities in terms of I_E, I_C and the above parameters. If this is done we obtain:

$$(-\ I_{Eo}) \left[\exp \frac{qV_E}{kT} - 1 \right] = I_E + \alpha_I I_C \tag{26}$$

$$(-\ I_{Co}) \left[\exp \frac{qV_C}{kT} - 1 \right] = I_C + \alpha_N I_E \tag{27}$$

Equations (26) and (27) could be derived from intuitive reasoning, as follows. If the emitter be open-circuited, the collector current-voltage relationship would be that of a diode with reverse saturation current I_{Co} and hence be given by

$$I_C = (-\ I_{Co}) \left[\exp \frac{qV_C}{kT} - 1 \right] \qquad \text{(lecture 1, equation 30)}$$

However, with a finite emitter current I_E, the above collector current would be augmented by $-\ \alpha_N I_E$, leading immediately to equation (27). Equation (26) is derived by similar argument.

Equations (26) and (27) can now be applied for the purpose of determining the voltage drop across the switch, when in the 'on' state under defined current working, this voltage drop being given by the difference between V_C and V_E.

Figure 2 shows the circuits to be considered, (a) being for <u>normal</u> and (b) for inverted operation of the transistor. In these circuits,

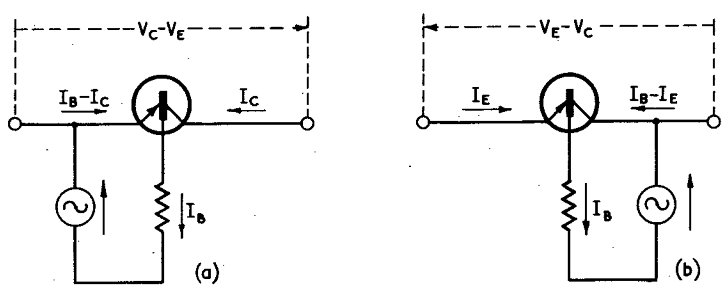


Figure 2

the resistance drawn in series with the base lead and the switching voltage generator should be regarded as including the base resistance r_{bb}' of the transistor itself, and the total resistance is presumed to be sufficiently great to ensure a constant input current when the transistor is on under all conditions of load current.

5.1. The Voltage Error at Zero Load Current

Applying equations (26) and (27) to the Fig. 2 (a) circuit, with $I_B - I_C$ written for I_E, we have:

$$V_E = \frac{kT}{q} \log \left[1 - \frac{I_B - I_C + \alpha_I I_C}{I_{Eo}} \right] \qquad (28)$$

$$V_C = \frac{kT}{q} \log \left[1 - \frac{I_C + \alpha_N (I_B - I_C)}{I_{Co}} \right] \qquad (29)$$

Hence $$V_C - V_E = \frac{kT}{q} \log \frac{I_{Eo}}{I_{Co}} \cdot \frac{I_{Co} - I_C - \alpha_N (I_B - I_C)}{I_{Eo} - I_B + I_C (1 - \alpha_1)} \qquad (30)$$

We now set $I_C = o$, and make the assumption that the drive current is sufficiently great to ensure that $\alpha_N I_B >> - I_{Co}$ and that $I_B >> - I_{Eo}$, and obtain:

$$V_C - V_E \simeq \frac{kT}{q} \log \frac{\alpha_N I_B I_{Eo}}{I_B I_{Co}}$$

Finally, since $\frac{I_{Eo}}{I_{Co}} = \frac{\alpha_I}{\alpha_N}$, we have:

$$V_C - V_E \simeq \frac{kT}{q} \log \alpha_I \qquad (31)$$

By a similar argument for the inverted circuit, Fig. 2 (b)

$$V_E - V_C \simeq \frac{kT}{q} \log \alpha_N \qquad (32)$$

If α_I and α_N are both close to unity, we have the approximations:

$$V_C - V_E \simeq -\frac{kT}{q}(1 - \alpha_I) \quad [\text{ Normal }] \qquad (33)$$

$$V_E - V_C \simeq -\frac{kT}{q}(1 - \alpha_N) \quad [\text{ Inverted }] \qquad (34)$$

The superiority of the inverted mode of operation for minimum voltage error at zero load current should be noted carefully. If we put $\alpha_N = 0.98$, $\alpha_I = 0.8$, the voltage error in the normal mode amounts to about -5 milli-volts, but with inverted operation is reduced to about -0.5 milli-volts.

It should also be noted that the error is, in theory, independent of the drive current I_B, provided that this is large compared with I_{Eo} and I_{Co}. However, at high values of drive current, voltage drops may occur due to the finite resistivities of the emitter and collector regions. These resistivities are neglected in the present discussion, but with grown junctions, the resistivity of the collector region is likely to be high and to cause considerable errors at high values of drive current. This and other effects are serious drawbacks to the use of normal kinds of grown junctions in switching circuits.

5.2. The Slope Resistance of the Switch at Small Load Currents

To determine this, we require to differentiate equation (30) with respect to I_C, and then equate I_C to zero. The expressions obtained are a little involved and the analysis will not be reproduced. The final results, however, are:

$$\left.\frac{d(V_C - V_E)}{dI_C}\right|_{I_C = 0} = \frac{1}{I_B}\cdot\frac{kT}{q}\cdot\frac{1 - \alpha_I\alpha_N}{\alpha_N} \quad [\text{ Normal }] \qquad (35)$$

$$\left.\frac{d(V_E - V_C)}{dI_E}\right|_{I_E = 0} = \frac{1}{I_B}\cdot\frac{kT}{q}\cdot\frac{1 - \alpha_I\alpha_N}{\alpha_I} \quad [\text{Inverted}] \qquad (36)$$

It will be observed that the slope resistance is inversely proportional to I_B and that there is not very much to choose between normal and inverted modes of operation, the normal mode having slightly lower slope resistance. For the transistor previously considered operating with a base current of 1 mA in the inverted mode, the value of slope resistance is about 7 ohms.

5.3 The Load Current Handling Capacity of the Switch

The collector and emitter junctions must both be forward biassed for the switch to remain on. Application of equations (26) and (27) shows that the relevant conditions to be fulfilled are that $I_E + \alpha_I I_C > 0$ and $I_C + \alpha_N I_E > 0$. First, consider the normal mode of operation of the switch with I_C positive and of sufficient magnitude just to reduce $I_E + \alpha_I I_C$ to zero. Since $I_E = I_B - I_C$, we can write :

$$I_B - I_C (1 - \alpha_I) = 0$$

Hence $$I_C = \frac{I_B}{1 - \alpha_I}$$

Next consider I_C negative and of sufficient magnitude just to reduce $I_C + \alpha_N I_E$ to zero. We can write :

$$I_C + \alpha_N(I_B - I_C) = 0$$

Hence $$I_C = - \frac{\alpha_N I_B}{1 - \alpha_N}$$

In the normal mode, therefore, the switch can operate between load current limits $-\alpha_N I_B/(1 - \alpha_N)$ to $+I_B/(1 - \alpha_I)$. The corresponding limits in the inverted mode are by symmetry, $-\alpha_I I_B/(1 - \alpha_I)$ to $+I_B/(1 - \alpha_N)$.

5.4 A Basic Equivalent Circuit for Steady State Conditions

By means of equations (26) and (27) we have been able to determine the performance of a transistor switch under conditions of defined load and drive currents. Although the analysis was relatively straightforward and convenient for the purpose required, we shall now adopt a rather different approach with a view to deducing a most useful equivalent circuit, which may be regarded as of fundamental importance for obtaining a real insight into the steady state behaviour of a transistor in the on regime.

Equations (11) and (12) may be re-written, substituting from equations (19) to (22) for the 'a' coefficients. If this be done, we obtain:-

$$I_E = -\frac{I_{Eo}}{1-\alpha_N\alpha_I}\left[\exp\left(\frac{qV_E}{kT}\right)-1\right] + \frac{\alpha_I I_{Co}}{1-\alpha_N\alpha_I}\left[\exp\left(\frac{qV_C}{kT}\right)-1\right] \quad (37)$$

$$I_C = -\frac{I_{Co}}{1-\alpha_N\alpha_I}\left[\exp\left(\frac{qV_C}{kT}\right)-1\right] + \frac{\alpha_N I_{Eo}}{1-\alpha_N\alpha_I}\left[\exp\left(\frac{qV_E}{kT}\right)-1\right] \quad (38)$$

Furthermore, by writing I_{Ei} for $-\frac{I_{Eo}}{1-\alpha_N\alpha_I}\left[\exp\left(\frac{qV_E}{kT}\right)-1\right]$

and I_{Ci} for $-\frac{I_{Co}}{1-\alpha_N\alpha_I}\left[\exp\left(\frac{qV_C}{kT}\right)-1\right]$, we have:-

$$I_E = I_{Ei} - \alpha_I I_{Ci} \quad (39)$$

$$I_C = I_{Ci} - \alpha_N I_{Ei} \quad (40)$$

I_{Ei} should be regarded as the hole current injected at the emitter junction, and I_{Ci} as the hole current injected at the collector junction. $\alpha_I I_{Ci}$ is to be regarded as the hole current collected at the emitter as a result of the injection of I_{Ci} at the collector and, similarly,

$\alpha_N I_{Ei}$ is to be regarded as the hole current collected at the collector as a result of I_{Ei} injected at the emitter. Equation (39) states that the external current flowing to the emitter is the difference between the current injected, I_{Ei}, and the current collected, $\alpha_I I_{Ci}$. Equation (40) is similarly interpreted. We thus arrive at the equivalent circuit shown in Fig. 3, in which the relationships previously deduced between I_{Ei} and V_E and between I_{Ci} and V_C are represented by the two ideal diodes, having reverse saturation currents $I_{Eo} / 1 - \alpha_N \alpha_I$ and $I_{Co} / 1 - \alpha_N \alpha_I$, respectively.

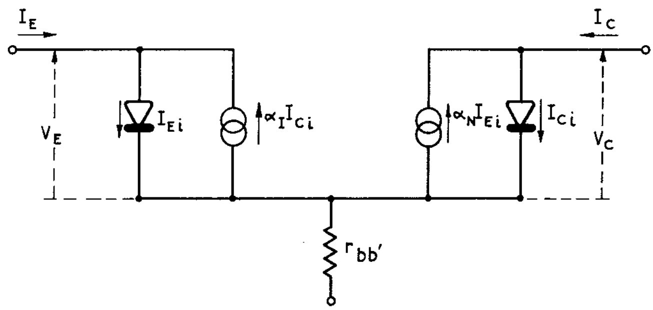


Figure 3

I_{Ei} and I_{Ci} can, of course, be expressed in terms of I_E and I_C by manipulation of (39) and (40) yielding:-

$$I_{Ei} = \frac{I_E + \alpha_I I_C}{1 - \alpha_N \alpha_I} \tag{41}$$

$$I_{Ci} = \frac{I_C + \alpha_N I_E}{1 - \alpha_N \alpha_I} \tag{42}$$

We are now in a position to take note of the important differences in the voltage-current relationships of the transistor emitter and collector circuits between normal operation without collector bottoming (active state) and bottomed operation (on state).

First, consider a transistor with 1 mA flowing in the emitter, and with a negative bias of several volts on the collector. If we neglect I_{Co}, we can say that $I_C = -\alpha_N I_E$, and equations (41) and (42) yield $I_{Ei} = I_E$ and $I_{Ci} = o$. We also find, if r_{bb}' be neglected, that the small signal emitter input impedance is:

$$\frac{dV_E}{dI_E} = \frac{dV_E}{dI_{Ei}} = \frac{kT}{q} \times \frac{1}{I_E} \quad \Omega \quad 25 \text{ ohms}$$

Next, consider the same transistor, but with 1 mA flowing to both emitter and collector. We now find:

$$I_{Ei} = \frac{1 + \alpha_I}{1 - \alpha_N \alpha_I} \text{ mA}, \quad I_{Ci} = \frac{1 + \alpha_N}{1 - \alpha_N \alpha_I} \text{ mA}$$

Both I_{Ei} and I_{Ci} are much greater than I_E and I_C, and the voltages across the two junctions will adjust themselves to values appropriate to these injected currents. A most important consequence of the injected emitter current being much greater than the actual external emitter current now follows. Suppose we hold the collector voltage constant so that I_{Ci} remains constant at the value noted above, but vary V_E by a small signal. The effective emitter input impedance is now:

$$\left. \frac{dV_E}{dI_E} = \frac{dV_E}{dI_{Ei}} \times \frac{dI_{Ei}}{dI_E} \right|_{I_{Ci} = \text{const}}$$

From equation (39), $\frac{dI_{Ei}}{dI_E} = 1$, and hence:

$$\frac{dV_E}{dI_E} = \frac{dV_E}{dI_{Ei}} = \frac{kT}{q} \times \frac{1}{I_{Ei}} = \frac{kT}{q} \times \frac{1}{I_E} \times \frac{1 - \alpha_N \alpha_I}{1 + \alpha_I}$$

It will be observed that the value of small signal input impedance is lower than the value obtained for the previous case by a factor $(1 - \alpha_N \alpha_I)/(1 + \alpha_I)$. Physically, this can be thought of as being due to the emitter diode being forward biassed to a much higher value of D.C. current than I_E, and therefore producing a much greater incremental injected current for a given voltage increment than a diode biassed to a current I_E.

At this stage reference may be made to equations (35) and (36) for the incremental slope resistance of a transistor switch. The remarkably low value of this resistance will now be seen to be due to effects of the type just discussed.

5.5 'Hole Storage' in a Transistor in the On State

Reference may now be made to Fig. 4, which shows very roughly how the hole density across the base region of a transistor may be expected to vary according to conditions of operation. All three diagrams are for a fixed value of I_E, (a) being for operation in the active state with no collector injection and therefore zero hole density at the collector, (b) for operation in the on state with a small amount of collector injection, and (c) for nearly equal amounts of emitter and collector injection.

The first diagram (a) requires virtually no explanation. As there is no injection from the collector, the total hole density, p_{Ti}, in the base region is that produced by injection from the emitter, p_{Ei}, and we can write $p_{Ti} = p_{Ei}$. The total external emitter and collector currents are proportional to the slopes of the p_{Ei} curve at E and C respectively. If $\alpha_N = 1$, the curve is a straight line with equal slopes at E and C*. If, however, $\alpha_N < 1$, the slope of the curve at C will be less than that at E. In the second diagram (b), the transistor is in the on state and a small amount of injection is taking place from the collector indicated by the curve marked p_{Ci}. It will be noted that the p_{Ci} curve has been drawn with considerably less slope at the emitter than at the collector, indicating that α_I

---

*Strictly speaking, of course, 'one-dimensional' carrier flow is implied if this statement is true, but the present discussion is only intended to be qualitative.

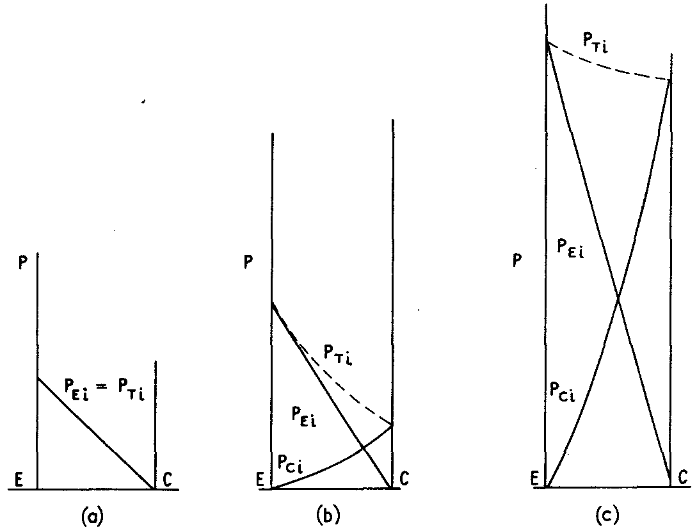


Figure 4

is substantially less than unity. The curve, p_{Ti}, for the total hole injection level is obtained by the addition of the two curves p_{Ei} and p_{Ci}. Again, the slopes of the p_{Ei} curve at E and of the p_{Ci} curve at C are proportional to the injected emitter and injected collector currents respectively, whilst the slopes of the p_{Ti} curve at E and C are proportional to the external emitter and external collector currents respectively. The third diagram (c) relates to much heavier injection levels, but the argument is similar. It will be noticed that all the three diagrams have been drawn with p_{Ti} curves of similar slopes at E, indicating that the external emitter currents for the three situations are equal. The slopes at C of these p_{Ti} curves, however, are all different, that in diagram (b) being less than that in diagram (a), and that in diagram (c) being still less, indicating progressively lower values of external collector current (measured away from the collector). We could have gone on to draw more diagrams to cover situations in which the slopes of the p_{Ti} curves at E and C were of opposite sign, indicating that the external collector current was, in such cases, flowing towards the collector, and we would find that still higher levels of p_{Ei} and p_{Ci} would be required to satisfy such conditions. The point that should be appreciated by a study of such diagrams is that if we specify the slope of the p_{Ti} curve at E and thereby fix the value of external emitter current, the levels of the individual p_{Ei} and p_{Ci} curves will depend upon the algebraic difference between the slopes of the p_{Ti} curve at C and E, which in turn depends upon the value of the external collector current. In general, the greater this algebraic difference, the higher will be the levels of the p_{Ei} and p_{Ci} curves necessary to satisfy the condition that the p_{Ti} curve should be the sum of the p_{Ei} and p_{Ci} curves.

Although the above discussion and diagrams are really no more than an illustration of the requirements to be met for a correct solution of equations (41) and (42) they afford a useful insight into the steady state conditions existing in the base region of a transistor switch. In particular, it is to be noted that the area under the p_{Ti} curve represents the total storage of holes in the base per unit cross section. Before a transistor switch can be turned off fully, these excess holes must be withdrawn by a suitable flow of external reverse current to the emitter or collector or be left for a sufficiently long time to enable them to recombine with the excess electrons which have flowed in from the base connection. We shall examine the implications of this requirement in section (6.1) below.

It is also to be noted that even when a transistor switch has been turned on and the external currents have become steady, a finite time is required for the hole density to build up to its steady state so that equations (41) and (42) are satisfied. It is instructive to consider briefly this situation for a 'perfect' transistor in which $\alpha_I = \alpha_N = 1$. The curves for p_{Ei} and p_{Ci} will each be straight lines for such a transistor, and it will be impossible for these quantities to reach sufficiently high levels to satisfy any differences in the slopes of p_{Ti} at E and C set by external current demands. When the transistor is turned on, therefore, the hole density in the base will build up without limit and steady state conditions will never be reached! In consequence, the longer the switch is left on, the longer it will take to turn it off! Again, this result can be obtained by consideration of equations (41) and (42) which predict infinite values of I_{Ei} and I_{Ci} unless $I_C = - I_E$.

There is yet another way of looking at the situation governing hole storage. The nett number of holes entering the base region per unit time is proportional to the algebraic sum of I_E and I_C.* When equilibrium in the on state has been reached, the nett number of holes entering per unit time must equal the number lost per unit time by recombination. We can regard the recombination rate as being proportional to hole density and inversely proportional to the effective lifetime of the injected holes in the base region. For a given rate of recombination, the hole density must therefore be proportional to the effective lifetime, which must tend to infinity as α_N and α_I each tend to unity.

We will now proceed to a highly simplified and approximate treatment of the transient behaviour of a transistor switch.

6. THE TRANSIENT BEHAVIOUR OF A TRANSISTOR SWITCH

The circuit which we shall analyse is shown in Fig. 5. It will be seen that it is essentially a common-emitter driven circuit and that the load current flow during the on state is into the emitter and out of the collector (where speed of operation is a prime requisite, inverted operation of the transistor is not a good thing). The transistor is therefore operating in the manner in which it probably finds most application in the field of fast pulse circuitry.

The rise-time transient is relatively easy to evaluate and it is considered to have been completed when, in response to an input current I_{B1}, the collector has fully bottomed and the transistor is in the on state. During the rise-time period, the emitter diode is forward biassed and the collector diode reverse biassed. After rise-time has been completed, a finite time must elapse before I_{Ei}, I_{Ci} and the carrier density in the base reach equilibrium values, as discussed in section (5.5). Turn-on of the transistor can be said to be completed when these equilibrium conditions have been reached. We shall not calculate this extra period of time necessary to complete the turn-on process. In most applications it will be found that it has been virtually completed before turn-off is required, but for the purpose of achieving the minimum possible period of hole storage during turn-off, it is a good thing if it has not, in fact, been completed. During the on time of the transistor we assume that changes in the voltage drops occurring across the emitter and collector diodes are insufficiently great to alter the external currents significantly.

---

*This assumes that current flow in the base lead is by electrons.

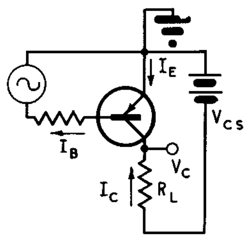


Figure 5

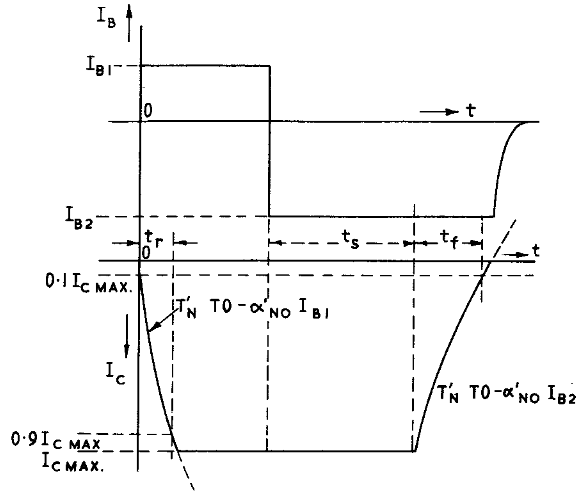


Figure 6

The turn-off process is initiated by reversing the polarity of the base input current and changing it to a value I_{B2}. The required period for turn-off is divided into two parts, the hole-storage time and the fall-time transient. During the hole storage time, the two diodes remain forward biassed because the hole density in the base has nowhere fallen to zero. Holes, however, are being extracted at the collector and also, if $|I_{B2}| > |I_{C\ max}|$, at the emitter. Provided that the reverse base current drive is not so great that the reverse emitter current is considerably greater than the collector current, the hole density at the collector will reach zero before that at the emitter. The collector diode will then 'unhitch' and the fall-time transient start. Even during this transient period in which $\alpha_I I_{Ci}$ (the current 'collected' at the emitter) is zero and the emitter current is negative, the emitter diode remains forward biassed on account of charge storage in the base (emitter diffusion capacitance, lecture 1, section 6.2). The fall-time transient is relatively easy to compute if the assumption is made that $\alpha_I I_{Ci}$ has effectively reached zero just before the fall in collector current starts. Finally, the emitter diode will 'dry up' and the reverse base current will cease, the base voltage rapidly rising to that of the drive generator. The transistor will now be fully off.

6.1 Calculation of Transient Response

Perhaps the most important matters in the following discussion are the frequency variations of α_N and α_I. We shall assume that the following operational expressions are sufficiently accurate for these parameters:

$$\alpha_N(p) = \frac{\alpha_{No}}{1 + \frac{p}{\omega_N}} \qquad (43)$$

$$\alpha_I(p) = \frac{\alpha_{Io}}{1 + \frac{p}{\omega_I}} \qquad (44)$$

In the above expression, α_{No} and α_{Io} are the D.C. and low-frequency values of α_N and α_I, p is the differential operator, d/dt, and ω_N and ω_I are the angular cut-off frequencies of α_N and α_I respectively. As we shall use linear operational circuit analysis for our calculations, it is necessary to assume that α_N and α_I are not dependent on current levels - in practice, this assumption is not fully justified, but suitable average values based on measurements with a collector current of about $-V_{CS}/2R_L$ will generally prove satisfactory.

I_{Eo} and I_{Co} will be completely neglected, and it will also be assumed that the time constant of the collector barrier capacitance and collector load resistance is small compared with $1/\omega_N$, so that the effect of this capacitance in slowing up the rise-time and fall-time transients may be neglected.

6.1.1 Rise-time Transient

We assume that, at t = o, the base current, I_B is switched from zero to a positive value, I_{B1} (remember that we consider the positive direction of base current as being out of the base so that we can write $I_B = I_E + I_C$). We can write:

$$I_E = I_B - I_C = - \frac{I_C}{\alpha_N}$$

Hence,
$$I_C = - I_B \frac{\alpha_N}{1 - \alpha_N} ,$$

and
$$I_C(p) = - I_B(p) \frac{\alpha_{No}}{1 - \alpha_{No}} \quad \frac{1}{1 + \frac{p}{\omega_N(1-\alpha_{No})}}$$

For a step function of input current, I_{B1}, the solution for I_C is:

$$I_C = - I_{B1} \frac{\alpha_{No}}{1 - \alpha_{No}} \left\{ 1 - \exp\left(- \frac{t}{T'_N}\right) \right\} \quad \text{where } T'_N = \frac{1}{\omega_N(1-\alpha_{No})}$$

The collector current waveform will therefore set-off towards an asymptotic value $-I_{B1}\alpha'_{No}$, where $\alpha'_{No} = \alpha_{No}/(1-\alpha_{No})$, with a time constant T'_N. However, the transistor will bottom when $I_C = -V_{CS}/R_L$, and hence, provided that $I_{B1}\alpha'_{No} > V_{CS}/R_L$, the exponential build up of current will not be completed.

For most practical purposes, we can consider the rise-time transient to be completed by the time the current has reached 90% of its bottoming value. Hence, putting $I_{C\,max}$ for $-V_{CS}/R_L$, we find

$$t_r = T'_N \log \frac{1}{1 - \left| \frac{0.9 I_{C\,max}}{\alpha'_{No} I_{B1}} \right|}$$

6.1.2 Hole Storage Time

We assume that the transistor has been on for a sufficiently long time to permit equilibrium conditions to have been established in the base region before the base input current is reversed for the purpose of initiating the turn-off process. Both the emitter and collector diodes are forward biassed and since $I_C = I_{C\ max}$ and $I_E = I_{B1} - I_{C\ max}$, equations (41) and (42) give, for the values of I_{Ei} and I_{Ci}:

$$(I_{Ei})_1 = \frac{I_{B1} - I_{C\ max}(1 - \alpha_{Io})}{1 - \alpha_{No}\alpha_{Io}} \qquad (46)$$

$$(I_{Ci})_1 = \frac{I_{C\ max}(1 - \alpha_{No}) + \alpha_{No} I_{B1}}{1 - \alpha_{No}\alpha_{Io}} \qquad (47)$$

The base current is now suddenly changed to the negative value I_{B2}. The incremental change in base current is $\Delta I_B = I_{B2} - I_{B1}$ and since the collector current cannot change whilst the transistor remains bottomed, an incremental change equal to ΔI_B appears in the emitter current. By combining equation (38) with the operational equations (43) and (44) for α_N and α_I, we can determine the incremental change produced in I_{Ci}, and we find:

$$\Delta I_{Ci} = \Delta I_B \frac{\alpha_{No}\ (1 + pT_I)}{(1 + pT_N)\ (1 + pT_I) - \alpha_{No}\alpha_{Io}} \qquad (48)$$

where we have written T_N for $1/\omega_N$ and T_I for $1/\omega_I$. Bearing in mind that ΔI_B is a step-function and that the denominator in (48) is one degree higher in p than the numerator, it is easily seen that no discontinuity in ΔI_{Ci} can occur at the moment of application of ΔI_B. Furthermore, except in cases where the magnitude of I_{B2} is very large (say, more than twice $I_{C\ max}$), we are not particularly interested in the value of I_{Ci} at times very close to the initiation of turn-off since the object of the present analysis is to determine how long I_{Ci} takes to fall to zero. We will therefore simplify (48) by eliminating the term in p in the numerator and the term in p^2 in the denominator. We thus write:

$$\Delta I_{Ci} \;\Omega\; \Delta I_B \frac{\alpha_{No}}{1 - \alpha_{No}\alpha_{Io} + p(T_N + T_I)} \tag{49}$$

If t be measured from the time of initiation of turn-off, we find:

$$\Delta I_{Ci}(t) = \frac{\Delta I_B \, \alpha_{No}}{1 - \alpha_{No}\alpha_{Io}} \left[1 - \exp\left(- \frac{t}{T_S} \right) \right] \tag{50}$$,

where $T_S = \dfrac{T_N + T_I}{1 - \alpha_{No}\alpha_{Io}}$

Now, in order to determine the time at which I_{Ci} will have fallen to zero and the collector diode will have reached reverse bias conditions, we must equate $-\Delta I_{Ci}(t)$ to $(I_{Ci})_1$, the value of I_{Ci} at the start of turn-off, and solve for t. This gives us:

$$t_S = \frac{T_N + T_I}{1 - \alpha_{No}\alpha_{Io}} \log \frac{I_{B1} - I_{B2}}{-\dfrac{I_{Cmax}}{\alpha'_{No}} - I_{B2}} \text{, where } \alpha'_{No} = \frac{\alpha_{No}}{1 - \alpha_{No}} \tag{51}$$

It should be noted that the above expression for t_S will, in fact, hold even if I_{B2} is not negative, provided that it is insufficiently positive to hold the transistor bottomed ad infinitum. The criterion is that $-I_{C\,max}/\alpha'_{No}$ shall be greater than I_{B2}. However, to reduce t_S as much as possible, I_{B2} should be negative and of large magnitude. However, a limiting condition may be reached, if I_{B2} is of excessive magnitude, in which the emitter diode becomes reverse biassed during the hole storage decay time before the collector diode has reversed. This is on account of an excessively high value of reverse external emitter current. If this happens, equation (51) ceases to be valid, and a somewhat complicated situation ensues. Provided, however, that $|I_{B2}| < 2|I_{C\,max}|$, this contingency is remote.

Generally speaking, with an unsymmetrical type of transistor structure as shown in Fig. 1, $\omega_I << \omega_N$ and hence $T_I >> T_N$. This is because of the excessively large area of the collector, which, when it is acting as an emitter floods a much bigger volume of the base region with holes than does the normal emitter. An alternative way of expressing this is to say that the diffusion capacitance of the collector diode when forward biassed is much greater than that of the emitter diode during normal operation.

6.1.3 Fall-time Transient

At the end of the hole storage time, the carrier density at the collector will have reached zero, the collector diode will have become reverse biassed, and the collector current will start to decay. In the simplified treatment which follows, it will be assumed that the transition between the hole storage regime and the fall-time transient regime is a sudden one, but this is not really so in practice and a more gradual transition between the two regimes actually occurs, this leading to a slight rounding off of the front edge of the fall-time transient. The reason for this is that even after I_{Ci} has become zero, a short time must elapse before the emitter has succeeded in collecting all the holes previously injected into the base by the collector diode, whilst it was forward biassed. The smaller the value of T_I, the shorter will be the time required for this operation to be completed. The effect, however, is a relatively slight one, and we shall neglect it.

As remarked in the last paragraph of section 6, the emitter diode remains forward biassed during the fall-time transient (at least, we shall assume that it does, though excessively fast turn-off demands may cause it to 'dry up') even though I_E is negative and $\alpha_I I_{Ci}$ is zero, the reverse emitter current being supplied by the discharge of the emitter diffusion capacitance.

At the beginning of the transient, the collector current is still $I_{C\,max}$, and we can say that if, at this moment, the base current had suddenly been changed from I_{B2} to $-I_{C\,max}/\alpha'_{No}$, the collector current would have remained at $I_{C\,max}$, and the transistor would have been in equilibrium. Thus, the fact that the base current input is in reality I_{B2} is equivalent to a base current increment, $I_{B2} + I_{C\ max}/\alpha'_{No}$. By an analogy with the argument used for calculating the rise-time transient, we can now calculate the incremental change in collector current, and we find:

$$\Delta I_C = -(\alpha'_{No} I_{B2} + I_{C\ max}) \left\{ 1 - \exp\left(-\frac{t}{T'_N}\right) \right\}$$

The actual value of collector current can be obtained by adding to the increment the initial value $I_{C\ max}$. If this is done, we see that the collector current waveform sets-off towards an asymptotic value $-\alpha'_{No} I_{B2}$ with a time constant T'_N, but must obviously terminate at zero current. For most practical purposes, we can regard the fall-time transient as having been completed by the time I_C has reached a value of 10% of $I_{C\ max}$. The time required for this is:

$$t_f = T'_N \log \frac{\alpha'_{No} I_{B2} + I_{C\ max}}{\alpha'_{No} I_{B2} + 0.1 I_{C\ max}} \tag{52}$$

Provided that I_{B2} is negative, which will generally be true, we can write:

$$t_f = T'_N \log \left| \frac{\alpha'_{No} I_{B2} \ | + | \ I_{C\ max}}{\alpha'_{No} I_{B2} \ | + | \ 0.1 I_{C\ max}} \right| \tag{53}$$

The transient behaviour of the collector current waveform is summarised in Fig. 6.

6.1.4 A Simple Example of Transient Response Calculation

For a Mullard OC44 transistor, reasonable values of the important parameters might be:

$$\alpha_{No} = 0.98$$

$$\alpha_{Io} = 0.9$$

$$\omega_N = 31 \times 10^6 \text{ radians/sec.}$$

$$\omega_I = 11 \times 10^6 \text{ radians/sec.}$$

Suppose we set $I_{C\ max} = -5$ mA, $I_{B1} = 1$ mA, $I_{B2} = -5$ mA

Then $$t_r = \frac{1}{31 \times .02} \log \frac{1}{1 - \frac{4.5}{49}} = 0.15\ \mu \text{ sec.}$$

$$t_s = \frac{\frac{1}{31} + \frac{1}{11}}{1 - 0.88} \log \frac{1 + 5}{\frac{5}{49} + 5} = 0.17\ \mu \text{ sec.}$$

$$t_f = \frac{1}{31 \times .02} \log \frac{245 + 5}{245 + 0.5} = 0.03\ \mu \text{ sec.}$$

LECTURE 6. THE TRANSISTOR IN HIGH SPEED SWITCHING CIRCUITS*

by G. Ord

1. Introduction

A thorough exposition of the theoretical aspects of the alloy junction transistor as a switch was given in the previous lecture. It is the aim of this lecture to discuss some of the practical applications of that theory with particular emphasis on circuits working at high speeds. The approach will be empirical rather than quantitative as at high speeds too many factors come into a thorough calculation to make it amenable to complete analysis. However, it is hoped to indicate the main principles involved, and an estimate will be given of how nearly the circuit performance agrees with the basic calculation.

At the present time high speed switching circuits are used to their greatest extent in digital computers and therefore it is from that field that examples will be drawn. As there has been no lecture on computing previous to this one, it will be necessary to say a little about basic computer circuits, but the main emphasis will be on their high speed aspects. Many of the principles used in the circuits are applicable not only in digital computers, but also in other high speed circuit fields.

---

*Since this lecture was prepared a new approach to the subject of transistors in circuits (particularly switching circuits) has been suggested by R. Beaufoy and J.J. Sparkes. See:

1. Beaufoy R. and Sparkes J.J. "The Junction Transistor as a Charge-Controlled Device". A.T.E. Journal Oct. 1957. 13 No.4 p.310.

2. Beaufoy R. "The Transistor Switching Circuit Design Using the Charge-Control Parameters". Proc. I.E.E. May 1959. 106B p.1085.

3. Sparkes J.J. "The Measurement of Transistor Transient Switching Parameters". May 59. 106B p.562.

It is hoped that these papers will be the subject of a later article in the R.R.E. Journal.

It is proposed to deal only with circuits using p.n.p. transistors and thus the whole range of circuits using p.n.p. and n.p.n. transistors in association will not be considered. This is a branch of circuitry which has not been investigated a great deal in this country as it is only recently that n.p.n. transistors with reasonable high frequency properties have become available.

The subjects which will be considered are as follows:-

2. There will be some discussion of the circuit shown in Fig.1, which is the one considered by Ebers and Moll, and some numerical examples will be given of the rise, hole storage and fall times using the Ebers and Moll equations.

3. Methods of reducing the hole storage time.

4. Bi-stable circuits, in particular the transistor equivalent of the Eccles Jordan circuit;

5. A short consideration of emitter followers as buffer stages.

6. Methods of setting the state of the bi-stable circuit: modifications needed to make the circuit act as a scalar; ways of interconnecting several bi-stable circuits so that they will act as shifting registers: the simplifications which can be introduced into these circuits if the transistor is used as a bi-directional switch.

7. Direct coupled circuits.

8. Logic circuits. Only AND and OR circuits will be considered and only one example of each.

2. Rise, Fall and Hole Storage Times

The circuit shown in Fig.1, which is the one dealt with by Ebers and Moll, is the circuit which is basic to all switching applications. It may be considered as (a) a gate which allows current to flow through it (from emitter to collector) when base current is provided.

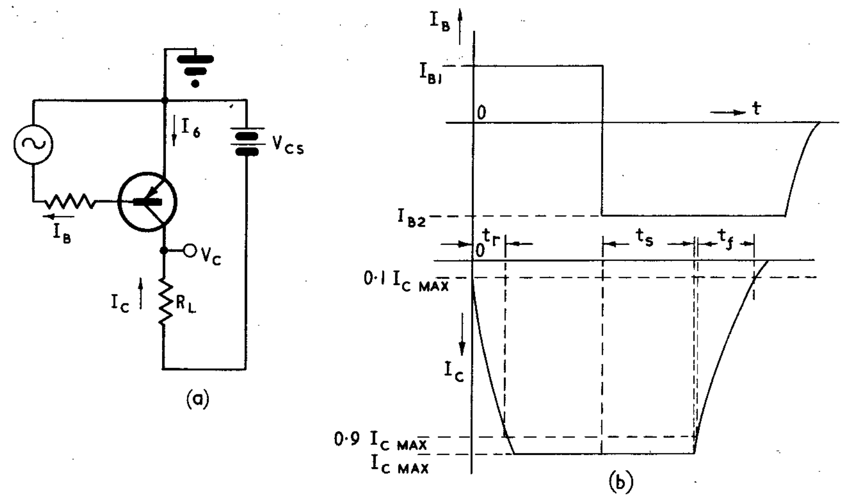


Figure 1

(b) a switch, which is the limiting case of (a) when the resistance between emitter and collector becomes very small. As has been shown, this switch can be a bi-directional one, allowing current to flow in either direction. Under these conditions the transistor can be considered to be the analogue of a relay, the emitter and collector being the relay contacts and the base current the current to the relay coil.

(c) an amplifier and also as a restandardising circuit if the collector is allowed to bottom.

(d) an inverter.

2A. Rise and Fall Times

The examples given in the previous lecture showed that by driving a transistor with a large forward and reverse base current, it is possible to achieve short rise and fall times using transistors of limited frequency response. In many computer circuits it is not possible to provide drive currents as great as was suggested because the output current of one transistor may be required to drive the inputs of several other transistors. This suggests the use of transformers to step up the collector currents available but this presents problems, though not insoluble ones, at high speeds when non-repetitive waveforms are used. In the subsequent discussion only non-transformer type circuits will be described.

The formula for the rise time was given as

$$t_r = T'_N \log \frac{1}{1 - \left| \dfrac{0.9\ I_{cmax}}{\alpha'_{NO}\ I_{B1}} \right|}$$

This is Equation 45 of the previous lecture.

When $\dfrac{I_{cmax}}{\alpha'_{NO} I_{B1}}$ is very small $t \simeq T'_N \left| \dfrac{0.9\ I_{cmax}}{\alpha'_{NO}\ I_{B1}} \right|$

$$\simeq T_N \left| \frac{I_{cmax}}{I_{B1}} \right|$$

The formula for the fall time, equation (53) of the previous lecture, can have similar approximations made to it and becomes

$$t_f \simeq T_N \left| \frac{I_{cmax}}{I_{B2}} \right|$$

For an OC 44 having the parameters

$$\alpha_{NO} = 0.98$$

$$\omega_N = 31 \times 10^6 \text{ radians/sec}$$

Typical currents which might be used in a computer application are

$$I_{B1} = - I_{B2} = 1\text{mA} \; ; \quad I_c = 5\text{mA}$$

thus $t_r \simeq$ 160 millimicroseconds

$t_f \simeq$ 160 millimicroseconds.

For an SB 240 surface barrier transistor

$$\alpha_{NO} = 0.96$$

$$\omega_N = 300 \times 10^6 \text{ radians/sec.}$$

$$t_r \simeq 15 \text{ millimicroseconds}$$

$$t_f \simeq 15 \text{ millimicroseconds}$$

In practice the measured values of rise and fall time, particularly for surface barrier transistors, are usually greater than this by anything up to 100%. This may be due to the fact that the conditions under which the Ebers and Moll equation hold, do not apply; in particular the current levels are rather high. Another factor which may account for the error is a large value of the product $C_c R_L$. At these high signal levels any simple equivalent circuit is a gross approximation to the truth, as most of the parameters in it are varying as the voltage and current levels vary. Bearing in mind these limitations when using the equivalent circuit of Fig.2, which is based on Fig.11 on p.68 of R.R.E. Journal, April 1960, the time constant $\frac{r_\epsilon}{1-\alpha_o} \times C_{DE}$ is the time constant T_N' in the Ebers and Moll expression for rise and fall times. The effect of Miller capacity due to C_c and finite R_L will increase this time constant. This may account for the greater rise and fall times.

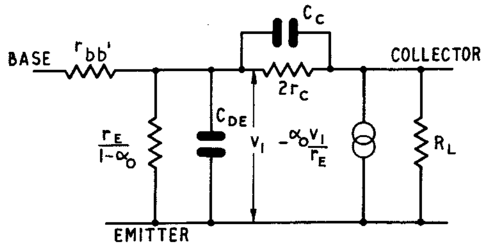


Figure 2

2B. Hole Storage Time

The formula for the hole storage time is

$$t_s = \frac{T_N + T_I}{1 - \alpha_{NO}\,\alpha_{IO}} \log \frac{I_{B1} - I_{B2}}{-\dfrac{I_{cmax}}{\alpha'_{NO}} - I_{B2}} \quad \text{where } \alpha'_{NO} = \frac{\alpha_{NO}}{1 - \alpha_{NO}}$$

When $T_I >> T_N$ (i.e. $\omega_I << \omega_N$) and $\alpha_{NO}\,\alpha_{IO} \rightarrow \alpha_{IO}$.

this formula simplifies to

$$t_s \simeq \alpha'_{IO}\; T_I \log \frac{I_{B1} - I_{B2}}{-\dfrac{I_{cmax}}{\alpha'_{NO}} - I_{B2}}$$

and with the sweeping assumption that $\left|\dfrac{I_{cmax}}{\alpha'_{NO}}\right| << |\; I_{B2}\;|$ this formula simplifies further to

$$t_s \simeq \alpha'_{IO}\; T_I \log \frac{I_{B1} - I_{B2}}{- I_{B2}}$$

The significance of this is very obvious. $|I_{B2}|$ should be large relative to $|I_{B1}|$ to achieve low hole storage time. For an OC 44 with

$$\alpha_{NO} = 0.98$$
$$\alpha_{IO} = 0.9$$
$$\omega_N = 31 \times 10^6 \text{ radians/sec.}$$
$$\omega_I = 11 \times 10^6 \text{ radians/sec.}$$

and $I_{cmax} = -5$ mA $I_{B1} = 1$mA and $I_{B2} = -1$mA the exact formula gives $t_s = 0.61$ µsec

the first approximation gives $t_s = 0.54$ µsec

and the second approximation gives $t_s = 0.63$ µsec

For a surface barrier transistor SB.240, typical transistor parameters are $\omega_N \simeq 300 \times 10^6$ radians/sec.

$$\omega_I \simeq 90 \times 10^6 \text{ radians/sec.}$$

$$\alpha_{NO} = 0.96$$
$$\alpha_{IO} = 0.92$$

For $I_c = -5$mA $I_{B1} = 1$mA and $I_{B2} = -1$mA.

t_r (Ebers and Moll) = 64 millimicroseconds.

t_r (first approx.) = 64 "

t_r (second approx.) = 86 "

However, in practice, these figures are more likely to be 20-30 millimicroseconds. This departure from calculated figures is thought to be due to the fact that one of the basic assumptions on which Ebers and Moll based their theory has been violated; with surface barrier transistors the emitter injection efficiency is low and is not near unity as Ebers and Moll assumed. However, the surface barrier transistor, although having a high α cut off frequency and a hole storage time disproportionately low considering its transistor parameters, has some disadvantages which will be shown when examples are given - in particular a low collector impedance and rapid fall off of α' with collector current.

3. The Reduction of the Hole Storage Time

The hole storage time t_r occurs when a transistor is allowed to bottom. If the primary interest of the designer is to achieve maximum speed of operation, then bottoming must be prevented. But the transistor when bottomed has some very interesting properties which result in more economic circuits and thus in design it is necessary to weigh speed against economy. In this section therefore, methods of reducing the hole storage time, as well as preventing it altogether, will be considered.

To achieve a short rise time a transistor is driven hard i.e. a large I_{B1} is used. This current need only be provided during the rise time period. Similarly $|I_{B2}|$ needs to be large, preferably larger than $|I_{B1}|$, but need only be available during the hole storage and fall times. The circuit of Fig.3 thus suggests itself. The values of R_I and R_L are chosen such that $\frac{R_I V_{cs}}{R_L V_{B1}}$ is just less than α' i.e. the steady current fed to the base is just sufficient to keep the transistor turned on. When driving the transistor hard on at the base, the input condenser can be considered to be feeding a low impedance (the input impedance is a variable quantity depending on the current drive).

When switching off the transistor, initially a large current flows, and the input impedance appears small until the transistor is nearly turned off; it then becomes large. If the value of C has been chosen to deal with a transistor storing many holes, then in a similar circuit with a transistor storing few holes, the holes will soon be removed and the voltage at the base will jump to nearly $+(V_{B1} + V_{B2})$ and will only approach $+V_{B2}$ on a time constant of CR_I which may be large compared with the time interval at which it is wanted to apply input pulses. This undesirable partial paralysis time can be reduced by putting a diode between the base and a voltage source of $+V_{B2}$. Fig.4.

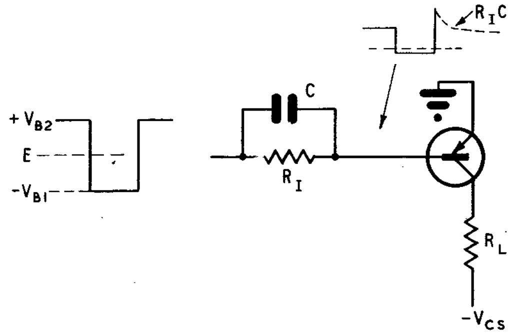


Figure 3

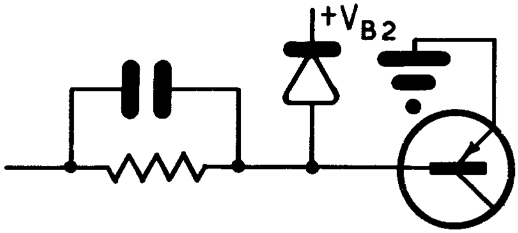


Figure 4

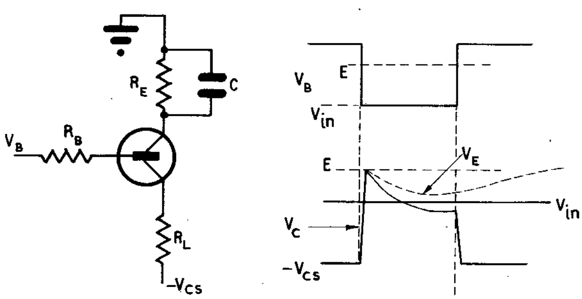


Figure 5

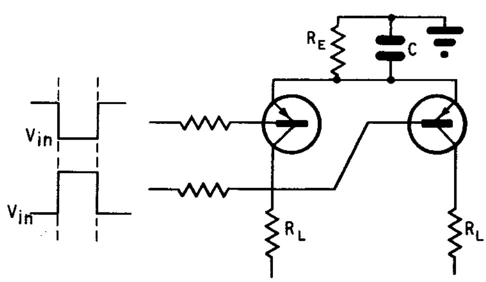


Figure 6

Another method of providing a temporary surge of current during the edges of the input waveform is by including an RC network in the emitter lead of the transistor, as shown in Fig.5. The condenser provides the low impedance for edge currents and the resistance defines the amount of steady current flowing through the transistor when it is conducting. V_B is a voltage source. The waveform which appears at the collector is as shown. It is presumed that

$$\alpha' \frac{V_{in}}{R_B} > \frac{V_{cs}}{R_L} .$$

Momentarily after the turn-on, the emitter can be considered to be earthed and the collector achieves a voltage near earth. But the current turned on charges up C, and the emitter voltage moves in a negative direction until, under steady conditions, V_c is at a voltage

$$\underline{\Omega} - V_{cs} + \left| \frac{V_{in}}{R_E} \right| \times R_L \text{ (assuming } \alpha' \text{ large)}$$

This type of circuit finds a particular application in flip flop type circuits which have a common R_E and C as shown in Fig.6. The cross coupling network between each collector and the base of the alternate transistor is not shown. However, in this circuit either one or the other transistor is <u>on</u> and thus the emitter potential before the change-over is V_{in} and is the same after change-over. This is an example of defined current working, and is half way on the road to non-bottomed transistor circuits.

To come now to transistor amplifiers working under non-bottomed conditions. It is possible to prevent the transistor from bottoming by clamping the collector voltage with a diode at some more negative voltage. Under these conditions the amount of current passed by the transistor is $\alpha' i_b$ (supposing the diode to have a very low resistance) and this current may be very much greater than V_{cs}/R_L. The hole storage in the diode may exceed the hole storage which would otherwise have occurred in the transistor. As the current available to remove holes stored in the diode can only be provided by the collector load resistor, the delay on turn off may be just as great, or even greater, than if the transistor had been operated in the bottomed region without the catching diode.

A method which is preferable to the above is to prevent the collector voltage becoming too near its bottoming voltage by arranging a feedback path from collector to base via a diode. See Fig. 7. When a current is fed into the base sufficient to bottom the transistor if the feedback were not present, the collector voltage moves from $-V_{cs}$ in a positive direction until the diode conducts at, say, -1.5 volts. Some of the current flowing through R_B now flows through the diode, the battery, and the collector to emitter path of the transistor, leaving only sufficient current flowing into the base of the transistor to maintain the collector current at a value which will give a collector voltage equal to 1.5 volts - the collector current being the sum of the current through the load resistance and also through the feedback path. Usually the current through the feedback path is much less than the collector current, and hole storage recovery time in the diode is very much less than the hole storage time which would have occurred if the transistor had been allowed to bottom. In practice the battery is provided by some form of resistor network as shown in Fig. 8.

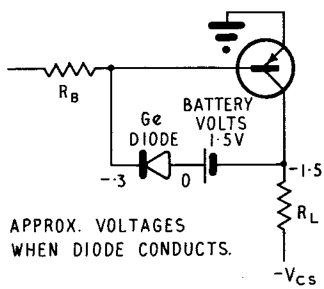


Figure 7

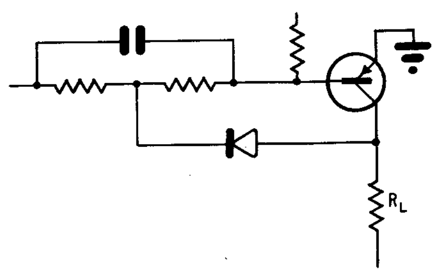


Figure 8

The circuit examples which will now be considered will employ bottomed transistors, mainly because they can act as bi-directional switches. A certain sacrifice in speed has been made on this count but the hole storage time has been reduced by the methods explained above. By extending the technique of non-saturating circuits to them, some improvement in speed can be obtained.

4. The Basic Bi-stable circuit

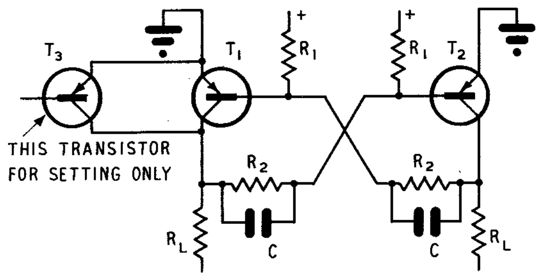


Figure 9

The basic circuit which is often used as the high speed memory element in a digital computer, is the transistor equivalent of the Eccles Jordan circuit. This may be looked at as if it were two inverting amplifiers connected back-to-back. As with the Eccles Jordan circuit, one of the transistors is arranged to be off whilst the other is on. If the on transistor is arranged to bottom, the voltage drop across it is so small and close to earth, that the collector voltage can be considered to be a reference voltage, namely earth. To cut off the off transistor completely, it is necessary to have a slightly positive bias on the base; this is the ideal, but in practice the amount of collector current flowing is very very small even with -100 millivolts on the base. The variation of this bias with temperature is fairly small.

The low input impedance of the base is a disadvantage because the values of cross coupling network resistance need to be low. When the transistor T_1 is <u>off</u>, the current through R_1 must be capable of providing the I_{co} as well as developing the bias across R_2; I_{co} may be 50 μA at elevated temperature.

When the transistor T_1 is <u>on</u> the base current must be sufficient to keep the transistor bottomed. With surface barrier transistors this is a point which has to be watched carefully as the current gain is much smaller (say $\frac{1}{2}$) at voltages near bottoming than what it was at, say, -3 volts. The condensers are not essential to the operation of the flip flop when it is acting as a flip-flop only, and the stimulus which sets it is applied for a long enough time. Under these conditions they only speed up the transitions between the two states of the circuit. But if the stimulus is very short compared with the changeover period, then they are essential to the operation of the circuit.

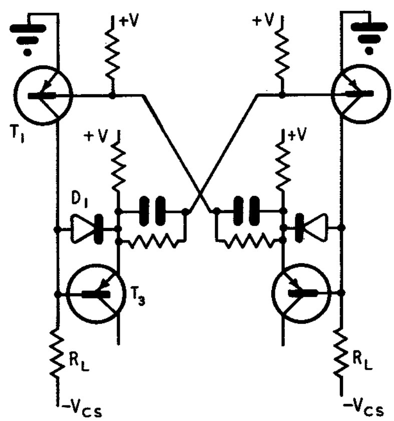


Figure 10

When the circuit is loaded, as it usually is, additional allowances have to be made for the load currents. When the load currents are large, it is necessary to use emitter followers, either as buffers between the load and the collectors, or connected inside the flip-flop loop as shown in Fig.10. In large signal operation however, emitter followers with a capacitive load become momentarily open circuit when a large positive voltage is applied to their bases; they can provide a high current output only in the one direction. Under these conditions the base to emitter voltage of T_3 exceeds the forward conducting voltage of D1 which turns on, and current to the load is now provided via the low collector impedance of T1. Thus the total impedance is the sum of the forward impedance of the diode and saturated collector impedance of the transistor. The emitter follower in fact, is not acting as a buffer during the positive-going direction of the waveform, and as the low impedance depends on the transistor T_1 being bottomed, even more care must be taken to arrange that enough base current is provided to it. Using normal alloy transistors (f_{co} 5-10 Mc/s) the time of a complete changeover is approx. 0.5 μsecs to 1 μsec. By preventing bottoming this speed can be increased by about a factor of 2. With surface barrier transistors of the SB.240 type a changeover period of a little over 100 m μsec. is possible (perhaps 150 m μsec if appreciable load currents are required). By preventing bottoming changeover times of perhaps 70-80 m μsec. are possible, (about 100 m μsec if load currents are required).

5. The emitter follower

At this stage it is as well to be reminded that the emitter follower current gain is a function of frequency and falls off at high frequency until it becomes unity at f_{co} (as explained in the lecture on a.c. amplifiers). From a large signal point of view, when a step function of voltage is applied at the base of the transistor, the diode between base and emitter passes current to the load with only a short delay (approx. T_N), but only after a time T_N' is the transistor acting as an emitter follower with its equilibrium value of current gain. For a surface barrier transistor $T_N' \simeq 75$ mμsec. Thus the input impedance of the emitter follower only gradually builds up to its high value. Until then, the current from the emitter must be partly provided by the source which it was intended that the emitter follower should buffer.

6. Setting the State of the Bi-stable circuit

There are many ways of setting the flip flop, e.g. diodes may be used to feed pulses to the circuit either at the base or collector points. A very common method using transistors, is to place the setting transistor T_3 in parallel with the flip-flop amplifier transistor T_1, see Fig.9, so that when the setting transistor is turned on, the collector voltage of the amplifier transistor approaches bottoming voltage. This positive movement of voltage turns off transistor T_2 if it was not already <u>off</u>, and brings <u>on</u> T_1. It is to be noted that to set the flip flop into one state or the other, requires 2 transistors, or 2 diodes if diode gating be used.

By using a transistor as a bi-directional switch only one switch is required to set the flip flop into either the on/off or the off/on state. This is illustrated in Fig.11. A large voltage pulse is applied to the base resistance R_B of T_3; the pulse being large compared with the voltages which are available at the emitter or collector of T_3. Thus T_3 is turned on and bottomed every time the input pulse is applied. The impedance at the base of T_2 is always large compared with the setting voltage sources and thus, when the input pulse is applied, the base voltage of T_2 is clamped to the setting voltage. When the setting voltage is +1, T_2 is turned off; when the setting voltage is -0.5, T_2 is turned on. In practice the source of the setting voltage is usually another flip flop having a low impedance output.

An extension of this system to provide a shifting register is shown in Fig.12. The delay line provides a delay which is long compared with the time for which the switch is closed. Let us consider that equilibrium has been reached and thus a voltage is available at the left hand side of each switch corresponding to the state of the flip flop. When all the switches are closed this voltage sets up the flip flop to which it becomes connected. New information about the state of the flip-flop now progresses down the delay lines but, before it reaches the switches, these are opened.

By interconnecting the output of one flip flop collector via a delay line and switch to the base of the same transistor, the flip flop changes over once for each input pulse and therefore acts as a counting circuit or halver. See Fig.13.

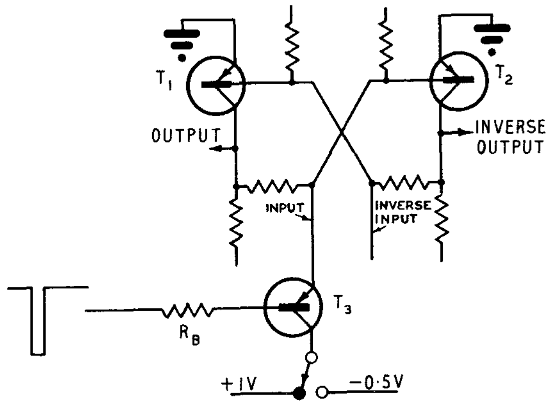


Figure 11

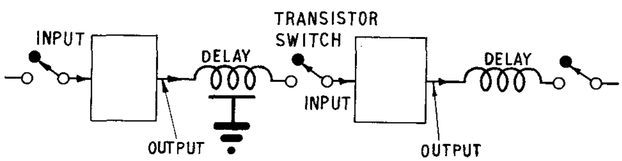


Figure 12

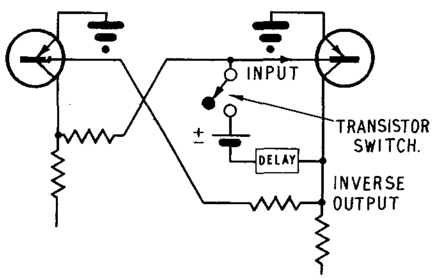


Figure 13

7. Direct coupled circuits

The circuits shown up to now, have had couplings from the collector of one transistor to the base of another via a potentiometer chain. In the class of circuit known as direct-coupled circuits, the couplings are made directly from collector to base. The principle is illustrated in Fig.14 and Fig.15.

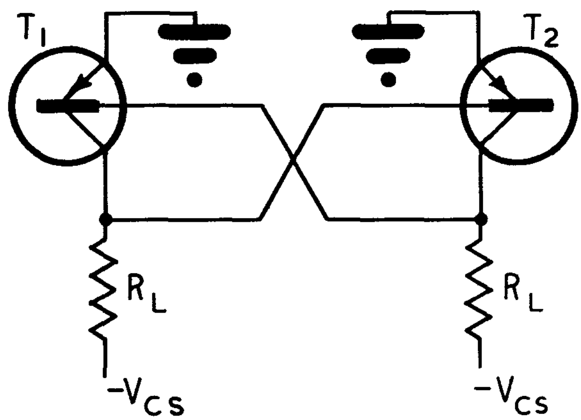


Figure 14

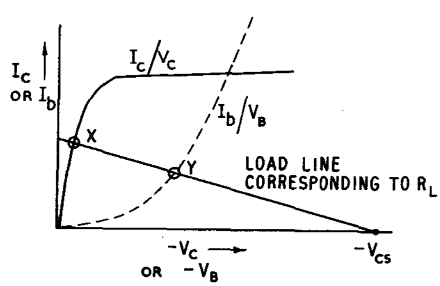


Figure 15

Let us suppose that transistor T_1 is conducting and is bottomed, the voltage at its collector is therefore that corresponding to the point X. This, applied to the base-emitter of transistor T_2 is sufficient to allow only a very small amount of current to flow to the collector of T_2. Effectively T_2 is an open circuit between both emitter-base and emitter-collector and thus all the current flowing through R_L of T_2, flows to the base of T_1 holding this transistor bottomed. The circuit is stable in this condition. By symmetry there is another stable state with T_2 conducting and T_1 non-conducting. The voltage at the top end of R_L therefore swings between voltages corresponding to the positions X and Y. If V_{cs} is large compared with this voltage swing, the current flowing through each load resistance is nearly constant and flows either to the base of one transistor or to the collector of the other.

As the base current to the transistor is so large when the transistor is conducting, it might be expected that the hole storage time would be very long indeed. The amount of holes stored certainly is large, but the reverse current is also very large. The path for the reverse current of one transistor is via the saturated collector impedance of the opposite transistor to ground, and therefore at turn-off, the base of the transistor is virtually connected to earth. The reverse current is limited therefore by the base resistance of the transistor which is being turned off. Even so, the hole storage time is usually much longer than the rise and fall times, these being short because the voltage excursion is only of the order of 0.3 volts. With surface - barrier transistors, rise and fall times are approx. 30-40 millimicroseconds and hole storage times are from 50 to 100 millimicroseconds. Surface-barrier transistors are particularly suitable for use with this technique, having such relatively short hole storage times.

Outputs can be taken to circuits of a similar type and such output loads actually make the hole storage times shorter, as current which would otherwise flow to a base and give excess holes, is diverted to the load. On the other hand, the diversion of the current from the base when a transistor is being turned on, makes the turn-on transient longer.

The circuit is very simple, but its output swing is insufficient to control bi-directional gates and is also insufficient to drive more than about 2 or 3 circuits in an AND gate. For these reasons more transistors are required to carry out the same operations as could be done with a few less transistors and a great many more other components. As far as speed is concerned, direct coupled circuits are as fast as a good number of other types of circuits, and only by using very complex circuits can a significant improvement in speed be obtained.

For small scale equipments the technique seems very suitable, but when larger scale equipments are wanted, the small voltage swing is a serious disadvantage. At these speeds of operation it is very difficult to provide a good earth throughout a large equipment, and the pulse currents which occur in the lead inductance generate voltages which are significant compared with the voltage swing available from the direct coupled flip flop. On the other hand, some large airborne computers have been made using this technique. The technique is more suitable with silicon transistors which need a much larger base emitter voltage before collector current flows.

8. Logic circuits

It is proposed to deal only with two types of such circuits; those providing the facilities of OR and AND. An OR circuit has several inputs, and provides an output when any one or more of the inputs is stimulated. The most usual type of OR circuit has already been used in Fig.9. The two transistors and their common collector load form the OR circuit. If either or both transistors are turned on hard, the voltage at the common collector takes up a voltage equal to the bottoming voltage, and to a first order this is unaffected whether either or both transistors are conducting. The number of transistors which can be connected in this way is limited by the amount of movement of voltage which can be tolerated at the common collector when all the transistors should be off. At high temperatures this voltage drop across the load is $nI_{co} R_L$. At high speeds, the number is limited by the stray capacity which each extra input introduces. With this type of OR circuit there is a phase reversal between the input and output.

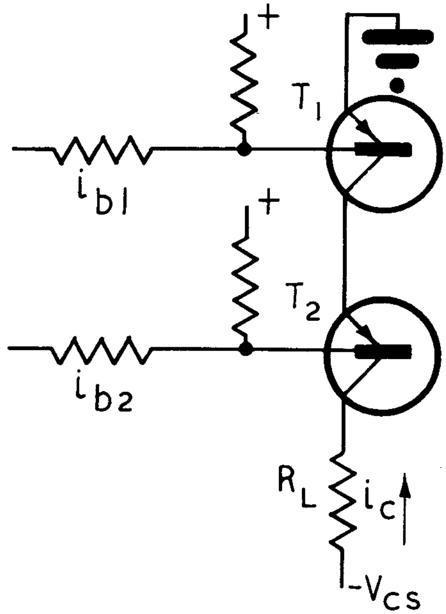


Figure 16

One simple form of AND circuit is shown in Fig.16. Only when both T_1 AND T_2 are conducting can current flow through R_L. It is usual to provide sufficient base drive to each transistor to enable it to bottom, and under these conditions the voltage excursion at T_2 collector is between $-V_{cs}$ and twice the bottoming voltage. For an n input AND circuit the upper limit of this voltage is n times the bottoming voltage and this may set a maximum limit to n. Another and more serious limit to n, is that set by the base current drive needed for the first transistor (i.e. the one which has its emitter connected to ground). If i_c is the current required in R_L, the base current drive to the n^{th} stage is $> \frac{i_c}{\alpha'}$, say $\frac{i_c}{5}$, and the current flowing from the emitter is $i_c + \frac{i_c}{5}$ i.e. $\frac{6i_c}{5}$. The base drive to the $(n-1)^{th}$ base is thus $\frac{6i_c}{5x5}$ and the amount of current flowing from the emitter is $\frac{6i_c}{5}\ (1 + \frac{1}{5})$. The collector current of the 1st stage is therefore $(\frac{6}{5})^{n-1} \cdot i_c$ and the base drive needed is $\frac{1}{5}\ (\frac{6}{5})^{n-1} i_c$. If n = 10 the base drive needed at the first stage $\simeq i_c$.

This type of AND, also provides a phase inversion of the input voltage.

THE FUNDAMENTAL PRINCIPLES OF TRANSISTOR CIRCUITS

LECTURE 7. DIGITAL TRANSISTOR CIRCUITS USING FERRITE MEMORY CORES

by G.H. Perry

1. Introduction

During the past ten years digital computers have changed from being rather speculative mathematical tools until today computer like machines are being exploited for a variety of data handling problems. There is consequently a great demand for simple electronic digital circuits that are both reliable and cheap; the circuits to be described are aimed at these requirements. Digital circuits are usually designed around a two state memory device, the most common of which is a bi-stable circuit consisting of a pair of valves or transistors. The circuits to be considered are based on the memory properties of magnetic cores having rectangular hysteresis loops, (Fig. 1) and the fact that a winding coupled to a core is an ideal low impedance drive to the base-emitter circuit of a junction transistor. For one of the two possible directions of flux change within the core, the polarity of the voltage developed across this winding is such as to turn the transistor 'ON' so that the emitter and the collector act as two terminals of a switch, and it is thus capable of changing the state of another core by controlling the current flowing in one of its windings.

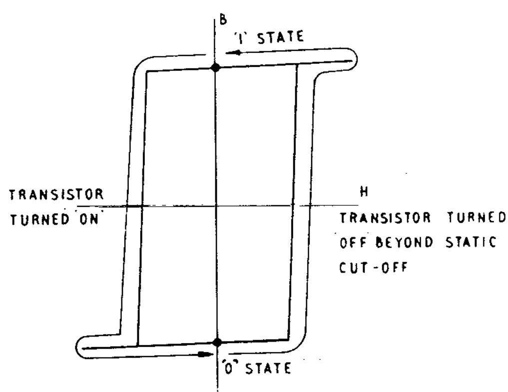


Figure 1 Idealised B/H Curve of Ferrite Memory Core

One of the big differences between these circuits and those of a more conventional type is that the digital information is not continuously available in the form of a D.C. potential. Methods of continuously reading the state of a magnetic core without permanently changing its state are invariably complex, so use is made of the simple system of clearing the core to a common state, and detecting whether or not a change of flux occurs. Despite the possible disadvantage that information is not available continuously, systems of this type have now been used for many practical applications. Ref. 2, 3, 4, 5, 6, 7, 8, 9. These circuits only require current during the change of state, a period of a few μ sec., and are consequently particularly economical at low PRF's and for apparatus with long periods of quiescence. This system, where information is only available when requested, does however require a completely different approach from the more conventional type of circuit, but it is hoped to show that except for a few cases where continuous indication is essential, this is no great disadvantage.

In this lecture I plan first to consider the basic circuit and the design procedure, then to show the way simple pulse separator, shift register and counter circuits can be constructed, and finally some of the ways ferrite memory cores can be used in conjunction with controlled hole storage effect will be discussed.

2. The Basic Circuit

Consider first the simple arrangement of Fig. 2. The transistor is used in the common emitter condition making use of the fact that the non-linear characteristics of a junction transistor approximate very closely to those of an ideal switch. When the transistor is "ON" the emitter-collector potential will approach zero, and when it is "OFF" the collector current will be very small indeed. Most of the power will be dissipated during the times of transition between these two states, and consequently it becomes possible to switch relatively heavy currents without exceeding the dissipation of low power transistors. When the transistor is "ON" the collector current is defined by R, -V, and the sum of the back E.M.F.s of the cores to be driven. Only in cases where large random numbers of cores have to be switched in series is it necessary to provide special constant current drives and in practice this is not frequently required.

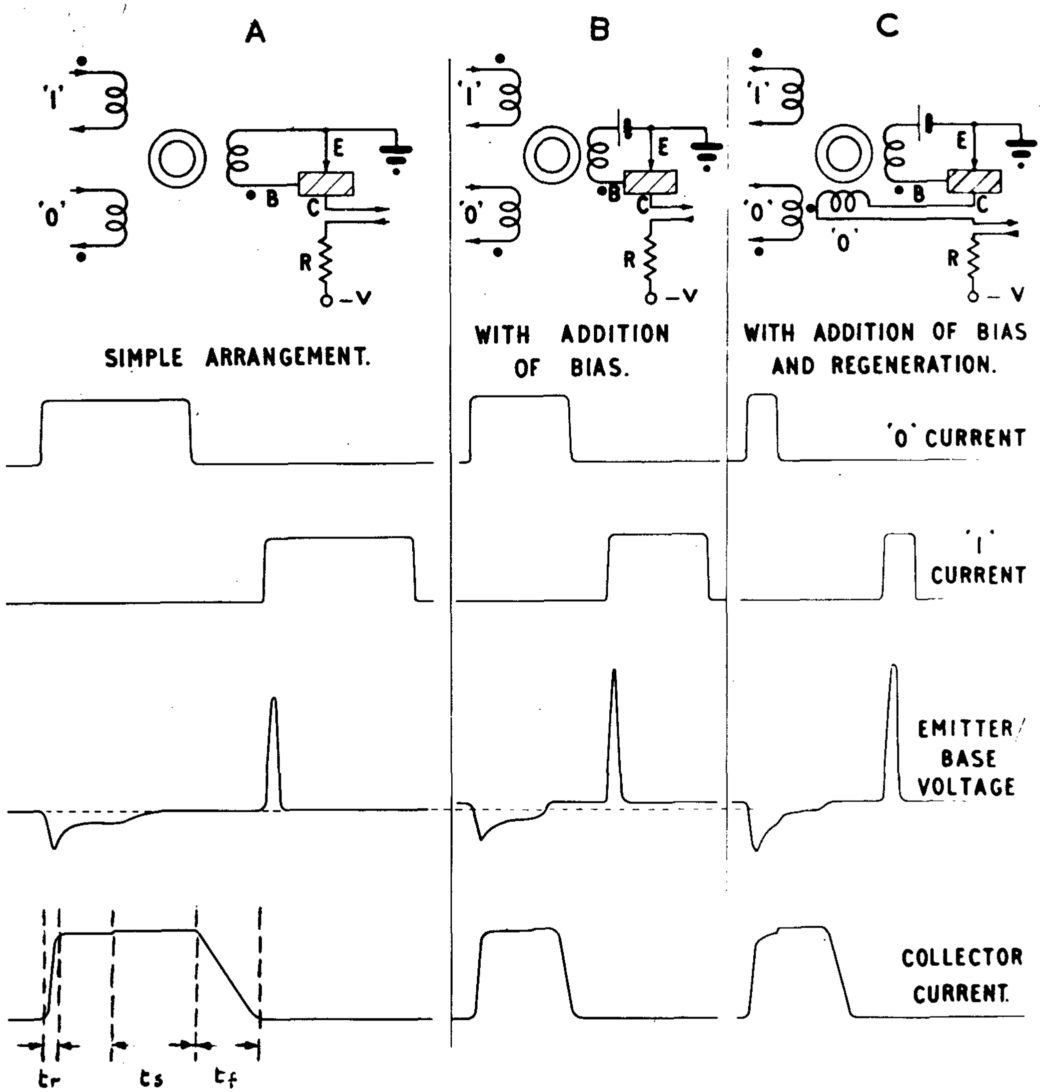


Figure 2 Ma netic Core and Transistor Combinations

In the quiescent condition the secondary winding on the core forms a short circuit between the emitter and the base of the transistor. Operated under these conditions typical collector currents are between 2-10 times I_{co}. This is still small compared with the tens of milliamps flowing when the transistor is in the 'ON' state.

The convention used to define the binary '1' and '0' is that a flux change from '1' to '0' causes the polarity of the voltage developed across the emitter-base winding to be such as to turn the transistor 'ON' thus the base of a p.n.p. transistor is taken negative during a '1' to '0' change of flux. Conversely the base is taken positive with respect to the emitter, so the transistor is taken beyond cut-off during a '0' to '1' transition. Drive windings are marked with multiples of '1' and '0' to indicate the magnitude and direction of the magnetizing force they produce e.g. a winding marked 2 x '1' would be used to produce a magnetizing force twice that required to switch to the '1' state.

Whilst a core is switching, the voltage developed across a winding coupled to it is

$$e = \frac{nd\phi}{dt} \qquad (1)$$

Where 'e' is in volts; 'n' is the number of turns; ϕ is the flux in Webers; and t is time in seconds. Where however a winding is coupled to the emitter-base circuit of a transistor, the voltage 'e' will, during the '1' to '0' transition, be limited, to Veb due to the input characteristics of the transistor. Since this voltage remains substantially constant during switching, $\frac{d\phi}{dt}$ remains constant thus:

$$Veb = N_2 \frac{\Phi}{T} \qquad (2)$$

or

$$N_2 = \frac{Veb\ T}{\Phi} \qquad (3)$$

Where Φ is total flux change in Webers from the '1' to '0' state, T is the total switch time of core in seconds, and N_2 the number of turns on the emitter-base winding.

The switching properties of square loop magnetic materials are characterised by the imperical formula

$$S = (H - Hc)\ T \qquad (4)$$

Where H and Hc are respectively the applied and coercive magnetizing forces, and 'S' is the switching constant. For a particular size of core (4) may be expressed in terms of mmf's as:

$$S = (Fc - K)\,T$$

$$\text{or} \qquad Fc = \frac{S}{T} + K \qquad (5)$$

Where Fc and K are mmf's in Ampere-Turns associated with the magnetizing forces H and Hc respectively, and S in this case is in Ampere turn - Seconds.

The current flowing in the emitter-base winding produces an mmf. Fs, which like Fc is in direct opposition to the mmf. from the drive winding Fp.

$$\therefore \quad Fp = Fs + Fc \qquad (6)$$

When the core is switched back into the '1' state the polarity of the voltage Veb is reversed. Under these conditions the emitter-base impedance is high, and no current will flow in N_2, so that the switching speed of the core will be fast, determined solely by the m.m.f.s. Fp and Fc.

The operation of the transistor as a switch has been analysed by Messrs. Ebers & Moll Ref. 1, and also by Mr. S.W. Noble in his last lecture. Unfortunately neither of these approaches completely characterizes the transistor when it is operated under large current conditions. Consequently it is at the present essential to use an experimental approach.

From Fig. 2A it will be noted that the storage time t_s of the collector current waveform is very long and that the turn-off transient t_f is slow. This is due to the fact that current I_{b2} is very small, being determined by the voltage drop across the emitter-base junction and the base resistance R_b, consequently a certain amount of minority carrier injection continues to take place.

In Fig. 2B the base of the transistor is biased positively relative to the emitter, this has the following advantages:-

1. In the quiescent period the collector current will be I_{co}, and even at 100°C this will only be in the order of one milliampere which is small compared with the tens of milliamperes that flow during the 'ON' time.

2. Equation (3) now becomes $N_2 = \dfrac{(Veb + VB)T}{\Phi}$ (7)

   thus making the switch time less dependent on the characteristics of the individual transistors.

3. When the core has completely switched, the base is taken positive of the emitter so that the emitter starts to act as a collector of stored holes and prevents any further injection. It will be seen that the storage time of the collector current waveform Fig. 2B now becomes small, and the duration of the collector current approximates closely to the switch time of the core.

In both Figs. 2A and 2B the '0' current pulse has to be equal to or longer than the time required for the core to switch. In Fig. 2C the collector current is passed through an additional winding on the core so as to continue the switching of the core after the '0' pulse of current has terminated. This pulse widening effect has the advantage that under certain circumstances it can be used as a temporary memory; at the same time the power required from the '0' pulse is reduced. It also enables circuit arrangements that are far less sensitive to pulse width to be designed.

3. Design Procedure

As an example a circuit will be designed that consists of two basic circuits of the type shown in Fig.2., coupled together to form a multi-vibrator with a 10 μsec. period, Fig. 3. Both cores are biased with a D.C. current through a winding to produce an mmf. capable of switching the cores from the '1' to the '0' state. The current from each collector is taken through a winding on the opposite core to produce an m.m.f. large enough to overcome the mmf. due to the D.C. bias current and switch the core from the '0' to the '1' state. It may be noted that this type of circuit is not self starting. Starting may be achieved by a momentary short circuit between the emitter and the collector of either transistor so as to force one of the cores into the '1' state.

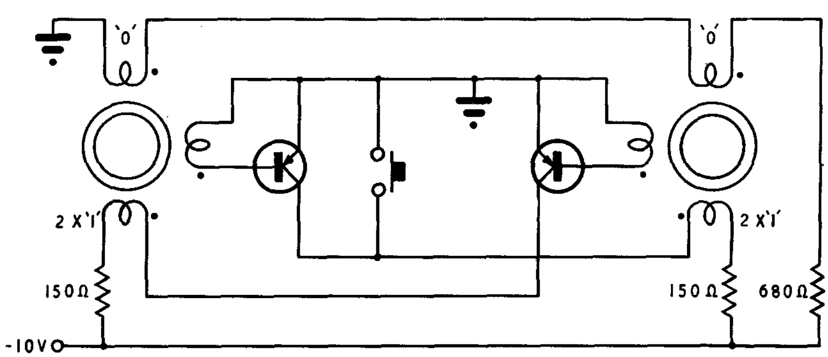


Figure 3 Multivibrator

Tests on a batch of the type of transistors used (XA102's) showed that for a I_c rise time of 0.5 µsec., maximum current gain, base to collector, occurred at I_c = 50mA, and that at this current and rise time the worst sample had a gain of 10; therefore the max I_b required is 5mA. Under these conditions the average value of Veb = 0.35 Volts.

The magnetic core to be used is of the type used in matrix stores, it is 2 mms. in diameter and of D3 material and chosen because of its low coercive force. The relevant details are :-

$$\Phi = 9 \times 10^{-8} \text{ Webers}$$
$$K = 0.35 \text{ AT}$$
$$S = 0.34 \times 10^{-6} \text{ AT. Secs.}$$

First, the emitter-base turns are determined by using equation (3) this gives an answer of 20 Turns. The mmf. due to the secondary current and turns Fs can now be obtained, Fs = $N_2 \times I_b$ i.e. 0.1 AT. Next the mmf. Fc. required to switch the core in 5 µsec is calculated by using equation (5) to be 0.42 AT
From equation (6) the drive mmf. Fp equals 0.52 AT

In this case Fp is derived from a D.C. current, and a reasonable compromise is 12.5mA through 40 turns. The mmf. used to set the cores to the '1' state must do so in less than 5 µsec at the same time as neutralizing the effect of the D.C. bias current. With a collector current of 50mA., 20 turns enables this to be performed in 2.25 µsec. Assuming a 10V supply we can now determine the resistor values required. Three resistors only are required, one for each collector circuit and one to define the D.C. current which biases the cores to the '0' state.

The collector resistors are determined by the supply of 10V, the collector current of 50mA, and the back emf due to the 20 turns 2 x '1' winding which will be roughly 2V peak.

Hence $R = \frac{10 - 2}{0.05} = 160$ ohms or 150 ohms nearest standard value.

For the bias current we have, 10V supply, 12.5mA current and a back emf which is Ve-b x turns ratio between the '0' and e-b windings i.e. 0.70V.

Hence $R = \frac{10 - 0.7}{0.0125} = 745$ ohms or 680 ohms nearest standard value.

4. Pulse Separator

A pulse separator circuit can accept a series of clock pulses and distribute them along a set of leads, so that pulse 1 appears on lead 1, pulse 2 appears on lead 2 and so on. A simple pulse separator can be constructed by using a chain of transistor-core combinations of the type shown in Fig. 2. These are connected so that the collector current of one transistor sets up the next core in the chain to the '1' state, all the '0' windings of the cores in the chain are driven by the clock pulses. Once a '1' state has been introduced into the beginning of the chain it will be transferred to the next core on the arrival of the clock pulse. If the last transistor is coupled to the first core this single '1' state can be made to circulate.

There are many applications for a circuit of this type, for example in dynamicizers, sampling systems, computer sequence control and ring counters. An example of a pulse separator is given in Fig.4. In this circuit a means is provided to ensure that the circuit will automatically assume and maintain the correct operating condition i.e. the circulation of a single '1' state.

At each input current pulse C_1, which is biased to the '1' state by a D.C. current, is switched to the '0' state causing T_1 to be turned 'ON'. It will be assumed that C_2 to C_n all start in the '0' state. When an input current pulse occurs the current of T_1 produces an mmf. in C_2 equal and opposite to the mmf due to the input current pulse, but since T_1 current pulse is designed to be of a longer duration, C_2 would be eventually switched to the '1' state.

The next input pulse will again cause C_1 to turn on T_1 and attempt to set C_2 to the '1' state, but C_2 will have already started to change to the '0' state causing T_2 to turn 'ON'. The collector current pulse of T_2 is longer in duration than both T_1 and the input pulses and performs two functions; it sets C_3 to the '1' state, and produces an mmf in C_2 of sufficient amplitude and duration to inhibit the effect of T_1 current. At each subsequent input current pulse the '1' state is moved one place to the right, and since the current of each transistor except the n^{th} is fed through the common winding on C_2 '1's are prevented from being set into C_2 until the n^{th} core is cleared from the '1' to '0' state.

Externally this circuit appears to behave like a continuous ring of core-transistor circuits. It has however the great advantages that it automatically gets itself into the correct working condition, and also enables the exact position of a fault anywhere in the chain to be readily located.

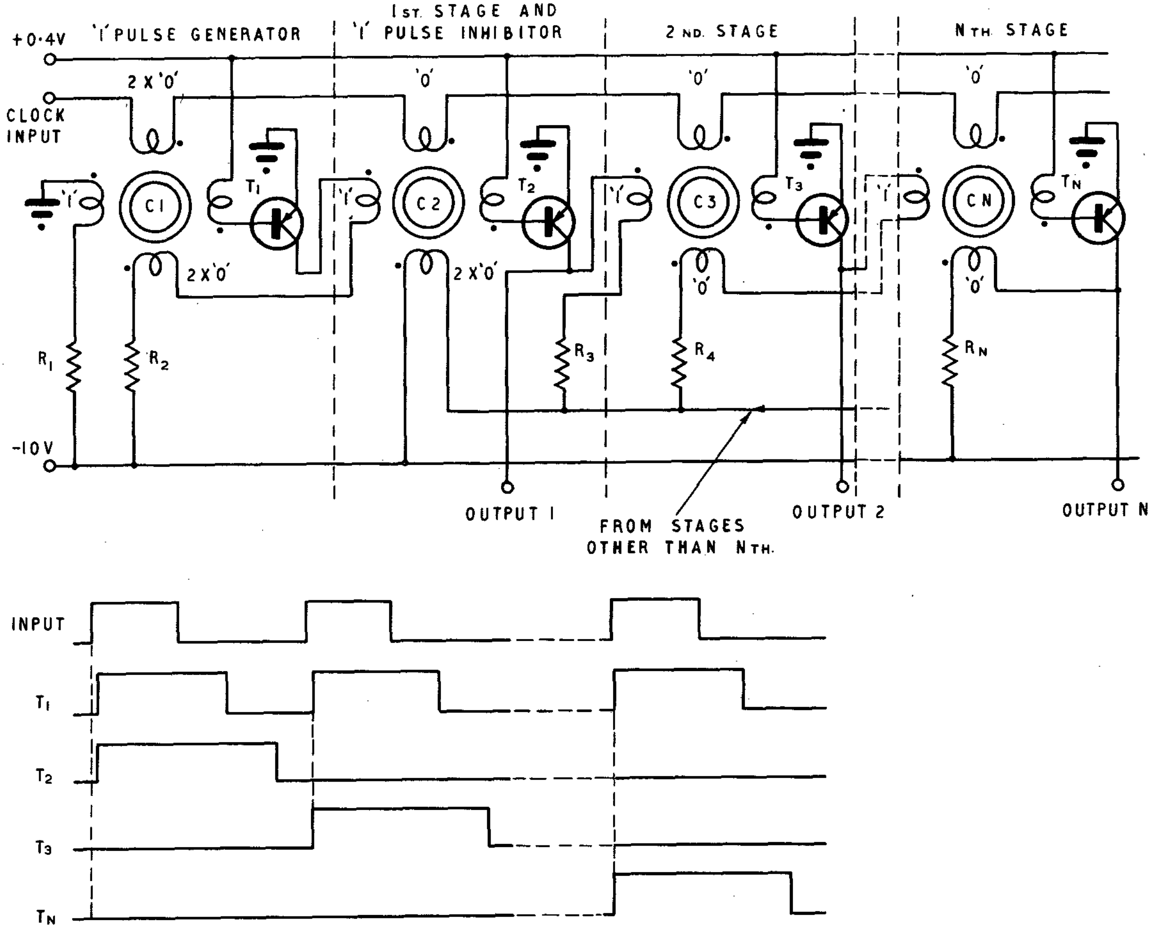


Figure 4 Pulse Separator

5. Shift Register

Whereas in pulse separators a single '1' state is shifted along a chain of memory cores; the shift register has to shift any combination of '1's and '0's, consequently it is essential to have a temporary memory between each of the permanent memory cores. One arrangement for this is shown in Fig.5. C_1 acts as a permanent memory and C_2 as the temporary memory. C_2 is biased into the '0' state by a D.C. current. If C_1 is in the '1' state when the shift pulse is applied, T_1 will be turned 'ON' and its collector current will switch C_2 into the '1' state. At the end of T_1 current pulse the bias current will cause C_2 to revert to the '0' state turning on T_2 and thus cause the next memory core to be set into the '1' state.

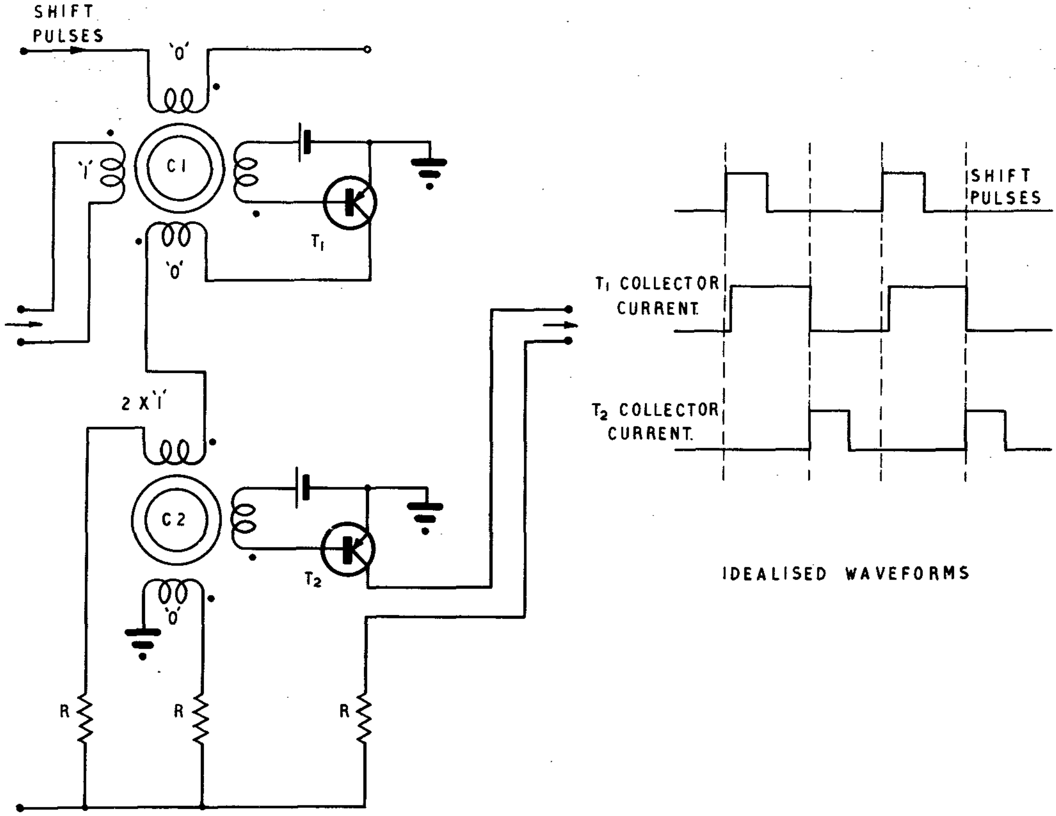


Figure 5 Section of Shift Register

6. Binary Counter

The minimum requirements of a binary counter would appear to be as follows :-

(a) A permanent memory device that has two stable states.
(b) A gate which will set up the route of the input pulse according to the state of the permanent memory.
(c) A temporary memory to remember the previous state of the permanent memory for at least the transition period.

It should however be noted that it is not necessary to have three physically separate units to fulfil these requirements.

A single stage of one of the many binary counters that are possible with cores and transistors is shown in Fig. 6. Core C_2 acts as the permanent memory, and the gating action is due to the voltage developed across the winding 'y' preventing the voltage developed across the winding 'x' from turning 'ON' on the transistor. The temporary memory is provided by the time required for the flux to change in C_2.

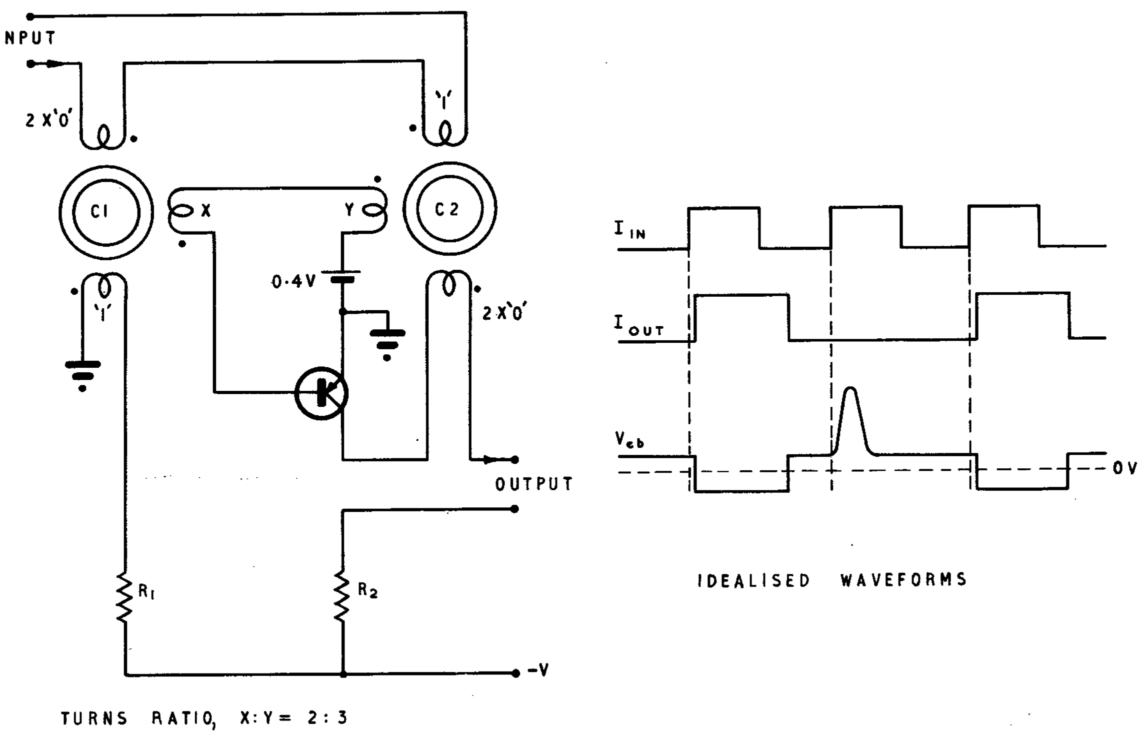


Figure 6 Binary Counter

If C_2 is in the '1' state, the arrival of the input pulse will have very little direct effect on the state of C_2, but as C_1 is normally biased into the '1' state the input pulse of current will always switch C_1 into the '0' state. The voltage developed across the winding 'x' will cause the transistor to be turned on, so that its collector current passing through a winding on C_2 will produce an m.m.f. to oppose the m.m.f. due to the input pulse, and cause the core to switch into the '0' state. As C_2 changes into the '0' state the voltage developed across the winding 'y' will be such as to keep the transistor 'ON' until C_2 has completely switched.

C_2 will be in the '0' state when the next input current pulse arrives; so the core will be directly switched into the '1' state and the voltage developed across the winding 'y' will prevent the voltage developed across the winding 'x' from turning on the transistor. Thus for every two input current pulses the transistor will only give one output pulse.

7. THE USE OF A JUNCTION TRANSISTOR AS A THREE TERMINAL TEMPORARY MEMORY FOR DIGITAL CIRCUITS

The phenomenon of carrier storage in junction transistors has generally been looked upon as one of their undesirable characteristics. Many authors have described both the effect and circuits to minimise it, but it will be shown that carrier storage when controlled can provide a three terminal temporary memory for digital computing circuits of a type not provided by any other single electronic component.

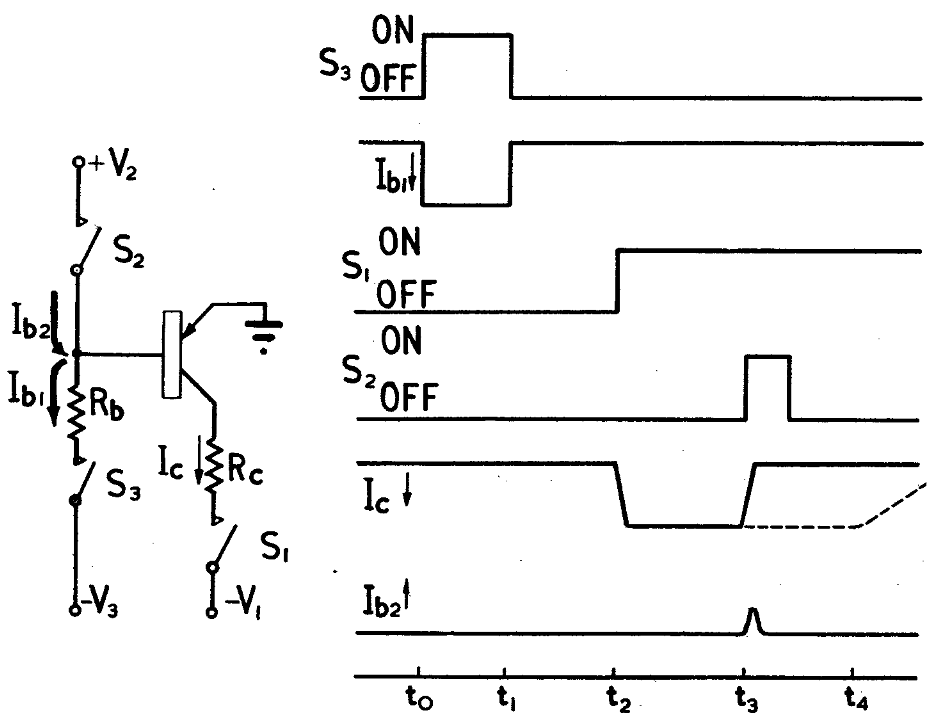


Figure 7

A p.n.p. transistor in this controlled mode of operation is shown in Fig. 7. Carriers are initially injected by the current pulse I_{b_1}. After t_1 the base is left open circuit. At t_2 the collector current is permitted to flow, finally being terminated by the closing of S_2. The switches S_1 and S_2, could in practice be common to a number of such transistors, and themselves would be transistors programmed to operate by a central control. The storage time $(t_4 - t_1)$ is dependent on both the amplitude and the duration of I_{b_1} and also the amplitude of I_c, but experimental measurements show that it is relatively independent of t_2, the time at which S_1 is closed; these relationships for a typical transistor are shown in the form of graphs in Fig.8. When the current I_{b_1} is terminated and the base of the transistor is left open circuit, the holes stored at this instant will start to die away according to the life time characteristics of the particular base material. At t_2 S_1 closes and the collector current is permitted to flow. Provided that there are still sufficient holes stored in the base the collector will bottom, and a diffusion gradient determined by $I_c = V_1/R_c$ will be established. Since the base of the transistor is still open circuit

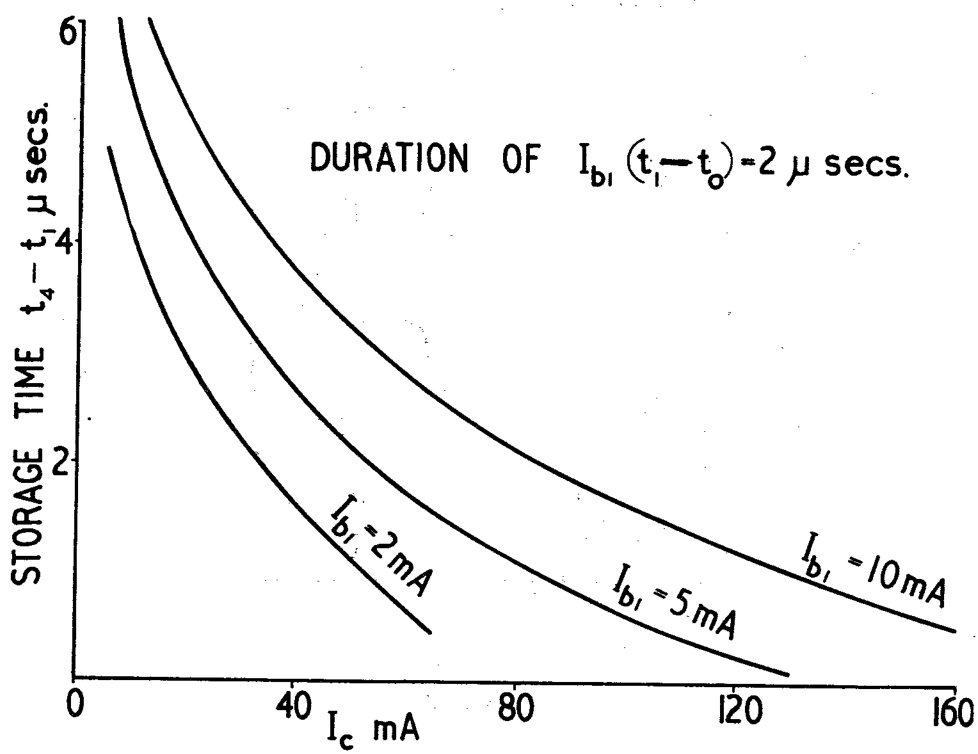


Figure 8

all collector current must flow through the emitter and thus cause further holes to be injected into the base. The number of holes stored in the base would decay with time until t_4 when the transistor would no longer remain bottomed. However, S2 is closed at t_3 so as to take the base of the transistor positive with respect to the emitter, thus preventing further injection of holes and enabling the emitter to help extract the holes already stored. In this way the collector current is terminated very rapidly at a predetermined time.

By suitable choice of operating conditions the quantity of electrical charge supplied to the base-emitter circuit may be small compared with the total charge thereby caused to flow in the collector-emitter circuit. The amplitude and duration of I_{b_2} required to terminate I_c at various delay times for a typical 5 Mc/s α cut-off transistor is illustrated by Fig. 9.

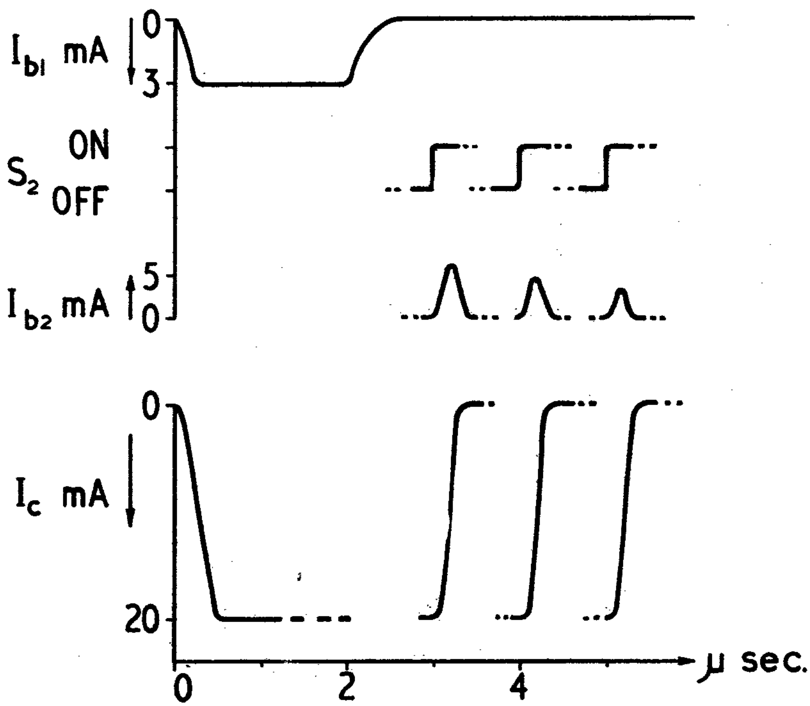


Figure 9

A transistor operated in this way performs the function of both an amplifier and a temporary memory at the same time, and is in many ways analogous to a black box containing a monostable circuit that can be returned to its stable condition by a clocking pulse before its natural relaxation time. One application of this effect has been in the design of a shifting register using rectangular hysteresis loop magnetic cores as the permanent memories. Fig. 10 shows the circuit of a simple register. After switch S_1 is closed, the shifting pulse clears all the cores to the '0' state of remanence. Negligible change of flux is produced in a magnetic core previously in the '0' state and consequently negligible secondary current will flow. However if the core is in the '1' state, a current will flow causing holes to be injected into the base of the following transistor. The S_1 switch is opened again before the end of the shift pulse to ensure that each core is swtiched completely to the '0' state. Finally, the bases of all the transistors are taken positive with respect to their emitters to ensure a well defined and rapid turn-off of the collector currents.

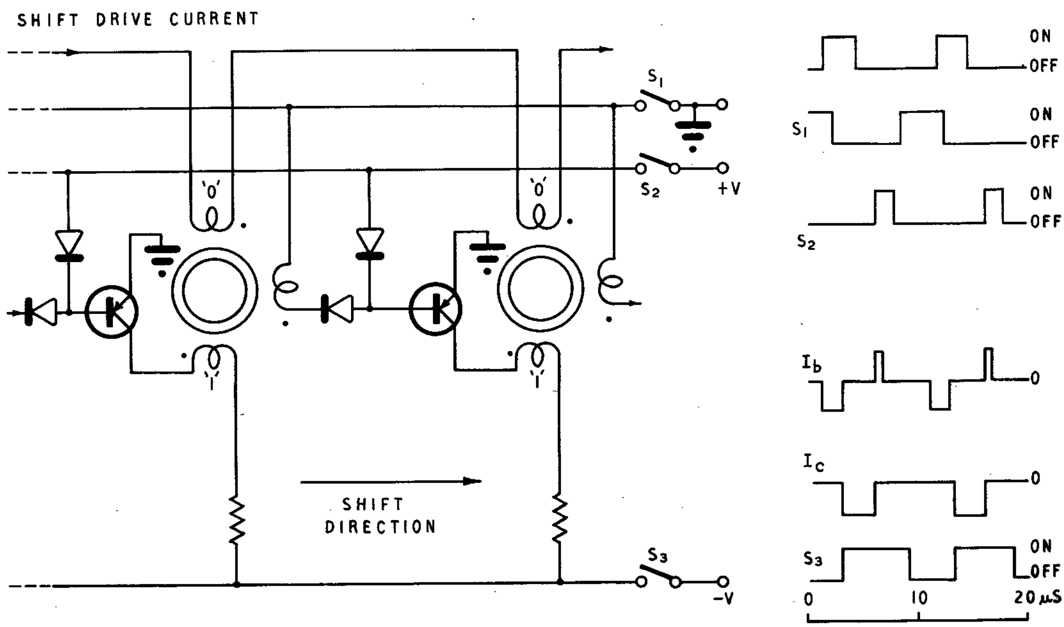


Figure 10

Fig. 11 shows how, by the use of a core with three secondary windings the register can have non-destructive read-out in addition to the ability to shift both left and right.

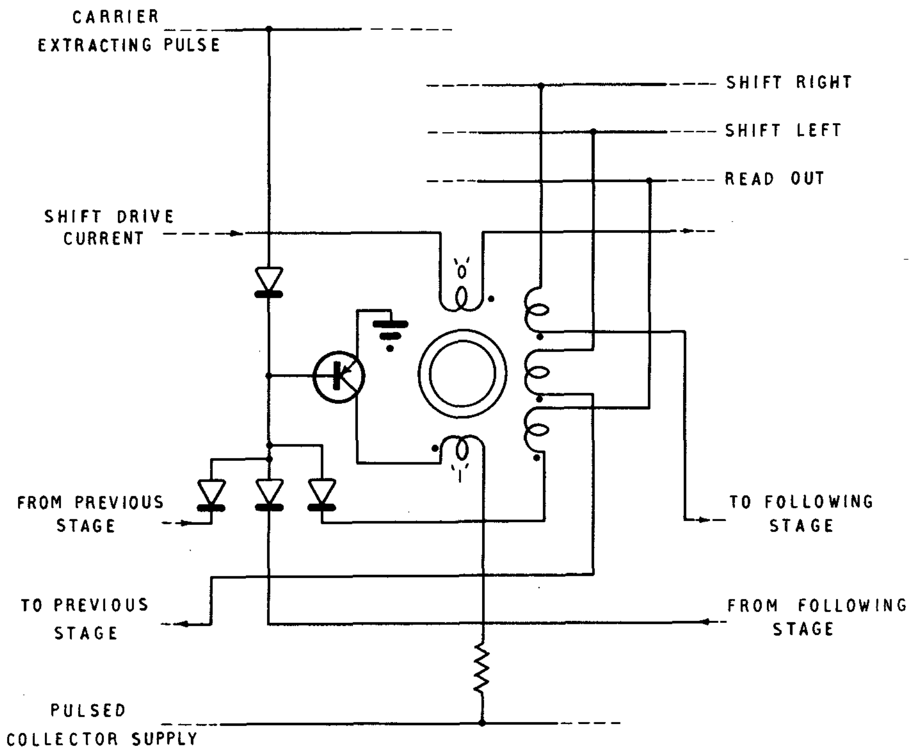


Figure 11

A circuit element of this type has many other practical applications. For example, its memory properties may be used as a delay to compensate for the propagation delays in logical networks, and the operation 'inhibit' can be performed by the extraction of the stored carriers before the collector is connected. A two state circuit can also be constructed as shown in Fig.12. When the switch S_2 is closed, the collector current that flows will depend on whether or not holes are stored in the base of the transistor. If no holes are present the transistor will remain 'OFF', however, if holes are present, collector current flows and since this current is coupled to the base by T_1 further holes are injected and the transistor is maintained in the 'ON' state.

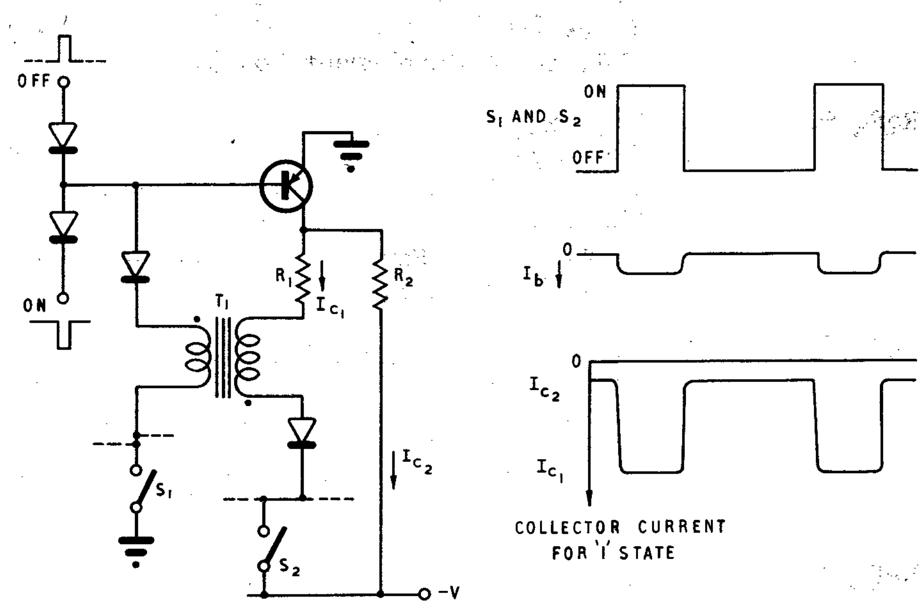


Figure 12

References

Ref. 1. Ebers, J.J., Moll, J.L.
"Large Signal Behaviour of Junction Transistors
Proc. I.R.E. December, 1954 p.1761

Ref. 2. Guterman, S.S., and Carey, W.M.
"A Transistor Magnetic Core Circuit: A new Device applied to Digital Computer Techniques"
Convention Record I.R.E. 1955 Part 4 p 83.

Ref. 3. Perry, G.H., Hoffman, G.R. and Shallow, E.W.
"A New and Simple Type of Digital Circuit Technique Using Junction Transistors and Magnetic Cores".
Proc. I.E.E. Paper No. 2112 M Nov.1956.
Vol. 103 B Supplement No. 3.

Ref. 4. Hoffman G.R., Maclean, M.A.
"Quiescent Core-Transistor Counters"
Proc. I.E.E. Paper No. 2122 Nov. 1956.
Vol. 103 B Supplement No. 3.

Ref. 5. Six, W., Koolhof, R.A.
"Some Applications of Square-Loop Ferrite Cores in Telecommunication Switching Circuits".
Proc. I.E.E. Paper No. 2206 Oct. 1956.
Vol. 104 Part B.

Ref. 6. Hayward, R.H., Bubb E.L.
"'ERNIE' - The Electronic Random Number Indicator Equipment for the Premium Savings Bond Prize Draw"
Post Office Electrical Engineering Journal
April, 1957.

Ref. 7. Eldridge, D., Dorey, P.F.
"A Core-Transistor Logical Element"
Proc. I.E.E. Paper No. 2347R July, 1957.
Vol. 104 B Supplement No. 7 p.512.

Ref. 8. "The Design of Logical Circuits using Transistor and Square-Loop Ferrite Cores.
Mullard Technical Communication
Vol. 4. No. 34 p 110.

Ref. 9. Eldridge, D.
"A new High-Speed Digital Technique for Computer use".
Proc. I.E.E. Paper No. 2672 Vol. 106 Pt. B. 1959.

THE FUNDAMENTAL PRINCIPLES OF TRANSISTOR CIRCUITS

LECTURE 8. TRANSISTORS DRIVING LARGE MATRIX STORES

by D.A.H. Brown

1. SUMMARY

The lecture begins with a brief reminder of the binary system and the requirements of a computer store. The properties of square loop ferrite cores are then described before considering how some 10^5 such cores can be arranged to form a store. The requirements and methods for driving this large array by transistors are then given. The amplifier makes use of the low core output impedance to drive a common base stage. Finally the extension of the driving system to n dimensions is considered.

2. THE BINARY SYSTEM OF NUMBERS

As computers frequently use the binary system of numbers a brief reminder of this system is given here for completeness.

The decimal number 256 is understood to mean

$$2 \times 10^2 + 5 \times 10^1 + 6 \times 10^0$$

where the 2 is the most significant digit and is by convention written on the left, while the 6, the least significant digit, is written on the right. There are ten digits in this scale of ten, the numbers 0, 1, 2 .... 9.

In exactly the same way in the binary or scale of two system the number 10011 means

$$1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 19 \text{ (decimal)}$$

There are only two digits in this system, 0 and 1.

3. A COMPUTER STORE

We assume that the store stores the numbers in binary form. (The distinction between numbers, instructions and words is deliberately avoided here.) If each number is, say, 40 digits long, then 40 two state elements are needed to store it. The two state elements can take many forms. A bistable circuit (Eccles - Jordan, sometimes called a "toggle") is one form. The two directions of current round a superconducting ring or flux round a magnetic toroid are others. If then the store is to have capacity for 1000 numbers, a total of 40,000 elements is required. Each element represents one "bit" of information and a store capacity of $10^5 - 10^6$ bits is now fast becoming conventional.

Clearly each element must be cheap if the store cost is not to be prohibitive. Ferrite cores are about 1/- each and a store of 10^5 bits works out at about £5,000, excluding the cost of the associated electronics which is of the same magnitude.

4. PROPERTIES OF FERRITE CORES WITH RECTANGULAR HYSTERESIS LOOP

4.1 General Features. The ferrite cores used in computer stores are small rings. The larger type is 2mm O.D. x 1.25 mm. I.D. x .67 mm. thick. The latest and smaller size is 1.25 O.D. x .75 I.D. x .37 thick (dimensions in mms.) The ferrite material has a hysteresis loop which, ideally, would be rectangular (Fig.1a) but in practice it is somewhat rounded (Fig.1b). The loop is characterised by a saturation flux density B_s, a remanence B_r, and a coercivity H_c. For practical materials $B_r > 0.9\ B_s$.

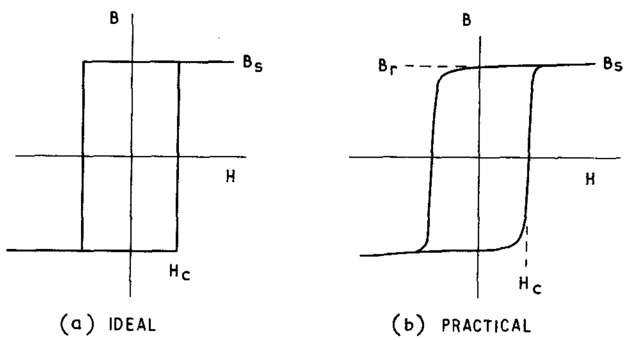


Figure 1 Hysteresis Loops of Rectangular Loop Material

4.2 The Storage Property. The storage property of these cores arises from the two possible states of remanence, positive or negative, and it is usual to designate, arbitrarily, one of these as the "0" state and the other as the "1" state.

4.3 The "AND" gate (Coincidence detector) Property.

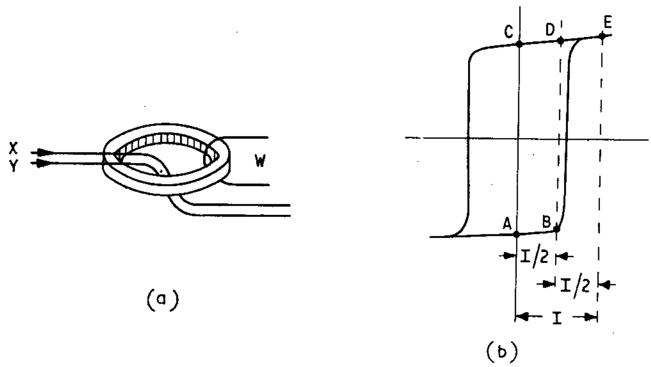


Figure 2 Coincident Current Working

Consider Fig.2(a) which shows a core threaded by three wires. X and Y are drive wires through which currents are passed, and W is an output wire. Suppose that the core is in the remanent state represented by A (Fig.2b). A current I/2 is now passed through X alone, the magnitude of I/2 being rather less than the m.m.f. corresponding to H_c, and its direction being to drive the core away from saturation. Because of the square loop characteristic this current drives the core to B only, but does not permanently alter the magnetic state. When the current is stopped the core returns to A.

If now currents of I/2 are passed simultaneously down both X and Y the m.m.f. is twice as great and this current exceeds H_c; the core now changes state, to point E and when the currents are removed is left at C. Therefore with the currents suitably adjusted the core is a detector of coincidence for currents in both wires or, in the language of computer logic circuitry, it forms an "AND" gate for currents in X and Y.

If the core stared at C, then the half level current would drive it to D and the full current to E. In both cases it will return to C when the current is removed. (This is an oversimplified statement).

4.4. <u>Voltage Outputs</u>. On the hysteresis loop of Fig.2b we can distinguish four different outputs. When the core switches from A to E on the application of a full current a large output pulse is produced and conventionally, this is called the "1" output. Driving the core from C to E gives a smaller output, normally called the "0" output. Half level currents give even smaller signals corresponding to the AB or CD transitions; these are half one (h1) or half zero (hz) signals.

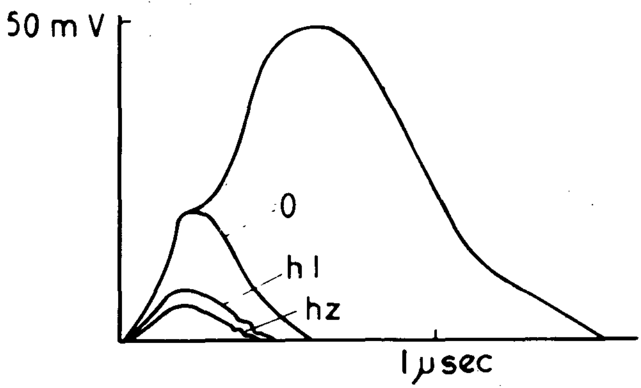


Figure 3 Voltage Outputs From Core FX 1899

These are shown superimposed in Fig.3. The 0, h1 and hz are not only of smaller amplitude than the 1, but are also over quicker. It is common practice therefore to increase the 1/0 discrimination ratio by strobing at the peak of the 1 signal when the amplitudes of the other signals have decayed to an almost negligible value.

5. THE COINCIDENT CURRENT TYPE OF STORE

5.1 The Arrangement of the Ferrite Cores

As a definite example we take the store of the projected transistor computer at R.R.E., the R.R.E.A.C.; this is to hold 4096 "numbers" of 36 binary digits each and therefore uses 4096 x 36 = 147456 cores. 4096 is 64^2.

The simplest arrangement of these cores is as 4096 lines, side by side, each of 36 cores. This arrangement has certain advantages but it requires a selection system to select one out of 4096 and this, though it can be done, needs a large amount of equipment.

A system needing less equipment arranges the 4096 "numbers" in a square array of side 64 and any number can then be identified by its X and Y coordinates. This requires two selection systems, each selecting one out of 64 wires, which is much easier than selecting one out of 4096 directly. This basic idea leads to the coincident current store.

For this type of store the cores are arranged in the form of a rectangular box with a square XY end, 64 x 64, and 36 sheets or planes deep in the Z direction (Fig.4.). The corresponding X lines of each plane are joined in series as shown in Fig.5, and similarly the Y lines are joined in series.

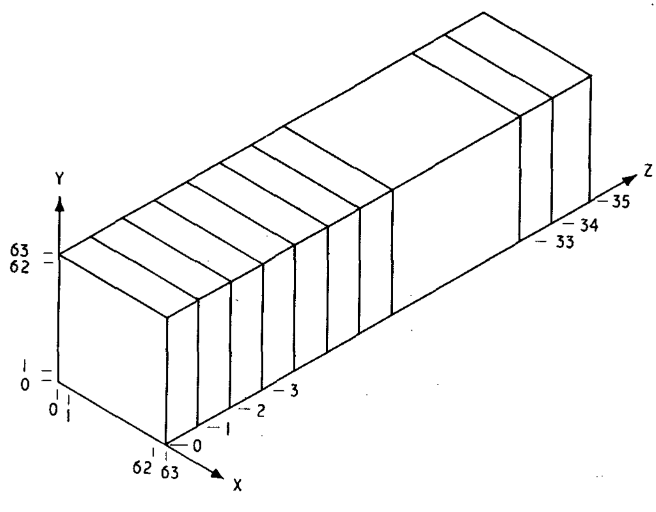


Figure 4 Arrangement of Cores in Matrix Stack

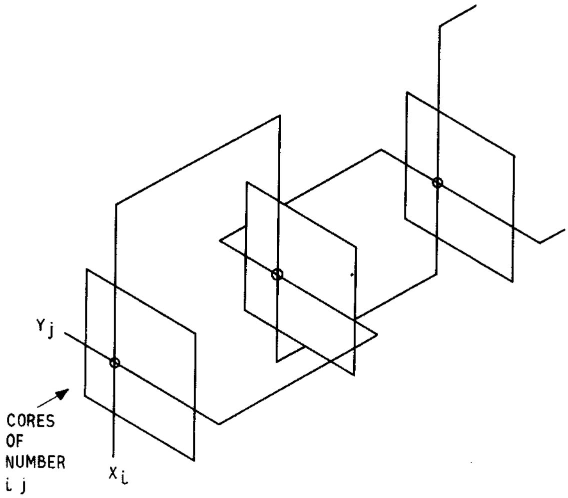


Figure 5 Series Wiring of X and Y Lines

Each plane corresponds to one digit position; for example the Z = 0 plane contains the 2^0 digit of all the 4096 numbers, the Z = 27 plane the 2^{27} digit of all the numbers and so on.

In addition to X and Y wires each plane also has two other wires, the Inhibit (sometimes called Digit) wire, and an output wire. The purpose of the Inhibit will be made clear later.

5.2 The Operation of a Coincident Current Store

5.2.1. Reading

To read the number given by X = i and Y = j half level currents are sent along wires Xi and Yj. In each plane the core at the coordinate ij will receive the two half level currents in coincidence. If the core was previously in the "1" state it will switch and change state giving a "1" input; if it was previously in the 0 state it will not switch and will give the smaller "O" output. Notice that the operation of reading the information destroys that information, because a core starting in the "1" state at A (Fig.2b) is left at C after switching.

The output pulse is first amplified, then applied to a discriminator to distinguish 0 and 1 pulses, and the output from this used to set up a toggle which can then hold the 0 or 1 state indefinitely. 36 such toggles, one for each digit, form the Store Register.

5.2.2. Writing

If the information stored is not to be progressively lost as it is read (and thereby lost from the cores) it is necessary to follow each read by a write operation. The information cannot be rewritten simply by reversing the direction of the X_i and Y_j currents because this would set all the cores at ij back to "1", whereas the binary number may contain both 0's and 1's. An additional current is necessary and this is supplied by the inhibit wire; this wire threads all the 4096 cores of one digit plane and the current in it is always in what may be called the "read" direction, and is another half magnitude I/2 current equal to the X and Y currents.

The rewrite operation then uses both X and Y currents and, conditionally the Inhibit current. The X_i and Y_j currents are in the write direction (opposite to read). If there is no inhibit current in any plane then the ij core is switched back to the "1" state. If it is required to leave a core at 0, then the inhibit current is used; this, being in the read direction, opposes one of the writing half currents and the net effect all three is I/2 in the write direction. This will not switch the core over, leaving it at "0".

6. X and Y Selection Systems

6.1 Driving Specification

We are now able to consider the driving requirements for the store. The nominal full driving current for small ferrite cores (FX 1899 - 50 thou. OD) is 440mA giving I/2 = 220mA. The core specifications are usually given in terms of rise times (10 to 90%) for this current of better than 0.2 μsec.

On each X or Y line there are 2304 cores and, although 36 of these occur at the wanted XY intersection, the other 2268 are disturbed by the half level currents. Although the disturb signals are only about 5mV, the total back e.m.f. of all these cores is about 11V. In addition, the rate of current rise is sufficiently large for the inductance of the wiring to be important and this contributes another 9V of back e.m.f. The IR drop of the wiring is about 3V, and the 36 selected cores may, if they all switch at once, produce 1.5V of back e.m.f.

The driving system must, therefore, produce 220mA with a rise time < 0.2 μsec. into an inductive load capable of generating about 20V of back e.m.f. The p.r.f. must be ≃ 167kc/s (equivalent to a 6 μsec. read/write cycle). As the back e.m.f. spike is a short transistory one the generator is supplying the current at full voltage during most of the pulse. The dissipation is, therefore, about .22 x 20 x 1/4 ≃ 1.1W. (for a duty cycle of $\frac{1}{4}$: - 1.5 μsecs. in 6).

While it is always dangerous to state categorically that this specification cannot be met, with presently available transistors, certainly it cannot using the only transistor readily available of suitable type the Mullard OC23. The Philco 2N601 is not readily available and in any case its maximum dissipation is only 750mW with heat sink. The Texas n-p-n silicon transistor 2S017 is difficult to use because of its high collector saturation resistance, and high and variable base/emitter resistance.

6.2 The Effects of Relaxing the Driving Specification

The Mullard OC 23 has a guaranteed rise time in a certain test circuit of < 0.5 μsec; this is to 350mA, not 220. The nominal cut off frequency is 2.5 Mc/s. The rise times obtained in practice are usually about 0.35 to 0.4 μsec, and this is an increase of about 0.2 μsec. over the core specification rise time of 0.2 μsec.

This slower rise delays the peak of the 1 signal by approximately 0.2 μsec. and reduces its amplitude 5 or 10%. The 0 signal is also lengthened and reduced in amplitude. The disturb signals are reduced in amplitude almost in proportion to the rise time so that, because the extra 0.2 μsec. represents a doubling of the rise time, they are halved; this is because the disturb signals are caused by a reversible flux change which is equivalent to an inductance L_r. The disturb signal may be regarded as L_r di/dt, so that if di/dt is halved so is the signal.

The slower rise also reduces the Ldi/dt term caused by the matrix wiring.

The overall effect is to reduce the back e.m.f. on the drive wires to about 10 to 12 V, to delay the signal and not to effect greatly the 1/0 discrimination ratio.

6.3 The Basic Matrix Scheme

The simplest X or Y coordinate selection system consists of a pulse current generator feeding 64 switches in parallel. The switches would in fact be transistors and one only would be closed at any one time to direct the current down the appropriate lines. A scheme of this type has been used in the EMIDEC 2,400.

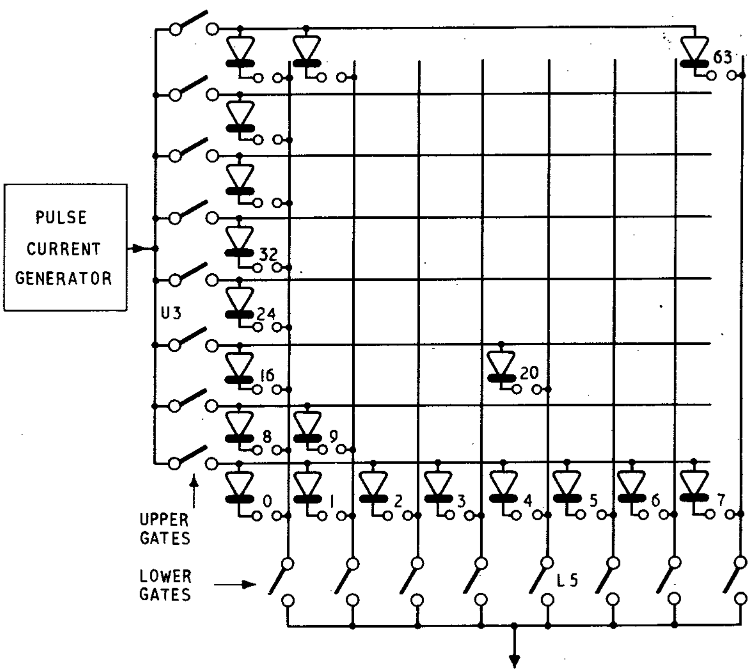


Figure 6 8 x 8 Selection System for One of 64 Wires

A rather more complex scheme but needing less equipment is an 8 x 8 selection system, shown schematically in Fig. 6. One upper gate switch and one lower gate switch is closed at any one time and the current is thereby constrained to flow down one wire only; for example if U3 and L5 are closed current flows down wire 20 only. The diodes in series with each wire are to prevent other wires being in parallel with the chosen one, causing the current to split between parallel paths.

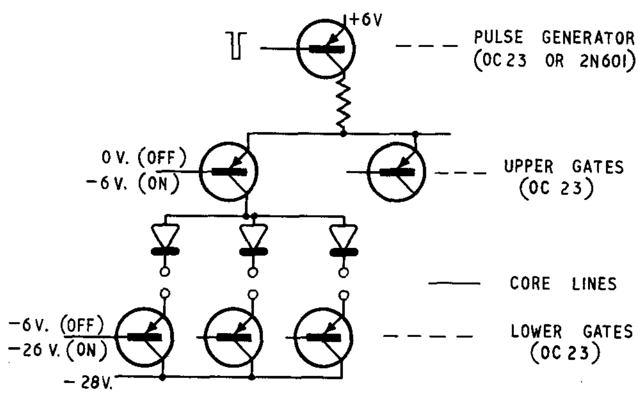


Figure 7 Basic Engineering of 8 x 8 Selection Scheme

The basic engineering of Fig. 6 is given in Fig. 7. The pulse generator upper and lower gates are all in series, with the output wires and blocking diodes between upper and lower gates; even so, the overall voltage is only 34V which compares with 720V (+420V to -300V) in a corresponding valve system (Quartly, Cain and Clarke - Ferrites Convention - Proc. IEE 104B Supplement No. 7 p. 457 1957). The normal (off) position of the gates is 0V (upper) and -6V (lower). A wire is selected by altering upper and lower gate levels to -6V and -26V respectively. The pulse generator is now turned on for a 2 μsec. current pulse. The resistor R defines the current; the bottom end of R will be held at -6V by the emitter of the on transistor while the top end will be near +6V if the pulse generator bottoms, so that the current is defined as 12/R. (Some refinements must be obviously made to take account of base emitter drops etc. when considering the exact value of R and its tolerance).

It will be noticed that the movement of the lower gate bases is 20V. This is to accommodate the large back e.m.f. on a line when the current is turned on. The lower end of the selected wire will be held at -26V but, for a back e.m.f. of 12V the upper end will rise, as a transient spike, to -14V. The connections of Fig. 7. show that this will drag up the emitters of all the lower gates and these must remain off during the transient. The 20V movement is then about the maximum which can be accommodated within the maximum collector-emitter voltage specification of the OC 23 (=24V) allowing for some tolerance on voltage lines and avoiding bottoming during the pulse.

6.4 Circuits

The actual circuits of the pulse generator, upper and lower gates are given in Fig. 8 (Pulse Generator), Fig. 9 (Upper Gate) and Fig. 10 (Lower Gate). Each can only be described as conventional.

Two features may be noted particularly. First, the use of emitter followers using the non/pnp combination to give current gain in both edges. This is a well known transistor circuit which has no valve equivalent. Secondly the fairly wholesale use of circuits which prevent bottoming. The design philoscophy of these circuits starts with a given collector current i (Fig.11a), and a minimum $\alpha 1$ (usually taken as 20 for OC 44). Fig. 11a then shows the currents in the on condition. About 0.5V is dropped across R_2 but the exact voltage will vary with the transistor α^1 because of the variation in base current; the variation is reduced by bleeding a larger current (5 times the maximum base current) down R_1; this fixes R_1. During turn on (Fig.11b) the diode current is available as additional base current until the collector potential rises to almost its final value; the value of 4i/20 is a reasonable excess giving a drive to the base of one fifth of the required collector current. This fixes R_3. Turn off could be just accomplished (at nominal values with no tolerance) by supplying the base with a current of 5i/20 to stop the transistor base current and the current in the diode. A current of twice this amount leads to conditions as in Fig. 11c during turn off, which is then quite rapid.

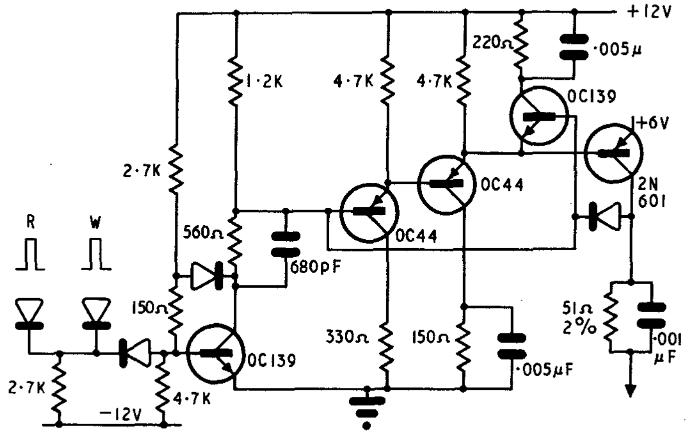


Figure 8 Pulse Generator

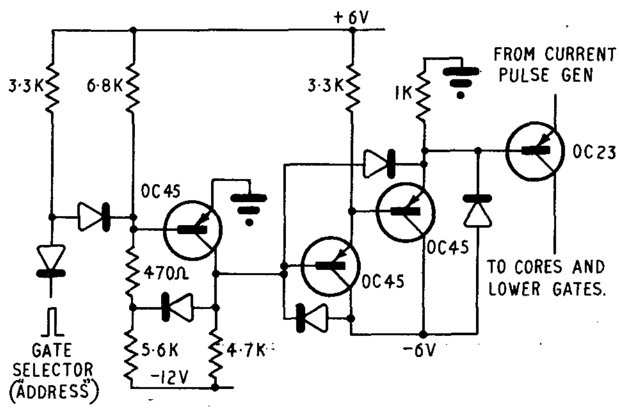


Figure 9 Upper Gate

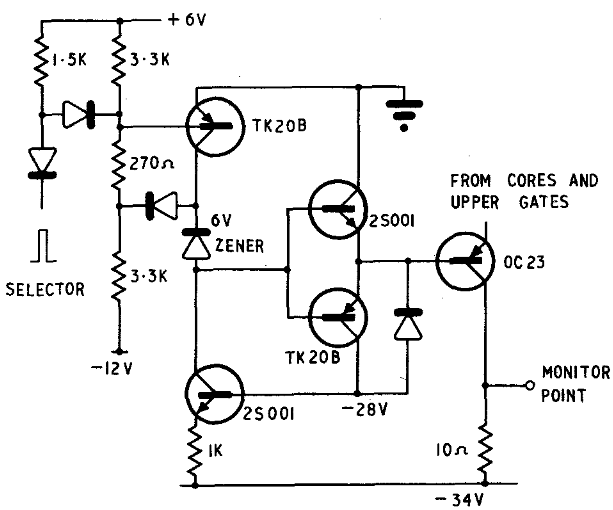


Figure 10 Lower Gate

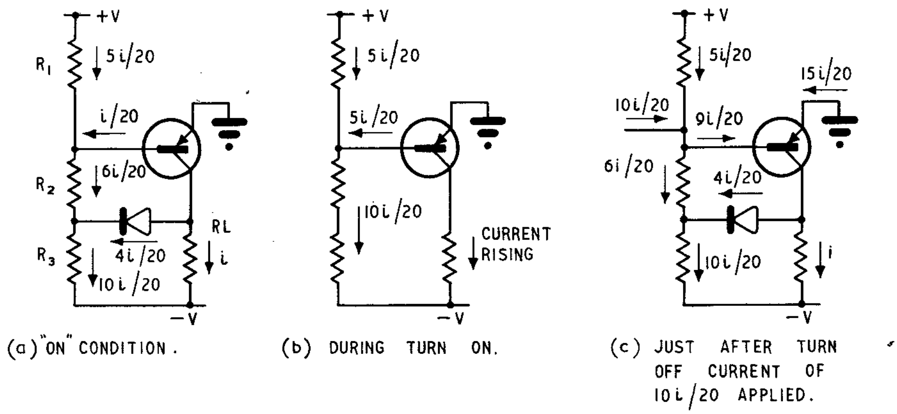
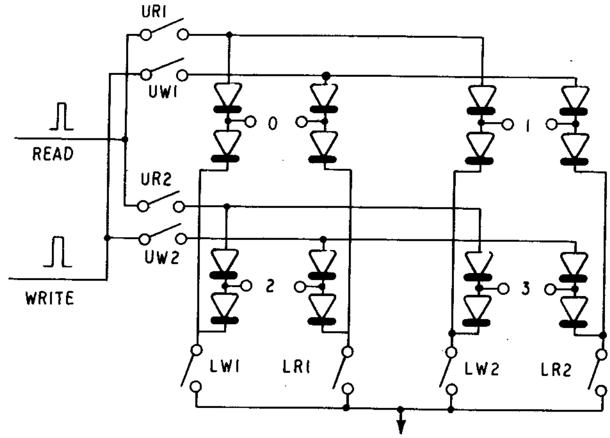


Figure 11

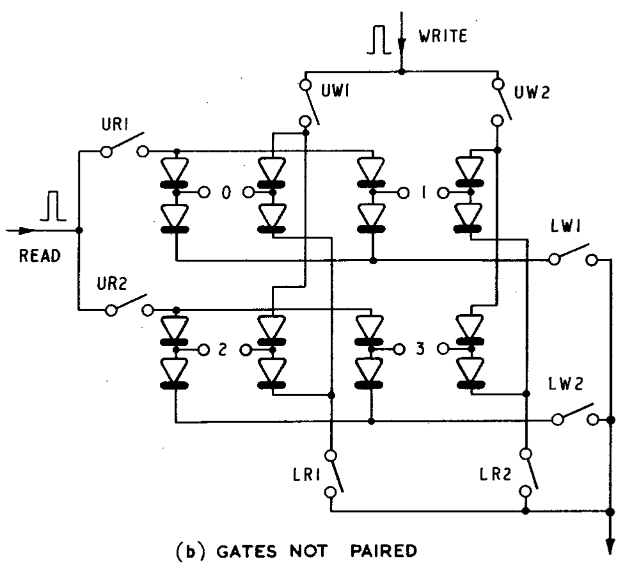
6.5 Read and Write on One Wire

The scheme of Fig.6 will only send current one way down each wire. If it is possible to have two separate wires for read and write on X and Y coordinates, then four such systems, being read and write for X and Y, could be used (though some compression is possible to save gate switches).

However, in large core matrix stacks only one wire is available for each coordinate and this must be used both for read and write, so that it must be possible to send current both ways along one wire. One way of doing this is to drive the wire from a transformer with two primaries, and this has been used (EMIDEC 2400).



(a) GATES PAIRED



(b) GATES NOT PAIRED

Figure 12 Four Diodes per Wire

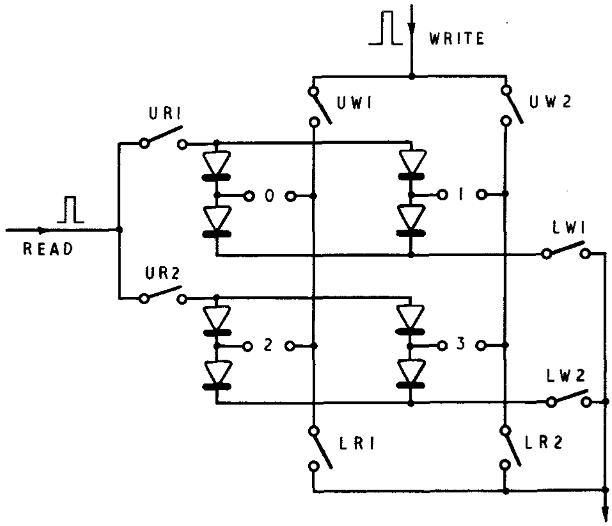


Figure 13 Two Diodes per Wire
(Same as Fig.12b with R.H. pair of diodes
of each set shorted out).

An alternative system uses read/write gates in pairs in both upper and lower gate positions. Extra diodes are needed in association with each wire. Four diodes per wire are to be used in the R.R.E. system (Fig.12a) but Mullards have recently shown an improved system using only two diodes per wire (Fig.13). Both Fig.12 and Fig.13 demonstrate the principle applied to a 2 x 2 system. In Fig.12a wire 0 has a read current by closing UR1 and LR1, and the current direction is reversed for write by closing UW1 and LW1. The read and write generators supplying the pulse currents can, in fact, be one generator pulsed twice and in this case all the upper (u) gates are in parallel. In Fig. 12b and 13 the operation is best explained by a table.

Wire	Switches closed for	
	Read	Write
0	UR1, LR1	UW1, LW1
1	UR1, LR2	UW2, LW1
2	UR2, LR1	UW2, LW2
3	UR2, LR2	UW2, LW2

As in Fig.12 the read and write generator can be combined if desired.

In both cases the diodes used are OA5 for the smaller size of core requiring only 220mA, and OA10 for the larger size of core requiring 350mA.

7. The Inhibit Generator

The Inhibit current disturbs all the 4096 cores of one digit plane and its design is therefore, dominated by the back e.m.f. generated. If the current pulse rise time were to be as fast as the 0.2 µsec. of the core specification the inductive back e.m.f. of the cores would be about 20V. To this must be added the wiring inductive e.m.f. of about 9V, plus IR drop.

In fact the inhibit rise time need not be very fast, provided the inhibit current is fully established before the writing currents are applied. The simplest method then is to allow the inhibit generator to bottom and to determine the current by a resistor in the collector in series with the inhibit wire. Taking a collector supply of 22V and 220mA as the required current R = 100 ohm. The cores are equivalent to an inductance of about 20 µH and the wiring adds another 5 µH, making 25 µH. The rise time constant is therefore, determined mainly by the circuit as 0.25 µS (25 µH/100 ohm) and the current should take nearly one microsecond to establish itself. In fact the figure of 20 µH for the equivalent core inductance is based on the maximum disturb signal expected. The average disturb signal is less than this and, with 4096 cores, the average disturb is probably a more useful figure. This appears to be about half the maximum and the current rise time constant is correspondingly less.

An alternative scheme is to determine the current other than in the collector circuit, and to provide as far as possible true current drive to the inhibit wire. This is the present R.R.E. approach which uses an OC23 with current determined by a series emitter resistance.

8. The Amplifier

If a true voltage source (Zero internal impedance) is applied between base and emitter of a transistor then one may consider that it produces a current of either $i_e = v/[r_e + r_b(1-\alpha)]$ in the emitter circuit or of $i_b = v/[r_b + r_e/(1-\alpha)]$ in the base. In both cases the input impedance is normally dominated by the r_e term and a first approximation becomes $i_e = v/r_e$ or $i_b = v(1-\alpha)/r_e$. In the second case the emitter current produced by i_b is $i_b/(1-\alpha) = v/r_e$, so that the two approaches lead to identical results. This must be so because it cannot matter which end of the source generator is earthed. However, in many practical cases the source impedance is much greater than r_e; then the common base connection puts the source impedance in the emitter circuit where it automatically becomes a feedback impedance, and reduces the gain. But the core output impedance is very low ($\sim 0.5\Omega$) and the resistance of the amplifier wire in the matrix stack is only 11Ω, so that the common base connection can be used. This was pointed out by Mr. R.C. Bowes. With a collector load of R_L the gain is then $R_L/(r_e + R_e)$ where R_e is the source resistance plus any additional resistance in the emitter circuit.

The amplifier must be able to handle signals of both polarity. Fig.14 shows the first stage using a "Bowes brick". The input transistor works at 2.5mA giving $r_e = 10$ ohms. The observed performance is as expected; 50mV input signals give an output of rather more than 2V.

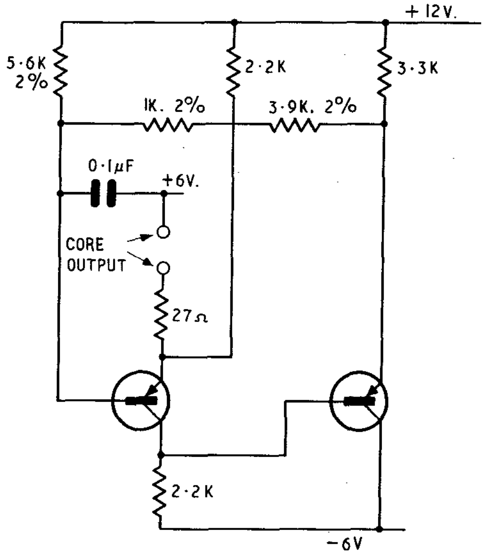


Figure 14 Amplifier Using Common Base Circuit

The more common input arrangement uses a large step up transformer on the core output. This is satisfactory at lower p.r.f.'s but leads to difficulty as the read/write cycle time is reduced. The difficulty is the common one of the loss of d.c. level in an a.c. coupling (the transformer). The pulse length of the core output "1" signal is 1.5 μsec. and if the droop on this pulse is to be small the L/R time constant of the transformer circuit must be several times longer. On the other hand the time constant should be several times shorter than the pulse repetition period if the magnetising current built up by one pulse is to have disappeared before the next pulse. With a read/write cycle of 6 μsec. it is impossible to fulfil both these conditions with a 1.5 μsec. pulse. The problem arises because the read and write signals are not necessarily balanced, for it is possible to read out a "1" and rewrite "0" and vice versa; this occurs when the store content is changed.

9. Extension of Coordinate Selection to n Dimensions

The hysteresis loop of practical "square loop" ferrite materials is not sufficiently square to allow reliable selection by more than two coincident currents. In principle, given a perfectly square loop material capable of distinguishing between $(n-1)I/n$ and I (where I is the current just sufficient to switch the core), n coordinate selection could be used. The advantage gained is a reduction in the number of access lines. For example 4096 lines require selection of one out of 4096 selectors in a one coordinate system. In the two coordinate XY coincident current system there are only $2 \times 64 = 128$ access lines; this is $2 \sqrt[2]{4096}$. In n coordinates only $n \sqrt[n]{4096}$ would be required; for $n = 4$ this is $4 \times 8 = 32$ only. The wiring of a four dimensional system is quite possible. Take the 4096 cores and split them into 8 blocks, $(8 = \sqrt[4]{4096})$ each of $4096/8 = 512$ cores. In each block of 512 $(= 8^3)$ one core can be selected by its X and Y and Z coordinate there being 8 lines per coordinate, with corresponding coordinates in each block all in series. This distinguishes 8 cores, having the same XYZ coordinates in each block. The unique core is distinguished by its W or block coordinate, a wire which would thread every core in one block.

10. Additional Notes.

(These have been added since writing the lecture).

NOTE 1. The complete matrix stack (64 x 64 x 36 bits) has been delivered and experience with this has led to some modifications.

NOTE 2. (Concerning the Inhibit Generator, Section 7).
It now appears that the method of the second paragraph of this section (inhibit generator bottomed, current determined by resistor in collector circuit in series with inhibit wire) is quite satisfactory. The power dissipation in the driving transistor is less by this method and it is therefore being used in preference to the method of the third paragraph of the section.

NOTE 3. (Concerning the Amplifier, Section 8).
The single ended form of the amplifier has now been replaced by a differential (push-pull) type which should be less susceptible to direct capacitive pick-up in the matrix stack.