WIRELESS ENGINEER

Vol. XXV SEPTEMBER 1948 No. 300

Cullwick's Electromagnetic Problems

IN the Editorial of November 1947 we discussed an electromagnetic conundrum which had been propounded by E. G. Cullwick in the American journal *Electrical Engineering* and gave what we considered the obvious solution. We have since been asked, however, if we have made calculations to confirm the correctness of the solution, and, judging from a letter from Cullwick in *Nature* of 19 June, he still regards the conundrum as unsolved. On looking further into the matter we are more than ever convinced that the solution given in November is the correct one. As Cullwick further suggests that there is some difficulty in reconciling the ordinary solution of the problem with the conservation of energy we propose to look also into this aspect of the question.

In Fig. 1, the charge $+Q$ is moving with the velocity v past a closed conducting ring C. The approaching charge produces an increasing magnetic flux through the ring, which reaches a maximum when Q is nearest to C and then decreases as it recedes. This varying magnetic flux induces in the ring an electromotive force in one direction as the charge approaches and in the opposite direction as it recedes. In Fig. 2 the charge $+Q$ is at rest and the conducting ring is moving with the velocity v as indicated. The relative motion of the charge and ring is the same as before but the ring is now moving through the electrostatic field of Q, and the line integral of this field around the ring is zero. Is there then an e.m.f. in one case and not in the other? No, there is the same e.m.f. in both cases. The charge $+Q$ induces a charge $-q$ on the part of the ring nearer to it and an equal charge $+q$ on the more distant part. These induced charges increase as the ring approaches Q and then decrease as it recedes. Although Q is at rest, $+q$ and $-q$ are moving with velocity v and produce a magnetic field in the space between them. The exact distribution of the charges over the surface of the ring and the determination of the resulting magnetic field at various points would involve a prohibitive amount of mathematics, but it is not difficult to obtain an approximation which will provide a rough comparison between the two cases.

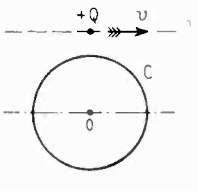


Fig. 1

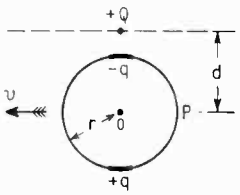


Fig. 2

Instead of being distributed over the whole ring we assume that $+q$ and $-q$ are concentrated more or less at the two extreme points of the ring as shown in Fig. 2. Their magnitudes must be such that their electric field strength exactly counteracts that due to Q. In reality they are distributed in such a way that this cancellation holds at all points of the ring, but when concentrated as shown in Fig. 2 we can determine the values of $+q$ and $-q$ which give zero resultant electric field at the centre of the ring. It makes little difference whether we assume the charges to be concentrated at points or spread over short arcs, and whether we assume the fields to be cancelled at the centre O or at the point P.

If the minimum distance from Q to the centre of the ring is d and the radius of the ring r we

have $\mathcal{E}$ at O due to $Q = Q/d^2$

$\mathcal{E}$ at O due to $+q$ and $-q = -2q/r^2$

If the resultant $\mathcal{E}$ is zero, $q = \frac{Q}{2} \cdot \frac{r^2}{d^2}$

Now $+q$ and $-q$ are moving to the left with a velocity v and their magnetic fields through the ring are in the same direction, viz., away from the reader as in Fig. 1 The value of H at the centre of the ring will be $2\,qv/cr^2$ and if we substitute for q the value just determined we have $H = Qv/cd^2$ which is exactly the value of H at the centre of the ring in Fig. 1 where the coil is stationary and the charge Q moving. Hence we can safely state that the magnetic field reaches the same maximum value in both cases, and there is no doubt that the e.m.f. induced in the ring is the same in both cases. It may be suggested that the induced charge $-q$ is not moving with the velocity v on account of its movement around the ring, but Fig. 3 shows that its shortcomings are counterbalanced by the extra movement of the charge $+q$.

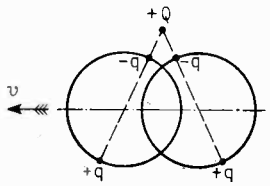


Fig. 3.

Up to this point we have not considered the effect of any current flowing in the ring; as in the November Editorial, we have assumed the ring to be open-circuited. We have also neglected any charge distribution set up by the e.m.f. induced by the changing magnetic flux, which would be superposed upon the electrostatic distribution.

We now turn to what, at first sight, appears to be a more difficult problem. The electromotive force induced in the ring in Fig. 1 will set up a current and power will be dissipated at the rate of i^2R watts where i is the momentary current and R the resistance of the ring. Where does this power come from? The only energy available is that of the moving charge and the only way to reduce its energy is to reduce its speed. The current in the coil produces a magnetic field at right angles to the line of flight of the charge, but such a magnetic field can only exert a force at right angles to the line of flight and is therefore incapable of reducing its speed. The principle of the conservation of energy demands that the speed of the moving charge be reduced, and we have to explain the mechanism whereby this is effected. A moving charge carries with it not only a magnetic field but also an electric field. The magnetic field produced by the currents in the coil is not a steady field but a variable one and variable magnetic fields produce electric forces. These facts indicate the importance of looking into the electric fields before jumping to the conclusion that the principle of the conservation of energy is being violated.

In Fig. 4 PG represents a pulse generator producing pulses of unidirectional current around the rectangular circuit within which is a closed metallic ring. The pulses I in the primary circuit produce pulses i in the closed secondary circuit and the energy dissipated in the ring is supplied by the pulse generator. The rectangular circuit and the ring constitute merely an air-core transformer. Without going into details we can say that the current I flows against a back electromotive force due to the current i, or in other words, the electrons moving along the straight wire experience an opposing force due to an electric field caused by the presence of the closed ring. Nobody would suggest that there is anything here contrary to the principle of the conservation of energy.

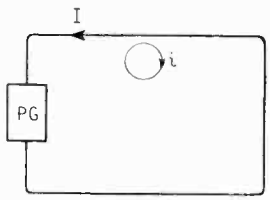


Fig. 4.

In Fig. 5 the upper side of the rectangle has been replaced by a cathode-ray tube, within which we may imagine the closed ring to be placed. Instead of a pulse of slowly moving electrons we now have single electrons moving at high speed, but there is no fundamental electromagnetic change, and the ring will still produce an electric field opposing the motion of the moving charge, thus slowing it down and reducing its energy.

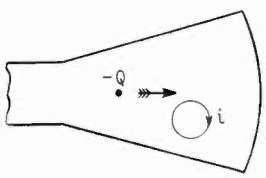


Fig. 5

This treatment of the problem has the great advantage that it is based on simple classical theory and does not invoke the aid of Lorentz, Ritz or Einstein.

G. W. O. H.

SECONDARY ELECTRON SUPPRESSION

New Effect at Large Transit Angles

By J. H. Owen Harries, A.M.I.E.E., M.I.R.E.

1. The Problem

In microwave valves energy is withdrawn from a modulated beam of electrons by passing the electrons through an alternating electric field. This output field is positioned, as a rule, in a gap between the walls of a resonant cavity. The electrons enter the gap through a hole in one of the walls[1]. If, after passing across the gap and through the field, the electrons strike the other wall, then secondary electrons will be emitted from that wall. These electrons will travel into the field and may withdraw energy from it while in flight and so decrease the power output efficiency. Indeed, the efficiency may become almost zero if, as is usual, the walls of the resonator on each side of the gap are maintained at the same steady potential.

The problem is an extension of the familiar one of trying to suppress secondary electrons emitted from the anode of an audio-frequency power tetrode and preventing their flow to the screening grid. At longer wavelengths than about a metre or so, a 'suppressor grid,' which is positioned between the screen grid and anode, is fairly successful in preventing the flow [Fig. 1(a)] but at fairly short wavelengths, lead inductances and other difficulties make the suppressor grid much less efficient. At the shorter micro-wavelengths it appears to be of no real use at all. The well-known space-charge effect found at the critical anode distance[2] [Fig. 1(b)] prevents the flow of secondary electrons down to a much shorter wavelength than does a suppressor grid; but even the critical distance method fails if the wavelength is short enough, because the necessary electrode spacing is too large to be used in appropriate cavity resonators.

Therefore, the designers of microwave valves have usually dealt with the difficulty of avoiding the effects of secondary electrons not by suppressing them, but by causing the electron beam to pass through the output field in the gap without striking the walls of the resonator so that secondary electrons are not emitted from the walls. The electron beam is caught by a 'bucket'* which is outside the resonant cavity, [Fig. 1(c)]. This artifice is effective, but it has mechanical disadvantages and seems only to be applicable to velocity-modulated beams and control-grid valves. It is difficult to see how this idea can be applied to the provision of, for example, two push-pull output fields positioned side by side in a cavity as is necessary in a deflection valve.

A problem therefore exists of devising means to suppress the flow of secondary electrons if they are produced in output fields in cavity valves.

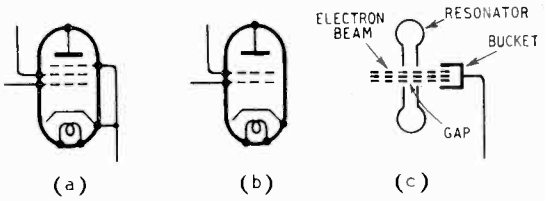


Fig. 1. *Known methods of avoiding the undesirable effect of secondary electrons in valves. (a) the suppressor grid; (b) the critical distance; (c) 'inductive output'.*

2. History

Some years ago the writer came to the conclusion that the phase of arrival of the primary electrons at the wall at the end of the gap, in relation to the instantaneous magnitude of the electric field at that wall, might perhaps be adjusted so that the secondary electrons were driven harmlessly back into the resonator wall by the field as soon as they were emitted. Some unpublished observations by the author and E. G. Autie in 1940 at the laboratory of Harries Thermionics Ltd. indicated that this might perhaps be true.

The writer subsequently carried out a research programme on this problem at the Electronics Department of Rediffusion Ltd., London.

3. Properties of Secondary Emission

Secondary-electron emission is very complex. Secondary electrons are emitted in accordance, very roughly, with a cosine distribution as shown in Fig. 2; that is, most of them are emitted in the direction from which the primary electrons arrive. This appears to be largely unaffected by

MS accepted by the Editor, May 1947

* It is sometimes stated that power-output efficiency can be greatly improved by maintaining the bucket at a relatively low potential. This is not true unless the efficiency of power transfer from the electron beam to the field in the gap is low in the first place. This is, for example, inherently the case if velocity modulation is used.

the angle of incidence of the primary electrons on to the emitting electrode. The number of secondary electrons emitted per primary electron varies very widely with different metals and different surface conditions and with the velocity of impact of the primary electrons (e.g., anode voltage). Secondary electrons are emitted (at the impact velocities used in radio valves) with a continuous spectrum of emission velocities from a peak which is equal to the impact velocity down almost to zero velocity. This is indicated in Fig. 3. It is not correct to look upon secondary-electron radiation in valves as consisting only of electrons of very low velocity (or energy) compared with the impact velocity (or energy) of the primary electrons.

The author carried out a preliminary survey[3] of the published knowledge of secondary-electron radiation as a preliminary to an attempt to analyse the problem set out above. The phenomenon was found to be far too complex, and too little quantitative information had been published about it, for a rigorous mathematical analysis to be attempted. It was felt however, that some useful information might be obtained by computing the energy relations, even though very drastic simplifying assumptions had to be made about the secondary electrons themselves. The problem is mathematically tractable if it is assumed (contrary to the facts) that all the secondary electrons are emitted in the direction from which the primary electrons arrive, that the number of secondary electrons emitted per primary electron is a constant, and, finally, that all the secondary electrons are emitted with velocities tending to zero.

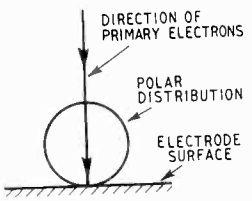


Fig. 2. *Approximate spatial distribution of secondary-electron radiation from a conductor.*

4. Mathematical Theory

Calculations based on these assumptions were made by T. S. Popham[4] during the progress of the experiments and are described by him elsewhere. In common with most calculations relating to microwave valves, when the electron transit time is not negligible compared with the periodic time, results of analysis can be expressed in terms of a 'small signal' transit angle

$$\phi = \frac{10^3}{V_b^{\frac{1}{2}}} \cdot \frac{l_1}{\lambda} \pi = 3.3 \times 10^{-8} f \cdot l_1 V_b^{-\frac{1}{2}} \pi \qquad (1)$$

where

$\frac{l_1}{\lambda}$ = the ratio between the length of the gap in cm and the wavelength of operation in cm

V_b = the steady voltage at which the walls at both sides of the gap are maintained.

Another important parameter is a ratio

$$M = \frac{\hat{e}_0}{V_b} \qquad (2)$$

where

$\hat{e}_0$ = the maximum instantaneous value of the alternating electric field integral (voltage) across the gap.

The particular value of M which is necessary (at a given transit angle ϕ) if the efficiency is to be at the theoretical maximum, will be indicated by M_0.

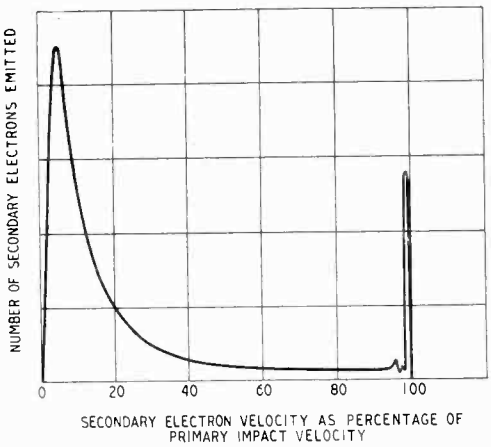


Fig. 3. *Velocity distribution of secondary-electron radiation in terms of the impact velocity of the primary electrons.*

ϕ may be looked upon as the transit angle of an electron under 'small signal' conditions; i.e., when $M \rightarrow 0$. Nevertheless, the relevant energy equations may be expressed in terms of ϕ when M is large.

When the transit angle ϕ tends to zero (low-frequency conditions) and if the electron stream is 100% sinusoidally modulated, then the theoretical maximum efficiency η_0 is equal to 50% and is attained when $M = M_0 = 1$. This will be recognised, for example, as the familiar condition for theoretical maximum efficiency of a class A audio-amplifier.

On microwavelengths, however, the resonant frequency and general electrical properties of useful resonant cavities result in the ratio l_1/λ [equation (1)] being of necessity appreciable in magnitude; therefore the transit angle ϕ cannot tend to zero and is unavoidably an appreciable fraction of π. This has been discussed by the

author in a previous paper.[1] The theoretical maximum efficiency is then reduced in comparison with its value when $\phi \rightarrow 0$.

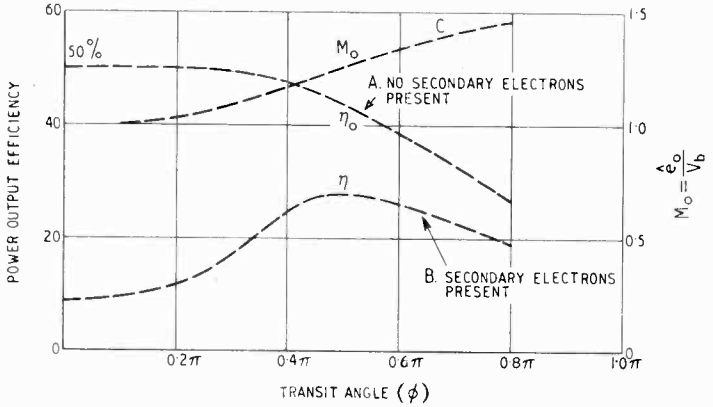
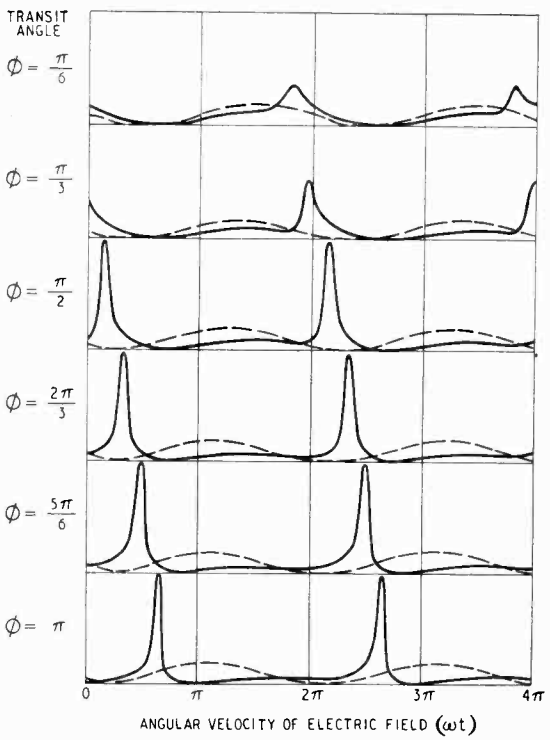


Fig. 4 (left). *Theoretical variations of power output efficiency η_0 and η and 'depth of modulation' M_0 of the electric field as a function of transit angle ϕ.*

Fig. 5 (below). *The entrance flow and exit flow of primary electrons across a gap as a function of angular velocity ωt of the electric field in the gap for various transit angles ϕ. When ωt is between zero and π, or 3π and 4π, the electric field is directed towards the exit wall. – – – – – –Primary electron flow entering the gap; ——— primary electron flow leaving the gap.*



Curve A of Fig. 4 (from T. S. Popham's work) shows the variation of η_0 with ϕ for a sinusoidally 100% modulated electron stream in the absence of secondary electrons. The reduction in η_0 with ϕ is not very serious until ϕ becomes greater than $\pi/2$. The optimum ratio M_0, between the peak alternating voltage $\hat{e}_0$ across the gap and the steady voltage V_b, rises with increases in ϕ (Curve C, Fig. 4).

Most designers of cavity-valves try to reduce the gap length l_1 as much as possible so that ϕ does not exceed at most $\pi/2$.

If secondary electrons are assumed to be knocked out from a wall on which the primary electrons impact and if these secondary electrons are further assumed to behave in accordance with the simplifying assumptions set out above, then Fig. 5 represents the mechanism of the relevant part of T. S. Popham's calculations.

This figure shows plots of the entrance and exit primary electron flows across the gap for various transit angles ϕ. The dotted plots show the flow of primary electrons when they enter the gap; the full-line plots show the flow of primary electrons in the plane of the wall which is assumed to bound the exit end of the gap. All the plots are versus the angular velocity of the electric field in the gap. The direction of this field is parallel to the paths of the primary electrons; that is, normal to the exit wall.

The primary electron stream is assumed to be 100% sinusoidally modulated when it enters the gap (dotted curves); but 'bunches' as it crosses the gap (full-line curves). If the transit angle ϕ exceeds about $\pi/2$ most of the primary electrons arrive at the exit wall in sharply defined 'bunches' (full-line curves). The significant fact is that these bunches arrive at the exit wall, and cause secondary electrons to be emitted from it, when the phase of the electric field across the gap is between zero and π, 2π and 3π and so on; that is, when this field is directed towards the wall, so that the secondary electrons are driven back harmlessly into the wall before they have time adversely to affect the output efficiency. Some of the primary electrons, however, arrive at the exit wall between π and 2π, 3π and 4π and so on, when the field is not towards the exit wall and when it does not tend to drive secondary electrons back into the wall. Only if the number of secondary electrons emitted per primary electron is very low, will these latter secondary electrons cause only a negligible loss of power efficiency.

If, on the other hand, the number of secondary electrons emitted per primary electron is large, these unsuppressed secondary electrons will by themselves be sufficient to reduce the overall power efficiency to a useless figure ; and the fact that some others are suppressed will not usefully affect this disappointing result.

For the purpose of the analysis, it was further and arbitrarily assumed that the ratio between the number of secondary electrons emitted per primary electron had some (unknown) value such that power-output efficiency was proportional to the ratio between the number of primary electrons in the bunches which arrive at the exit wall between zero and π and the total number of primary electrons which arrive per period ; i.e., between zero and 2π. This arbitrary ratio may be computed from the areas enclosed by the full line plots in Fig. 5 and may be seen to increase sharply if the transit angle ϕ exceeds about $\pi/2$. The power-output efficiency in the presence of secondary electrons may be found by multiplying Curve A of Fig. 4 by the arbitrary ratio. The resulting curve (B in Fig. 4) will be seen to possess a maximum. According to this theory, then the undesirable results of secondary-electron emission are reduced over a considerable range of transit angles around about $\pi/2$.

Only a very few out of the bulk of the secondary electrons emitted in practice will actually behave in accordance with the theoretical assumptions. The analysis thus does not apply to most of the secondary electrons because they are emitted with many different speeds and directions and many of them have high velocities comparable with those of the primary electrons. The overall result on the power-output efficiency cannot therefore be predicted, nor can one say, from this theory alone, whether any overall effect at all would be observed in practice. Experiments were therefore performed.

5. Experimental Results

The efficiency η of transfer of power from an approximately sinusoidally 100% modulated electron beam to an electric field in a gap was measured over a range of transit angles ϕ from almost zero to about π in the presence of secondary electrons. The apparatus used at the longer transit angles (higher frequencies) is illustrated in Fig. 6. The power delivered by the beam to the field was measured by the calibrated probe shown. At transit angles tending to zero (low-frequencies) ordinary power-output measuring gear was substituted for the resonant cavity and probe. At all transit angles the modulated electron beam, after passing through a gap of length l_1, impinged on an interchangeable target and knocked out secondary electrons from it. The amount and nature of the secondary-electron emission depended upon which of various available types of target was used.

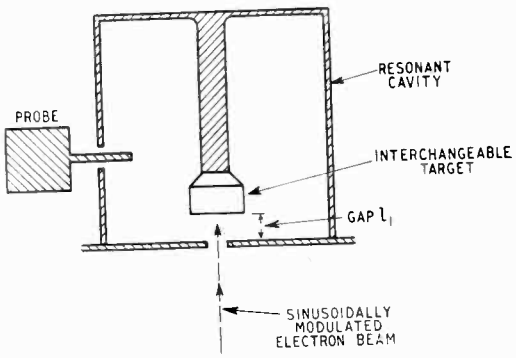


Fig. 6. *The experimental apparatus.*

The target and the entrance wall (through which the electron beam entered the gap) could be maintained at the same steady potential V_b ; or, alternatively, these electrodes could be maintained at different steady potentials. The gap l_1 was variable between about 1 cm and 1.5 cm. A range of frequencies from about 1,600 c/s to 750 Mc/s (40 cm wavelength) was used to give the required ranges of transit angles, [see equation (1)]. The percentage power efficiency at any transit angle is

$$\eta = \frac{P_0}{P_b}\% \qquad (3)$$

where

P_0 = the power delivered to the electric field

P_b = the primary-electron power entering the gap.

The average primary-electron current was measured at low frequencies and from this and the steady potential V_b of the target, P_b was found. P_b is approximately equal to the d.c. power supplied to the target when the voltage V_{s_A} of the entrance wall equals $V_T = V_b$.

Curves A of Fig. 7 show the measured results with a target having a plane polished copper surface. The efficiency has an optimum range and is at a maximum of about 28% at a transit angle of about $\pi/2$. This optimum transit angle of $\pi/2$ is about the same as that of the theoretical curve B of Fig. 4. This agreement is remarkable in view of the drastic assumptions upon which the mathematical analysis is based.

This comparison between the theoretical curve B of Fig. 4 and the measured curves A of Fig. 7(a) and (b) shows that the actual secondary-electron coefficients (at the impact velocities

used) of a commonly-used electrode material, namely polished copper, are in fact low enough for a suppression effect by the electric field to be observable. This is so despite the presence of secondary electrons which, contrary to the simplifying assumptions of the theory, are emitted in many different directions and by no means only in the direction from which the primary electrons arrive. They are not emitted in constant ratio to the number of primary electrons and only very few of them have velocities nearly zero. Most of them, indeed, have emission velocities which are very appreciable compared with the impact velocities of the primary electrons.

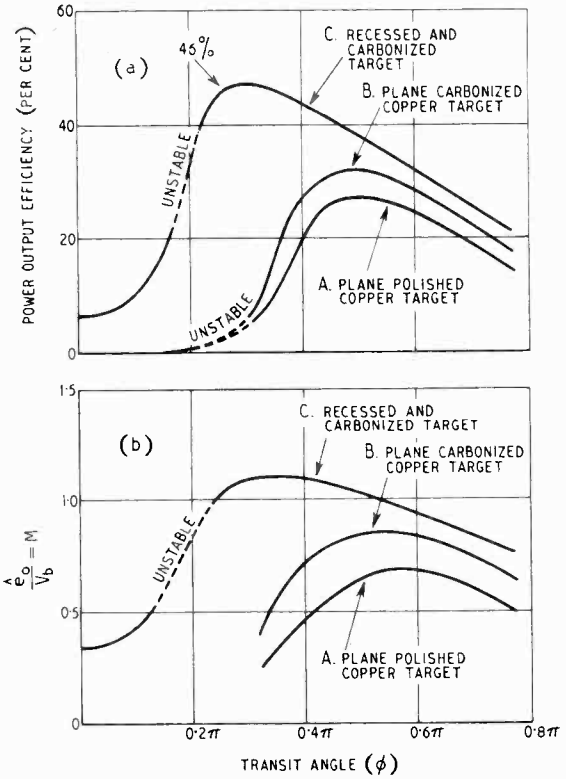


Fig. 7. *Experimental results; (a) shows the efficiency and (b) the modulation depth. $\hat{e}_0$ = the maximum instantaneous value of the voltage across the load.*

Nevertheless, it must be noted that the secondary-radiation coefficients of a plane polished copper surface are quite high enough to make the measured maximum efficiency (of about 28% at $\phi = \pi/2$) much less than the theoretical maximum efficiency of 44% at the same transit angle in the absence of secondary radiation, (theoretical curve A of Fig. 4). Moreover, it was not found at all easy always to repeat the measured curves A of Fig. 7 with a plane polished copper surface. Surface contamination and consequent increases of the secondary-electron coefficients of the surface, were found to be capable of removing the maximum, that is the suppression effect, almost entirely. This made the result of doubtful value in engineering design. It was concluded, indeed, that it would be unlikely that any suppression by transit-angle effect would be observable in practical resonator-valves in which the presence of a barium getter or a coated cathode might contaminate the appropriate parts of the copper resonator walls and that this would probably account for the apparent absence in the literature of any mention of the effect.

The substitution of a target having a sand-blasted and carbonized plane surface reduces the secondary-electron coefficients and was found slightly to improve the efficiency. The suppression effect tends to be more nearly independent of contamination, (curves B of Fig. 7). Nevertheless the improvement was not by any means all that could be wished for engineering purposes.

6. An Engineering Solution

A third kind of target surface was also used. It was recessed as shown in Fig. 8, so that most of the primary electrons struck it only at the bottom of the recesses. The results are shown in curves C of Fig. 7. The maximum efficiency is about 46% at an optimum transit angle of about 0.3π and the useful range of transit angles over which the efficiency approaches the maximum is wider than is the case with the other targets. This experiment was repeated a considerable number of times and no variations due to contamination were found.

The effectiveness of this recessed kind of target is broadly explainable in the light of the theory. Many of the secondary electrons which are emitted from a target which has a plane surface are emitted sideways and do not travel parallel to

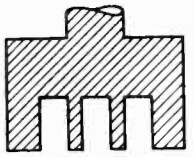


Fig. 8. *Recessed target.*

the electric field. These secondary electrons, therefore, may not be fully suppressed by the field; i.e., they may describe sideways trajectories in the field such that they withdraw energy from it. On the other hand, most of those secondary electrons which are emitted sideways from the bottom of a recess will be caught by the sides of the recess and therefore cannot travel through the field in such a way as to withdraw energy. The remainder of the secondary electrons which

are emitted, not sideways, but in the general direction from which the primary electrons travel, will not be caught in this manner by the sides of the recess; but, instead, will tend to be fully suppressed by the field because they travel in directions approximately parallel to it.

Suppression by the electric field only takes place over alternate half cycles; therefore, though most of the secondary electrons are emitted during those half cycles, it will be desirable to maintain the ratio of secondary electrons emitted per primary electron at as low a value as possible and to minimize the effects of surface contamination. A roughened and carbonized surface is preferable, for this reason, whatever kind of target is used.

It was concluded that a recessed and preferably carbonized surface, when used at the appropriate transit angle, is a satisfactory engineering solution to the problem of secondary electrons in microwave resonator-valves.

7. Unstable Points

The curves of Fig. 7 show portions which are unstable; but the instability does not occur at working values. The instability is not explicable in terms of the present theory and has not been further investigated. It would not be surprising, however, if, at intermediate transit angles, where most of the secondary electrons are not suppressed, unstable conditions exist due to some of the secondary electrons remaining in the field for transit times such that they give up, instead of absorb, energy.

8. Variation of Target Potential

Fig. 9 shows typical static plots of target current versus target voltage V_t for a fixed voltage V_{SA} of 2,500 volts on the entrance wall.

For all values of the target voltage V_t which are less than V_{SA}, the target current is greatly reduced because of the emission of secondary electrons from the target which travel unimpeded to the entrance wall of the resonator on the other side of the gap to the target. This is so for all the various target surfaces tried.

The suppression effect described in this paper is operative when the steady target voltage V_t is less than V_{SA}, as well as when they are equal. A characteristic like that of Fig. 9 was plotted while the electron beam was modulated to a small depth at an audio frequency. The transit angle thus tended to zero. An audio-frequency power output indicator was included in the target circuit and showed a lack of output when the target voltage V_t was less than V_{SA} as would be predicted from Fig. 9. This shallow modulation was then raised in frequency so that the transit angle had the optimum value of 0.3π [Fig. 7(a)] and an r.f. power output indicator was used.

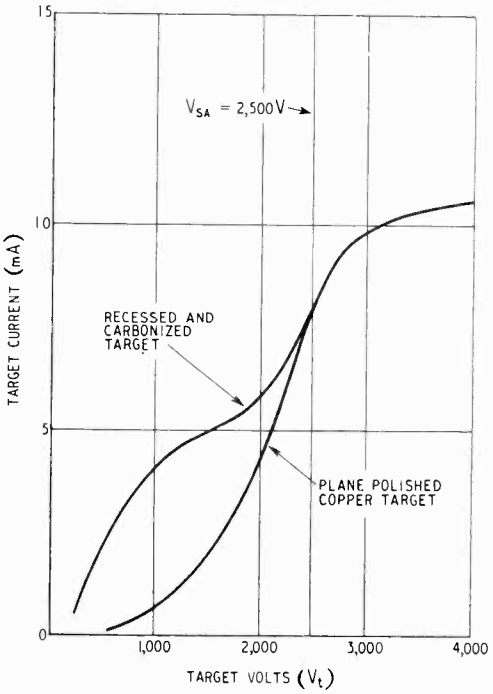


Fig. 9. *Static characteristics of the experimental apparatus illustrated in Fig. 6. V_T represents the target voltage, and V_{SA} the voltage of the entrance wall.*

The target voltage V_t was varied above and below V_{SA} but this time the r.f. power output indicator was fully operated for all values of the target voltage, both above and below V_{SA}, because the secondary electrons had been effectively suppressed by the effect described in this paper.

9. Acknowledgment

This research programme was carried out at the Electronics Department of Rediffusion Ltd., London.

REFERENCES

[1] J. H. O. Harries, "Cavity Resonators and Electron Beams," *Wireless Engineer*, March, April and May 1947, Vol. 24, pp. 71, 109 and 135.

[2] J. H. O. Harries, British Patents 380 429 and 385 968/31 and "Anode to Accelerating Electrode Space in Thermionic Valves," *Wireless Engineer*, April 1936, Vol. 13, p. 190.

[3] J. H. O. Harries, "Secondary Electron Problems," *Electronic Engineering*, January 1942, Vol. 14, p. 586. "Secondary Electron Radiation," *Electronics*, September 1944, pp. 100 and 180.

[4] T. S. Popham, "Transit Time Effects in Output Fields," *Wireless Engineer*, (to be published shortly).

SPACED LOOP AERIALS

Comparison of Simple and Twin Types

By F. Horner, M.Sc., A.M.I.E.E.

(Communication from the National Physical Laboratory)

SUMMARY.—A comparison is made between two types of screened loop which have been used in a coaxial spaced-loop aerial system at very high frequencies. Both types of loop were of square form, with a side of length 90 cm approximately. The 'simple' type of loop had the screen gap in the centre of one side and the output terminals at the centre of the opposite side. The 'twin' type had two screen gaps at the centre points of opposite sides and the output terminals were at the centre point of an additional member parallel to, and midway between, the sides containing the screen gaps.

It is shown that the use of the twin type of loop rather than the simple type in a coaxial spaced-loop direction finder leads to three advantages; (*a*) the system is easier to test for polarization error, (*b*) the geometrical tolerances in setting up the system are less strict, (*c*) the system is mechanically more stable and can be designed with a lower moment of inertia. These advantages are achieved with no apparent reduction in sensitivity or in directional accuracy.

1. Introduction

SPACED-loop aerial systems described in previous papers[1, 2] have consisted of coaxial square screened loops supported by horizontal transmission lines connected to the mid-points of the lower members, the screen gaps being at the mid-points of the upper loop members. This method of construction is satisfactory for small loops, but with large loops it becomes mechanically unwieldy and has undesirable electrical characteristics. It is then advantageous to use loops of mechanically balanced types which can be supported at their centres of gravity and which, at the same time, are free from the electrical disadvantages of large 'simple' loops One such type of loop is the 'twin' loop which has been described elsewhere [1,3] and which is illustrated in Fig. 1 as a component of a spaced-loop system.

In one of the above papers[1] it was demonstrated that the use of this or a more complex type of loop is essential in certain forms of spaced-loop system if satisfactory performance as a direction finder is to be obtained. While a coaxial system does not fall within this category, there are distinct advantages to be gained by using twin loops in this system. The present paper discusses these advantages and some features of the design of a coaxial system with twin loops.

2. Simple v. Twin Loops

The following discussion of the advantages of twin loops will be illustrated by a comparison between the electrical and mechanical characteristics of two rotating coaxial spaced-loop systems which have been constructed for use at frequencies in the neighbourhood of 30 Mc/s. The first system, described briefly in a previous paper[2], consists of simple screened loops, 92 cm square, supported at the mid-points of the lower members; the second is that illustrated in Fig. 1, the loops being 90 cm square and supported at their centres of area. The screen gaps are in the upper and lower members. These two aerial systems had identical transmission lines, the spacing between the loops being 2.2 metres, and were used with the same receiver, which had a balanced input circuit.

MS accepted by the Editor, May 1947.

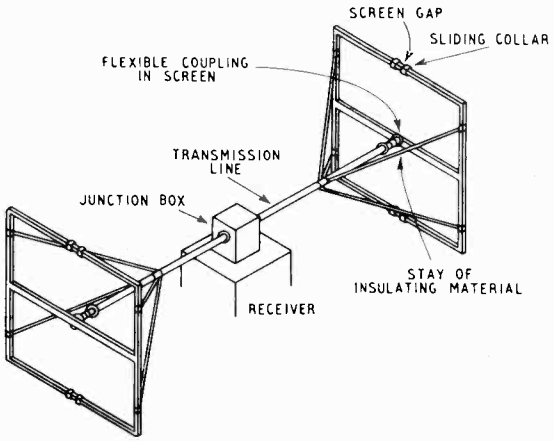


Fig. 1. *General view of twin-loop aerial assembly.*

2.1 *Electrical Characteristics.*

The differences in electrical behaviour between simple and twin loops are bound up with the non-uniform distribution of current in them. Provided the loops are not too large (perimeter $\not> \lambda/3$) the current has a cosine-law distribution with uniform phase. It has been shown[3] that as regards its polar diagram a simple loop with non-uniform current is equivalent to an ideal loop carrying a uniform current (which will be

referred to as an effective uniform loop) plus a dipole parallel to the side containing the screen gap. In a spaced-loop direction finder there are two undesirable characteristics associated with these effective dipoles.

(a) The presence of the dipoles makes the testing of the aerial system for polarization errors very difficult. When tested with the aid of a nearby elevated transmitting dipole the system exhibits large polarization errors which are due to the proximity of the transmitter and which obscure any genuine instrumental defects. This phenomenon has been fully described elsewhere[2] and it is necessary to state here only that the errors can be considered in two parts, due to unwanted pick-up on the effective uniform loops and the effective dipoles respectively. Of these, errors due to the dipoles are by far the greater in the system described. For example, with a test transmitter elevated and polarized at 30 degrees to the horizontal, and about 20 metres from the direction finder the dipole errors are of the order of 10 degrees while the loop errors are of the order of 1 degree with ground of average properties.

(b) As the pick-up in a simple loop is dependent on its orientation in its own plane (as represented by the effective dipole) it is necessary, in a coaxial spaced-loop system, to ensure not only that the loops lie in parallel planes, but also that corresponding sides are parallel. In the system described a rotation of one loop by 1 degree about the transmission line as axis would cause an error of 0.7 degree due to pick-up on the effective dipole, the radiation being assumed horizontally incident. (This error is independent of polarization and its value is a maximum at horizontal incidence.) The sign of the error would be the same for the direct and reciprocal bearings and so would not be revealed by a reciprocal error test.

Although a misalignment of one degree in setting up the loops is greater than one would normally expect, very careful setting up would be required to make full use of the potential accuracy of the spaced-loop system. It is worthy of note that, apart from the effects of ground reflections, the error is substantially independent of frequency, depending (with square loops) only on the ratio of loop side to loop spacing.

The twin loop, due to its non-uniform current, is equivalent to a uniform loop plus a pair of spaced dipoles of opposite senses[3]. A spaced-loop system with twin loops, therefore, shows similar effects to one with simple loops, but errors are of much smaller magnitude. For comparison the proximity errors may be quoted for the twin-loop system tested at the same frequency and under the same conditions as before. The error due to unwanted pick-up on the effective uniform loops under given test conditions depends only on the loop spacing and is therefore the same (1 degree) for both systems. That due to pick-up on the spaced-dipole pairs of a twin-loop system is, however, only about 0.4 degree, this being typical of values over quite a wide range of frequency and not due to a fortuitous phase relationship between wanted and unwanted pick-up. Furthermore the loop type of error, which is now the greater, is (with either system) introduced solely because of the imperfect reflecting properties of the ground and it should be possible to eliminate the error by use of a well-conducting ground mat during test work. Thus the system can be made almost free from proximity errors during polarization-error tests, which is a great help towards the detection of genuine polarization errors due to instrumental imperfections.

The pick-up on the effective spaced dipoles of either twin loop is of very small magnitude compared with that on the effective dipole of a simple loop. Consequently bearing errors of the type previously discussed, due to small misalignment of the loops, are negligible.

2.2 *Mechanical characteristics*

From the point of view of the direction-finder operator the main advantages of a system with twin loops lie in the improved mechanical stability and lower inertia. Experience shows that higher accuracy in aural-null bearing determination is obtained by swinging the aerial system rapidly. It is therefore desirable to reduce the moment of inertia of the system to a minimum. In a system with simple loops, however, the stresses during swinging are high and the loops and transmission lines must be made strong.

The most serious stresses are those tending to twist the loops about the transmission lines when an angular acceleration is imparted to the system. For instance, with the system with simple loops described above, an upper limit to the rate of swinging may be taken to be one complete swing per second for a swing of $\pm$ 20 degrees. Assuming that the swing is simple harmonic, the force acting at the centre of gravity of each loop, 18 inches above the point of support, is about 7 lb weight at the limits of the swing. This force tends to distort the loop, to shear the bolts securing the loop to the transmission line and to twist the transmission line. If the rate of swing approaches the mechanical resonance frequency of the system the forces may be very large indeed, although with the system under discussion the resonance frequency is about three swings per

second, which can be approached only on small swings. It is evident that the system must be strongly constructed to withstand these stresses, and the moment of inertia is therefore high (130 lb ft², of which 80% is due to the loops themselves). It was considered that loops of this type could not be made much lighter without impairing the rigidity, and no simple and satisfactory system of stays could be devised for preventing torsional strain. It seemed preferable to use mechanically balanced loops of the 'twin' type which can be supported at the centres of gravity. The design of such loops is discussed later, but it may be stated that it has been found possible to make the twin loops of much lighter construction than the simple loops, the overall moment of inertia being 78 lb ft².

3. Design of Twin-Loop System

Further details will now be given of the electrical and mechanical design of the twin loops described briefly in Section 2.

3.1 *Electrical design*

If the system has roughly the same overall dimensions with twin loops as with simple loops, there should be no significant difference in sensitivity, with proper design. The two types of loop would provide approximately the same available output power at the terminals when placed with their planes in the plane of incidence of an external field. It is, however, necessary to match or tune the loops to obtain high sensitivity. The system was required to be tunable over a range about 3 Mc/s on either side of 30 Mc/s with a 5–100 pF capacitor across the balanced input terminals to the receiver.

It was decided from considerations of sensitivity that the loops should be about 1 m square. They were to be used with existing transmission lines of which the characteristics were known, the loop spacing being fixed by the lines at 2.2 m. The internal construction of the loops is shown schematically in Fig. 2 (a) and equivalent circuits are given in Fig. 2 (b) and (c). The derivation of these equivalent circuits is discussed in the appendix.

There are several other forms of internal construction which could be used. For instance the inner conductors could be connected to the screen at the right-hand sides of the gaps at C_1C_2, since there is no pick-up on the portion C_1DC_2 of the internal transmission line, and it therefore plays no essential part in the action of the loop. The design shown is reasonably easy to construct, however, and appears to have no inherent disadvantages.

The impedance of the loop may be readily determined from the equivalent circuit of Fig. 2 (c). The loop impedance measured across the twin line at the point B is given by

$$Z_B = \frac{R + 4j(Z_{01}\tan\beta_1 s + Z_{02}\tan\beta_2 s)}{2Z_{02}(1-\tan^2\beta_2 s)+j(R+4jZ_{01}\tan\beta_1 s)\tan\beta_2 s} \cdot 2Z_{02}$$

assuming that there are no ohmic losses,
where $Z_{01} = cL_1$

c = velocity of propagation in the outside medium,
L_1 = the inductance per unit length of the four members of the screen comprising the loop proper,
β_1 = phase constant in the outside medium,
Z_{02} = characteristic impedance of the concentric transmission lines formed by the screen and inner conductor,
β_2 = phase constant within these lines,
R = radiation resistance of the loop proper,
s = length of the loop side.

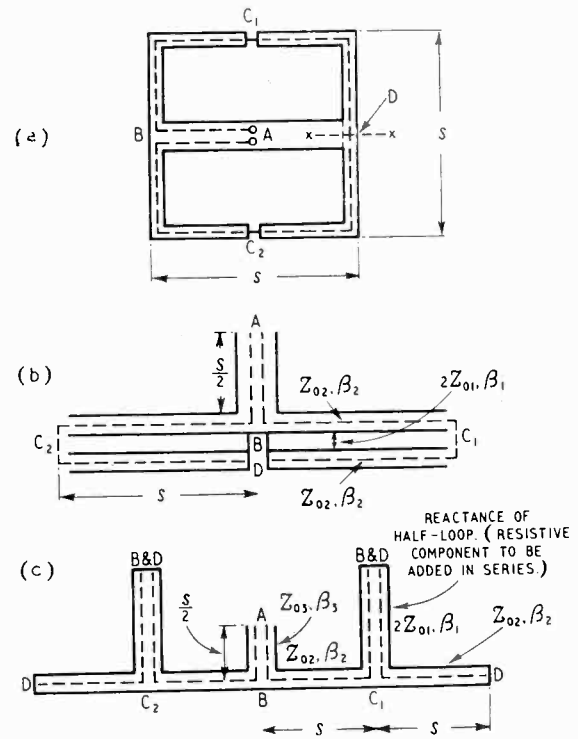


Fig. 2. *Equivalent circuits of a twin screened loop. Z_{01}, Z_{02}, and Z_{03} are respectively the characteristic impedances of the loop-screen perimeter, of the concentric line within the perimeter and of the screened twin line in the cross-member of the loop; β_1, β_2 and β_3 are the corresponding phase constants. The length of the loop side is denoted by* s.

Z_{01} has the dimensions of an impedance and may be termed the characteristic impedance of the loop screen. R is the ratio of the available power from the loop to the mean-square current at each gap. If ohmic losses cannot be neglected R must be suitably modified and allowance must be made in the equation for the losses in the internal transmission lines.

The impedance at the output terminals A of the loop is derived from Z_B by transformation along the transmission line AB (Fig. 2).

It should be noted that if the sides of the loop are short compared with the wavelength, the impedance is the same as that of a simple screened loop with the same perimeter.[4] This indicates that the two halves of the twin loop are effectively connected in series for the purpose of determining the output impedance.

The impedance formula was used to determine the necessary dimensions of the loop to give the required reactance at the centre of the system (the characteristics of the line AB and the main transmission lines of the aerial system being known). The required length of the loop side was 90 cm and the tuning range of the system with the 5–100-pF capacitor was 27–34 Mc/s.

It is worthy of mention that the polar diagram of the twin screened loop shown in Fig. 2(a), unlike that of a simple screened loop, can be affected by unbalance of the internal circuits to which the loop is connected. This would not lead to bearing errors in a spaced-loop direction finder, however, provided the loops were unbalanced identically.

3.2 *Mechanical design*

The mechanical design was required to provide maximum rigidity with a minimum moment of inertia. It was decided to design the loops initially with outer conductors of cross-section $\frac{3}{8}$ in $\times$ $\frac{3}{8}$ in made from 1/64-in thick copper sheet, with an inner conductor of $\frac{1}{8}$-in brass tube, supported by keramot insulators. Preliminary experiments were conducted on the strength of these members, tested as cantilevers, and the design carried out with due regard for the stresses which arise during swings. These are (a) the static forces due to the weights of the elements, (b) centrifugal forces dependent on the angular velocity of the system, (c) forces dependent on the angular acceleration of the system. Taking all these forces into account a satisfactory design was achieved with the dimensions quoted, insulated stays being added to increase rigidity as shown in Fig. 1. These should preferably be hollow for maximum rigidity/weight ratio but tube could not be obtained and $\frac{3}{8}$-in diameter rod was used with suitable clamps to the loops and transmission lines. The centre cross-members of the loops were placed horizontally, as this arrangement was rather more rigid when the system was swung rapidly and was also more easily connected to the existing screened-twin transmission lines in which the conductors were one above the other. The weakest parts of the structure were the upper and lower members of each loop, which tended to bend, so it was necessary for the insulation at the screen gaps to be of sufficient length to give adequate support. Sliding collars were provided to adjust the widths and positions of the gaps and these were of the same material as the outer conductors of the loops. Controls were provided in the central junction box for balancing the two halves of the aerial system and so minimizing reciprocal errors.

As mentioned previously the overall moment of inertia of the system was 78 lb ft^2, that of the loops alone being only half that of the simple loops.

3.3 *Instrumental performance*

The aerial system was found to be mechanically satisfactory, rapid swinging being possible without instability. Sensitivity was, as expected, about the same as with the simple loops with the same receiver. A bearing could be taken with a silent swing of $\pm$ 5 degrees on a horizontally incident, vertically polarized, 100% modulated signal with a field strength of 4 μV/m. Reciprocal error could be reduced to 0.2 degree at all frequencies within the tuning range (27–34 Mc/s) with a single setting of the balancing controls.

Polarization errors were measured with a dipole transmitting aerial elevated and polarized at 30 degrees to the horizontal, and distant about 20 m from the direction finder. Calculated proximity errors for these conditions at 30 Mc/s were 1.2 degrees for the loop type of unwanted pick-up and 0.4 degree for pick-up on the dipoles. These values were practically constant throughout the tuning range. The measured error was 1.9 degrees, also substantially constant throughout the range. The total error (including the flattening component) was not easily measurable with the low-power transmitter used for the tests, as the swings were all silent, but the value was almost certainly less than 2.5 degrees. It appears, therefore that there were no large polarization errors due to instrumental defects.

These results indicate that the advantages associated with the use of twin loops are gained without introducing any significant change in sensitivity or increase in reciprocal error. Although the limitations of measurement technique do not permit a precise comparison between

the polarization errors of the two systems, there is no evidence of large differences between them. Any polarization errors due to instrumental imperfections would be expected to be small with either system, but the twin loop system has the advantage that experimental confirmation of the smallness of these errors is obtainable.

4. Conclusions

1. The use of large, simple loops in a coaxial spaced-loop direction finder leads to the following undesirable electrical and mechanical characteristics :—

(a) the system is difficult to test for polarization errors,

(b) the loops require accurate alignment for the avoidance of errors,

(c) the system is mechanically unstable and has a high moment of inertia.

2. A system consisting of coaxial spaced twin loops is almost free from these undesirable features. The system is not difficult to test ; the loops must lie in parallel planes but their corresponding sides need not be exactly parallel ; and mechanical stability and a relatively low moment of inertia are readily attainable.

3. These advantages resulting from the use of twin loops are not accompanied by any significant loss in sensitivity or in directional accuracy.

5. Acknowledgments

The work described above was carried out as part of the programme of the Radio Research Board and this paper is published by permission of the Department of Scientific and Industrial Research.

REFERENCES

[1] W. Ross, "The Development and Study of a Practical Spaced-Loop Radio Direction Finder for High Frequencies", *J. Instn elect. Engrs*, 1947, Vol. 94, Pt III, No 28, pp. 99–107.

[2] F. Horner, "An Experimental Spaced-Loop Direction Finder for Very High Frequencies", *J. Instn elect. Engrs*, 1947, Vol. 94, Pt III, No 28, pp. 126–133.

[3] F. Horner, "Properties of Loop Aerials", *Wireless Engineer*, 1948, Vol. 25, p. 254.

[4] R. E. Burgess, "Reactance and Effective Height of Screened Loop Aerials", *Wireless Engr*, 1944, Vol. 21, p. 210

APPENDIX

Equivalent circuit of a twin screened loop

When a twin screened loop is connected to a balanced circuit two simplifications may be made which assist in the formulation of the equivalent circuit.

(a) The inner and outer conductors at D, Fig. 2 (a), are at the same potential, so a short circuit may be placed across the inner transmission line at this point (as shown by XX) without disturbing the current and potential distributions. It is then permissible to regard the upper and lower halves of the right-hand limb of the loop as entirely separate.

(b) The potential of the outer screen is the same at B and D (there being no current in the centre cross-member) so these points can be assumed to be connected by a perfect short-circuit. One-half of the loop screen as seen from the point C_1 (or C_2) appears as a short-circuited transmission line of length C_1B (= C_1D) and effective characteristic impedance $2Z_{01}$ where Z_{01} has the definition given in Section 3 of the paper.

One equivalent circuit of the loop is therefore that shown in Fig. 2 (b), where the lower short-circuited concentric transmission lines represent the parts C_1D and C_2D, and each short-circuited *twin* transmission line formed by the screens represents the reactance of half the loop screen. Associated with this reactance is a resistive component due to radiation, given by half the ratio of the power available from the loop to the mean-square current at the screen gaps. This resistance is in series with the half-loop reactance.

Dealing with the right-hand half of Fig. 2 (b) only, the lower concentric transmission line and the twin line formed by the two screens are effectively connected in series at the point C_1 and applied to the upper concentric line BC_1. (Note that the screen of each concentric line has a dual role ; the inner surface is part of a concentric line and the outer surface is part of the twin line, the two surfaces being connected in series at the point C_1). An alternative circuit representing this arrangement is shown in Fig. 2 (c), where the outer screened-twin sections represent the reactance of the loop proper (with which must be associated the resistive component of loop impedance) and the outer concentric sections represent the concentric transmission lines C_1D, C_2D.

It should be stressed that Fig. 2 (c) is a schematic equivalent circuit and that the inner conductors are not supposed to represent completely the inner conductors of the actual loop. The outer screened-twin transmission lines are merely convenient representations of the reactance of the loop proper and in fact the inner conductors of these lines correspond to the outer screen of the loop.

CALCULATING MUTUAL INDUCTANCE

Co-axial and Co-planar Multilayer Coils

By **R. C. de Holzer, D.Sc.(Tech.)**

SUMMARY.—Four diagrams provide simple formulae and graphs for the calculation of mutual inductance and coupling factor for the three possible cases of multi-layer coils; i.e., (a) identical coaxial coils, (b) different coaxial coils, (c) concentric coplanar coils.

The data covers the cases of coils of very small cross-sectional area (point-section), and those of finite cross-sectional area.

Formulae and graphs are derived from known equations and the results obtained are sufficiently accurate for most practical cases.

THE use of the graphs and formulae of Figs. 1-4 is largely self-explanatory and an example will suffice to illustrate it. After this, the derivation of the graphs and formulae is given.

1. Example

	Coil 1	Coil 2
Mean diameter	$= a_1 = 0.875$ in	$a_2 = 0.625$ in
Width	$= b_1 = 0.25$ in	$b_2 = 0.25$ in
Depth	$= c_1 = 0.25$ in	$c_2 = 0.125$ in
Turns	$= n_1 = 300$	$n_2 = 30$
Mean distance	$= d = 1$ in	

This example corresponds to the case of two different coaxial coils. The first step is to reduce them to two equivalent identical coils with the aid of Fig. 2. Following the procedure given there in order we get: for coils of point section, $a_0 = 0.75$ in, $d_0 = 1.0$. Then Fig. 1 gives $\beta_0 = 1.33$, $\zeta_0 = 0.0009$, $M_0 = 6.08\ \mu$H. The next step is to introduce a correction for coils of finite section by applying steps 4, 5 and 6, of Fig. 2,—giving $m_1 = 1.75$, $m_2 = 1.67$, $m = 1.71$, $\Delta = 0.03$, $M_1 = 6.26\ \mu$H.

To find k we use Fig. 4 and get $\beta_0 = 1.33$, $k_0 = 1.35\%$. Correcting for coils of finite cross-section gives, $g_1 = 2.8$, $g_2 = 2.86$, $g = 2.83$, $\Delta = 0.03$, $k = 3.9\%$.

2. Derivation of Charts

Mutual inductance between two identical circular coaxial point section coils.

With the symbols of Figs. 1 and 2, the known formula for mutual inductance is [1a]

$$M = 1.27\, a_0\, n_1\, n_2\, F \quad .. \quad .. \quad .. \quad (1)$$

where [1b] $F = \phi_1 (r_2/r_1)$ .. .. .. (2)

Now

$$\frac{r_2}{r_1} = \sqrt{\frac{(a_1 - a_2)^2 + 4d^2}{(a_1 + a_2)^2 + 4d^2}} \quad .. \quad .. \quad (3)$$

MS accepted by the Editor, March 1947

Writing

$$\beta_0 = d_0/a_0 \quad .. \quad .. \quad .. \quad .. \quad (4)$$

when $a_1 = a_2$ this becomes

$$r_2/r_1 = 1/\sqrt{1 + 1/\beta_0^2} \quad .. \quad .. \quad .. \quad (5)$$

or $$\beta_0 = 1/\sqrt{(r_1/r_2)^2 - 1} \quad .. \quad .. \quad (6)$$

putting

$$M_0 = a_0 n_1 n_2 z_0$$

(1) becomes

$$z_0 = 1.27\, F \quad .. \quad .. \quad .. \quad .. \quad (7)$$

The calculated values of z_0 are given in the Table and plotted in Fig. 1.

Mutual inductance between two different circular and co-axial coils of point-section.

If a pair of identical coils exists, which from the point of view of mutual inductance is equivalent to a pair of different coils, then (Fig. 2)

$$a_0 = \sqrt{a_1 . a_2} \quad .. \quad .. \quad .. \quad .. \quad (8)$$

and from (3) and (5)

$$\frac{(a_1 - a_2)^2 + 4d^2}{(a_1 + a_2)^2 + 4d^2} = \frac{1}{1 + 1/\beta_0^2} \quad .. \quad .. \quad (9)$$

Substituting (4) and (8) in (9) gives

$$d_0 = \sqrt{d^2 + \tfrac{1}{4}(a_1 - a_2)^2} \quad .. \quad .. \quad (10)$$

Thus such an equivalent pair of coils exists, where d_0 is equal to the smallest distance (r_2) between the centres of the cross-sections of these two different coils.

Mutual inductance between two circular concentric coplanar coils.

With symbols as in Fig. 3, Putting $d = 0$ in (10) gives

$$d_0 = \tfrac{1}{2}(a_1 - a_2) \quad .. \quad .. \quad .. \quad (11)$$

Coupling Factor.

Using symbols of Fig. 4, the known relation between coupling-factor and mutual inductance is $k = M/\sqrt{L_1 L_2}$ .. .. .. (12)

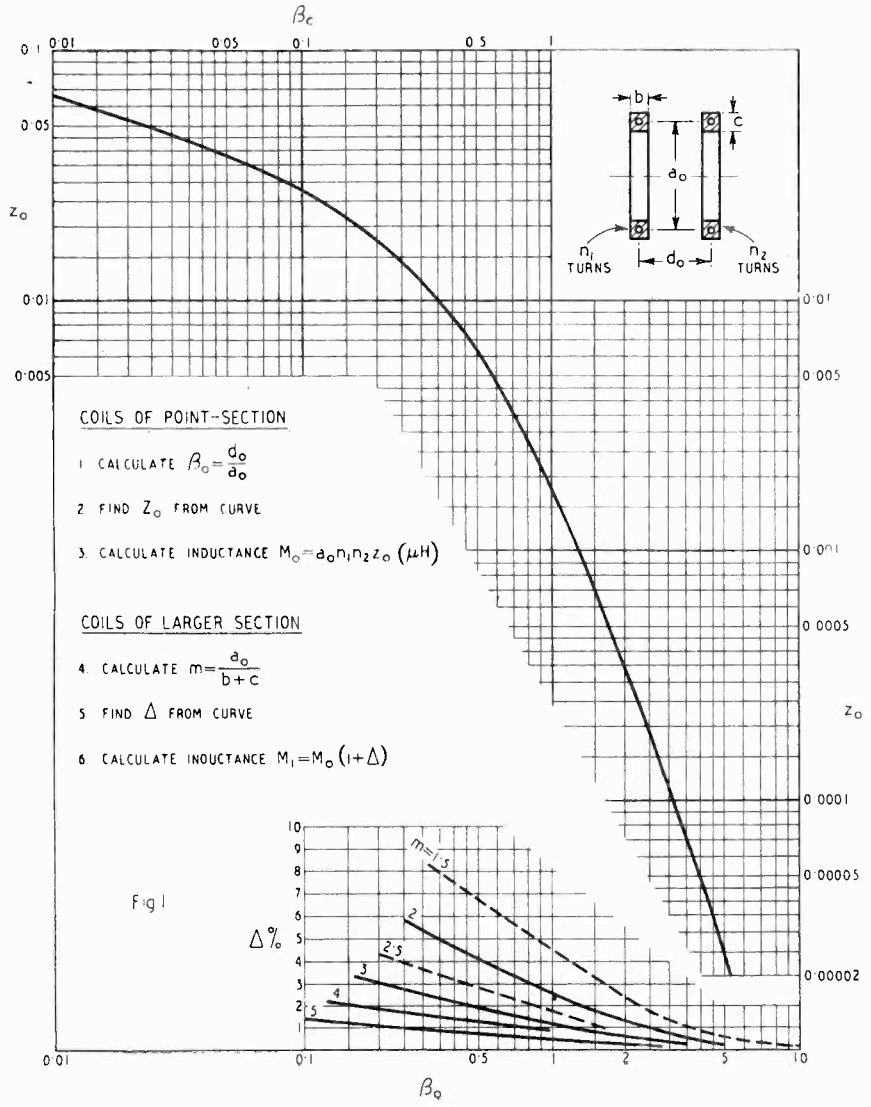
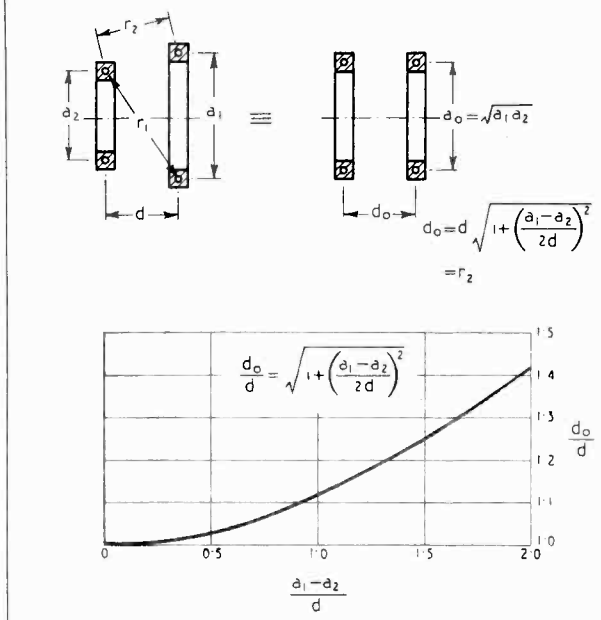


Fig 1



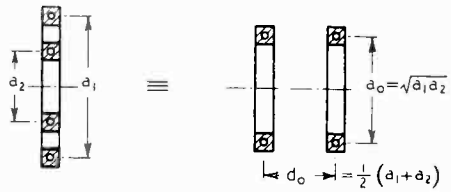
COILS OF POINT-SECTION

1. CALCULATE $a_o = \sqrt{a_1 a_2}$
2. FIND d_o FROM CURVE
3. PROCEED AS IN Fig. 1 (1)

COILS OF LARGER SECTION

4. CALCULATE FOR EACH COIL $m_n = \frac{a_n}{b_n + c_n}$
5. CALCULATE $m = \frac{1}{2}(m_1 + m_2)$
6. PROCEED AS IN Fig. 1. (5)

Fig 2



COILS OF POINT-SECTION

1. CALCULATE $a_0=\sqrt{a_1 a_2}$
2. CALCULATE $d_0=\frac{1}{2}(a_1-a_2)$
3. PROCEED AS IN Fig 1 (1)

Fig 3

COILS OF LARGER SECTION

4. PROCEED AS IN Fig 2 (4)

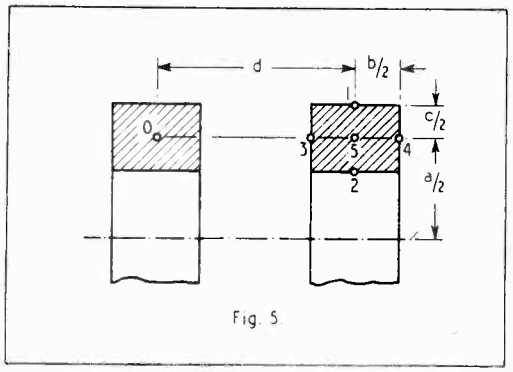


Fig. 5

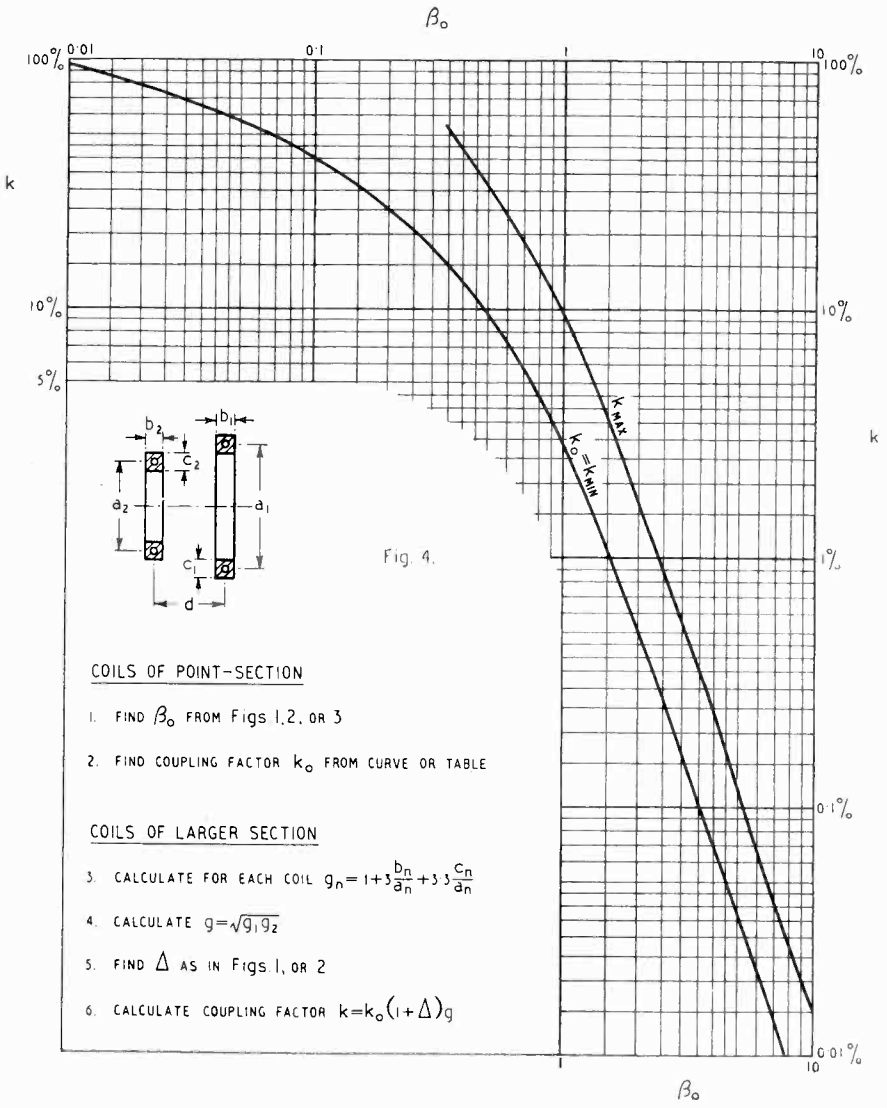


Fig. 4.

COILS OF POINT-SECTION

1. FIND β_0 FROM Figs 1, 2, OR 3
2. FIND COUPLING FACTOR k_0 FROM CURVE OR TABLE

COILS OF LARGER SECTION

3. CALCULATE FOR EACH COIL $g_n=1+3\frac{b_n}{a_n}+3\,3\frac{c_n}{a_n}$
4. CALCULATE $g=\sqrt{g_1 g_2}$
5. FIND Δ AS IN Figs 1, OR 2
6. CALCULATE COUPLING FACTOR $k=k_0(1+\Delta)g$

The self-inductance of a multilayer coil is given by Wheeler's formula [2a]

$$L = 0.067\, a\, n^2/g\ \mu\text{H} \quad .. \quad .. \quad .. \quad (13)$$

where the form-factor

$$g = 1 + 3b/a + 3.3c/a \quad .. \quad .. \quad .. \quad (14)$$

Substituting (7), (8) and (13) in (12) gives

$$k = 15\, z_0 \sqrt{g_1 g_2} \quad .. \quad .. \quad .. \quad (15)$$

For coils of very small cross-section a more accurate formula for the inductance is[3]

$$L = 0.037\, a\, n^2 \,.\, \log_{10}(2.45\ m) \quad .. \quad (16)$$

where $m = a/(b + c)$

For a coil of point-section

$b = c = 0$ therefore $g = 1$

and $k_0 = 15\, z_0 \quad .. \quad .. \quad .. \quad .. \quad (17)$

Coils of large cross-sectional area.

For two identical coils of large cross-sectional area, Fig. 5, M is given by [2b].

$$M_1 = (M_{01} + M_{02} + M_{03} + M_{04} - M_{05})\, n^2/3 \quad .. \quad .. \quad .. \quad .. \quad (18)$$

the subscripts of M define the point-section coils for which each individual M_0 is to be calculated from (7), ignoring the number of turns.

The computation of (18) is very tedious and the result in all practical cases is only a few per cent larger than the more easily obtainable result valid for point-section coils.

It can be demonstrated that (18) can be written

$$M_1 = M_0\,(1 + \Delta) \quad .. \quad .. \quad .. \quad (19)$$

where

$$\Delta = \phi_2\,(\beta_0 m)$$

In Fig. 1 values of Δ are given as calculated from (18) and (19).

The 'optimum' coil, for which $b = c = a/3$ and consequently $m = 1.5$ and $g = 3.1$ can be considered—for practical purposes—as the coil of maximum cross-sectional area; the coupling factor of two such coils is given in Fig. 4 and designated as k_{max}.

The factor $(1 + \Delta)$ applies actually to z_0, therefore, for all coils of finite cross sectional area:

$$z = z_0\,(1 + \Delta) \quad .. \quad .. \quad .. \quad .. \quad (20)$$

$$k = k_0\,(1 + \Delta)\sqrt{g_1 g_2} \quad .. \quad .. \quad .. \quad (21)$$

For two coils of different section

$$1 + \Delta = \sqrt{(1 + \Delta_1)\,(1 + \Delta_2)}$$

but as $\Delta \ll 1$

$$\Delta = \tfrac{1}{2}\,(\Delta_1 + \Delta_2) \quad .. \quad .. \quad .. \quad (22)$$

TABLE

β_0	z_0	k_0 (%)
0.01	0.063 7	95.6
2	52 7	79.1
3	46 2	69.3
4	41 7	62.6
0.05	0.038 2	57.3
6	35 3	53.0
7	32 9	49.4
8	30 8	46.2
9	29 0	43.5
0.1	0.027 5	41.1
2	17 2	25.8
3	11 8	17.7
4	08 47	12.7
0.5	0.006 31	9.49
6	4 74	7.11
8	2 83	4.25
1.0	0.001 78	2.67
2	1 20	1.80
4	0 827	1.24
6	0 595	0.893
8	0 438	0.657
2.0	0.000 332	0.498
2	257	0.386
4	197	0.296
6	159	0.239
8	130	0.195
3.0	0.000 104	0.156
5	071	0.107
4.0	047	0.071
5	034	0.051
5.0	0.000 025	0.038
6.	14	0.021
8.	06	0.009
10.	0.000 003	0.005

BIBLIOGRAPHY

[1] K. Henney, "The Radio Engineering Handbook," New York, 1941 (3rd edit.): (a) p. 99; (b) p. 98.

[2] F. E. Terman, "Radio Engineers' Handbook," New York, 1943 (1st edit.); (a) p. 62; (b) p. 73.

[3] F. L. Smith, "Radio Designer's Handbook," London, 1941, p. 149.

FREQUENCY MODULATION OF AN OSCILLATOR

By M. R. Gavin, M.B.E., M.A., B.Sc., M.I.E.E.

(Communication from the Staff of the Research Laboratories of The General Electric Company, Limited, Wembley, England.)

SUMMARY.—This note describes a simple method of increasing the frequency deviation of an oscillator by means of a suitably adjusted coupled circuit.

1. Introduction

USERS of self-oscillating circuits with more than one degree of freedom have all experienced the tantalizing readiness with which oscillation is obtained in unwanted modes. The oscillator with a primary tank circuit and a coupled load circuit has at least two possible frequencies of oscillation depending on the constants of the two circuits and the degree of coupling between them. Unless special precautions are taken it is possible for the frequency to change suddenly from one value to another when the operating conditions are varied. It is shown below that, by suitable choice of the circuit constants, this unstable state may be exploited to make an oscillator particularly sensitive to small changes of reactance so that a small modulating signal may cause a relatively large degree of frequency modulation.

2. Self-Oscillator with a Coupled Circuit

When a self-oscillatory tuned circuit has another tuned circuit coupled to it then the second circuit can have marked effects on the behaviour of the oscillator. Some of the effects are associated with the double-humped nature of the resonance curve of the two coupled circuits.

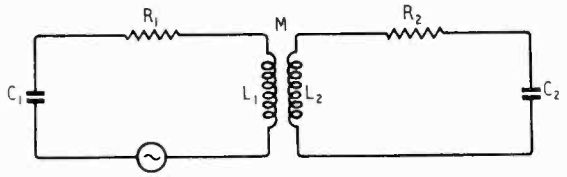


Fig. 1. *Generator with a coupled load circuit.*

Thus the frequency of oscillation may be that corresponding to one or other of the two humps of the resonance curve. The actual operating frequency is determined by the loading conditions. These effects have been known for a considerable time, and were discussed in some detail by Prince.[1]

MS accepted by the Editor, June 1947.

The circuit of Fig. 1 represents a generator in which L_1, C_1 and R_1 are the constants of the primary circuit, L_2, C_2 and R_2 are the constants of the secondary circuit and M is the mutual inductance between L_1 and L_2.

As far as the effects on the generator are concerned the secondary circuit can be replaced by equivalent resistance and reactance in series with the primary circuit as shown in Fig. 2, where

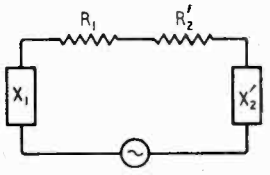


Fig. 2. *Equivalent circuit of Fig. 1.*

$$\left.\begin{aligned} X_1 &= \omega L_1 - \frac{1}{\omega C_1} = \text{primary series reactance,}\\ X_2 &= \omega L_2 - \frac{1}{\omega C_2} = \text{secondary series reactance,}\\ R_2' &= \frac{\omega^2 M^2}{Z_2^2}\, R_2 = \text{primary series resistance due to the secondary circuit,}\\ X_2' &= -\frac{\omega^2 M^2}{Z_2^2}\, X_2 = \text{primary series reactance due to the secondary circuit,}\\ \text{and } Z_2 &= \left\{R_2^2 + \left(\omega L_2 - \frac{1}{\omega C_2}\right)^2\right\}^{\frac{1}{2}} = \text{secondary series impedance.} \end{aligned}\right\} \quad (1)$$

In Fig. 3 the variation with frequency of the primary reactances X_1 and X_2' are shown. At the points where the two curves intersect $X_1 = -X_2'$, and the total primary reactance is zero; i.e., the points of intersection correspond to the resonant frequencies of the system. In Fig. 3 there are three such points, but the intermediate one f_2 is not obtained in practice for the following reasons. Suppose that the oscillator is operating at frequency f_2, and that for some reason the frequency is increased slightly, then the increase in $-X_2'$ is greater than the increase in X_1, so that the resultant change in reactance

is from zero to some capacitance. Now this capacitance is in series with the primary reactances and hence the resonant frequency must increase still further; i.e., the oscillator is in an unstable state. Under normal operating conditions, any change in frequency brings about a change in reactance which tends to oppose the frequency change and the conditions are stable. It may be observed from similar reasoning that f_1 and f_3 are stable frequencies.

The reactance curves in Fig. 3 have three points of intersection. It is possible to have only one intersection, in which case it can be seen that this point always gives stable conditions.

When there are two possible resonant frequencies such as f_1 and f_3 the oscillator, on being switched on, will in general pick that frequency at which it is the more lightly loaded. This statement is based on actual experience, and the mechanism is probably somewhat as follows. The impulse of the switching will set up oscillations at the natural frequencies of the circuit. As the valve acts as a negative resistance the amplitude of the oscillations will build up at a rate determined by the circuit resistance and the build up will be more rapid for the frequency giving the lighter load.

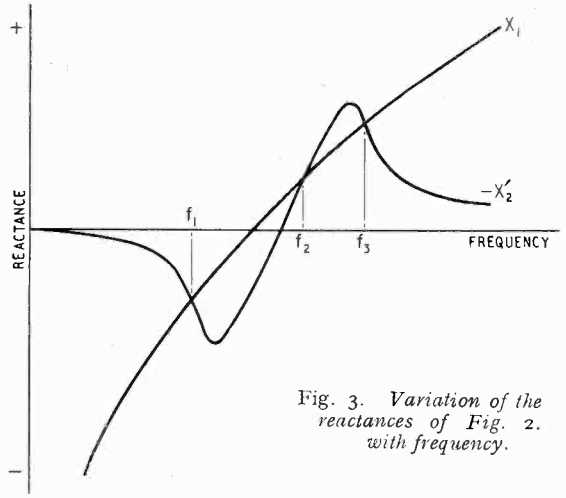


Fig. 3. *Variation of the reactances of Fig. 2. with frequency.*

If any element in the circuit is varied, then the frequency will vary also. It is possible that during the variation the loading conditions may alter so that the other resonant frequency becomes more favourable to the oscillator and there may be a sudden change of frequency.

If both circuits are tuned to the same resonant frequency f_0 so that

$$\omega_0 L_1 - \frac{1}{\omega_0 C_1} = \omega_0 L_2 - \frac{1}{\omega_0 C_2} = 0 \quad \ldots \quad (2)$$

then the two reactance curves cross the frequency axis at f_0 as shown in Fig. 4. If now the values of L_1 and C_1 are altered so that equation (2) still applies but the slope of X_1 increases then the two stable frequencies f_1 and f_3 approach one another until they finally coincide at f_0 when the curves have equal slope, as shown by the broken line X_1'

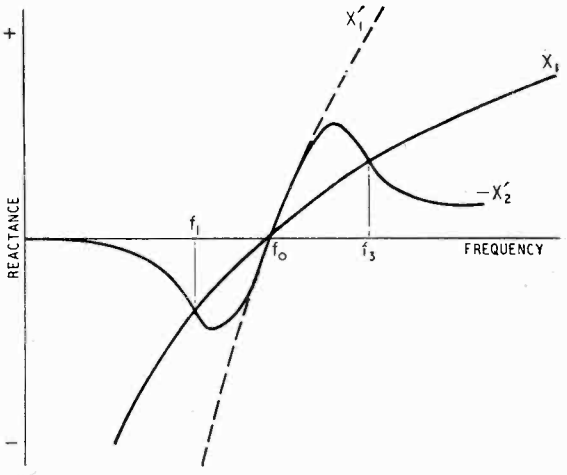


Fig. 4. *Variation of the reactances of Fig. 2 when both circuits are tuned to the same resonant frequency. The broken line shows the primary reactances adjusted to give large frequency modulation.*

in the figure. Under these conditions it is obvious that a small reactance change in either circuit will bring about a large frequency change. The circuit is on the verge of the unstable state already discussed. By bringing the circuits sufficiently close to this condition, relatively large frequency deviation may be achieved continuously by a small modulating reactance variation. With practical circuits the reactance curves are nearly straight at f_0 and linear frequency modulation can be obtained.

The adjustment of the circuit can obviously be effected by varying the constants of the secondary circuit and in practice M and R_2 are the main variables.

3. Relations Between the Circuit Elements

The quantitative relationships between the circuit elements can be determined as follows.

If $f_3 - f_1$ be small compared with f_0, then $\omega = \omega_0 + \delta\omega$ where $\delta\omega \ll \omega_0$.

From equations (1) and (2)

$$X_1 \approx 2\delta\omega L_1 \text{ and}$$

$$X_2' \approx -\frac{2\omega_0^2 M^2 \delta\omega L_2}{R_2^2 + 4(\delta\omega)^2 L_2^2} \quad \ldots \quad \ldots \quad (3)$$

The term in $(\delta\omega)^2$ is retained in the denominator since R_2 is also usually small. If R_2 is of the same order as $2\delta\omega L_2$, then $R_2 \ll 2\omega_0 L_2$; i.e., $\omega_0 L_2 / R_2 \gg \frac{1}{2}$ and the initial assumption that

$f_3 - f_1$ is small compared with f_0 is equivalent to assuming a reasonably large Q for the secondary circuit.

At the resonant frequency $X_1 + X_2' = 0$ and solving for $\delta\omega$ gives

$$\delta\omega = 0 \text{ or } \delta\omega = \pm\sqrt{\frac{\omega_0^2 M^2 L_2 - L_1 R_2^2}{4 L_1 L_2^2}}$$

The first value corresponds to f_0 and the other two to f_1 and f_3. For the condition for large frequency modulation, X_1' and $-X_2'$ coincide and $\delta\omega = 0$.

$$\therefore\ \omega_0 M = R_2\sqrt{L_1/L_2} \quad .. \quad .. \quad .. \quad (4)$$

From the above considerations it will be seen that the frequency range that can be covered is limited by the turning values of the reflected reactance, $-X_2'$. From equation (3) it can be seen that the turning values occur when

$$2\delta\omega L_2 = \pm R_2$$

and hence the maximum frequency deviation is given by $2\delta\omega = \omega_0/Q_2$ where Q_2 is the Q of the secondary circuit.

For some applications the degree of amplitude modulation is of importance. Since the useful output power is developed in R_2' some idea of the amplitude variation can be gained by considering the variation of R_2' over the range of the modulating cycle.

Using the same assumptions as before

$$R_2' = \frac{\omega_0^2 M^2 R_2}{R_2^2 + (2\delta\omega L_2)^2}$$

When $\delta\omega = 0$

$$R_2' = \omega_0^2 M^2 / R_2 \quad .. \quad .. \quad .. \quad (5)$$

At the extremities of the range $2\delta\omega L_2 = R_2$

and $R_2' = \omega_0^2 M^2 / 2R_2 \quad .. \quad .. \quad .. \quad (6)$

Thus there is a maximum variation in the effective output resistance of 2 to 1. If the circulating current were the same throughout the cycle then the useful power output would also vary 2 to 1. The actual variation in power output will depend on the value of the optimum load resistance. In one case investigated it was found that in order to exploit the condition under consideration the oscillator was overloaded at f_0. In that case the lighter load represented by equation (6) would give a greater efficiency at the extremities of the range than at f_0 so that the variation in power output should be less than 2 to 1 over the band.

4. A Practical Example

An example[2] of the use of the coupled-circuit technique to give wide frequency modulation is shown in Fig. 5. Two CV55 triodes are used as a self-oscillator operating in an earthed-anode push-pull circuit. The main oscillatory circuit L_1, C_1 is connected between the grids and the load circuit L_2, C_2, is coupled to it by means of the mutual inductance between L_1 and L_2. The effective resistance of the secondary circuit can be adjusted by varying the aerial-feeder tapping points on L_2. When the circuits are set up sufficiently near to the critical state then a small variation in any one of the circuit reactances produces a much larger frequency deviation than the same variation would produce otherwise. Factors of eight times the deviation have been obtained in practice with an oscillator of the type shown. In one particular war-time application the variation of reactance was produced by varying the mean current of the oscillator, thereby causing a small change in the effective interelectrode capacitances. The mean frequency was about 500 Mc/s and the change produced by the variation of anode current was about 2 Mc/s under ordinary conditions. When the circuit satisfied equation (4) the frequency deviation for the same change in anode current was about 10 Mc/s. In this case there was of course very considerable amplitude modulation which would be undesirable in some applications.

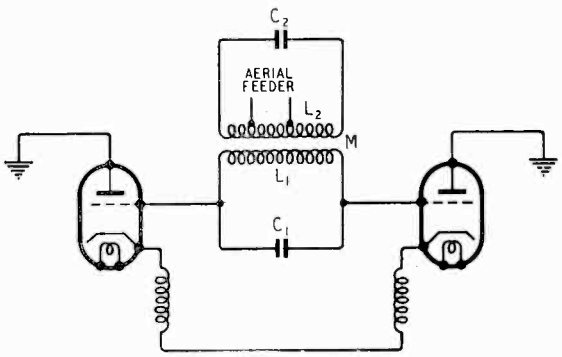


Fig. 5. *Push-pull oscillator with a coupled load circuit arranged for wide frequency deviation.*

Since there are several variable elements in the circuit, the setting up to the necessary condition requires care. The adjustment can be facilitated by the use of a 'panoramic' monitor; i.e., a receiver which scans a portion of the spectrum continuously and gives an oscillographic display of the output. The x-axis in the display indicates frequency and the y-axis shows signals entering the receiver.

The circuit is set up initially with a very loose coupling between L_1 and L_2, and with the secondary circuit lightly loaded. The primary circuit is tuned to the required frequency, as shown by the monitor and the secondary circuit is tuned to resonance, as determined by any suitable output indicator or by reaction on the oscillator mean currents. The coupling between

the two circuits is then increased and a point is reached where the frequency will jump suddenly to another value. At this stage the modulation should be applied and a condition can easily be found where the frequency coverage is discontinuous, showing that the modulating reactance is sufficient to cause a frequency jump. The aerial taps and the coupling between L_1 and L_2 are adjusted to reduce the gap in the spectrum and at the same time to increase the actual band covered. While the load and the coupling are the main variables, it may be necessary to adjust the circuit reactances too, as the various elements are interdependent to some extent.

5. Conclusion

The tendency to instability of a self-oscillator with two coupled circuits has long been a source of trouble in radio engineering, but by introducing a measure of control over this tendency it may be turned to good account by increasing the frequency deviation which can be obtained for a given reactance change. Increased deviations up to eight times have been obtained with careful adjustment. The process may be accompanied by a certain amount of amplitude modulation, but in many cases this is not of major importance. Where amplitude modulation is undesirable the coupled-circuit principle may still be used to give some increase in deviation, say two or three times. The adjustments for these smaller deviations are less critical.

6. Acknowledgments

The work described in this paper was carried out under a research contract for the Director of Scientific Research, Admiralty. The author desires to tender his acknowledgment to the Board of Admiralty for approval of the publication of this paper.

REFERENCES

[1] D.C. Prince. "Vacuum Tubes as Power Oscillators," *Proc. Inst. Radio Engrs.*, 1923. Vol. 11, page 405.
[2] British Patent No. 566,556.

The Late Sir Clifford Paterson

ELECTRICAL Research has suffered a great loss by the death on 26 July of Sir Clifford Copland Paterson. He was born on 17 October 1879. On leaving Mill Hill School he served a four years' workshop apprenticeship and attended classes at Finsbury Technical College and Faraday House. In 1901 he joined the staff of the National Physical Laboratory and remained there until 1919 as Principal Assistant in charge of the Electrotechnical and Photometric departments. In 1919 he took a step which had been under discussion with the officials of the Osram Lamp Works and the General Electric Co. for two or three years but which had been held up because of the war. The war being ended he left the N.P.L. and threw himself into what was to be his life's work, the planning, erection, equipment, and development of the General Electric Company's Research Laboratories at Wembley. Although originally conceived as a subsidiary of the Osram Lamp Works, it soon became the research organization of the whole field of General Electric activities. The early work on lamps naturally led to work on valves and much of the progress that has been made in both fields during the last thirty years has been due to the research carried out at Wembley.

Paterson's personal interest was always directed towards lamps and illumination and it is interesting to note that from 1939 to 1941 he was Master of the Worshipful Company of Tallow Chandlers. He was President of the Illuminating Engineering Society in 1928, of the I.E.E. in 1930-1931 and of the Institute of Physics in 1937-1938. In 1939 he was awarded the Honorary Degree of D.Sc. by Birmingham University and in 1942 he was made a Fellow of the Royal Society. Anyone who heard him lecture will understand why he was so much in request as a lecturer, the lectures and demonstrations being always meticulously prepared. Under his inspiring guidance the laboratories have extended their activities until the staff, which in 1919 numbered about half a dozen, now exceeds a thousand. He will be long remembered for his personal qualities and for his great achievement in the co-ordination of science and industry at the Wembley Laboratories.

G. W. O. H.

NEGATIVE-GRID PARTITION NOISE

By R. L. Bell, B.Sc.

(*King's College, Durham University*)

1. Introduction

IN the conventional negative-grid triode the electrostatic field about the grid is not uniform, owing to the discrete nature of the grid wires. It is of interest in the investigation of valve noise to determine what effect this departure from the ideal grid will have on the fluctuations induced at the grid.

It is clear that electrons taking different paths through such a field will in general induce differently-shaped current pulses in the electrodes maintaining the field, and that in the triode, an indefinitely large number of such paths is possible, since electrons are emitted from the cathode with random uncorrelated velocities and directions.

An extension of the Campbell mean-square theorem is used here to show that, subject to the assumptions of Section 2, this variation of pulse shape will give rise to an ultra-high-frequency noise effect analogous to partition effect in screen-grid valves.

Depending as it does on the variation of field configuration in planes parallel to the valve electrodes, the magnitude of the partition effect will in general be greater for large values of the ratio of grid-winding pitch to grid-cathode clearance.

2. Basic Theory and Assumptions

We are concerned with the calculation of mean-square noise fluctuations due to space-charge-limited currents (space charge being introduced specifically in order to show that it has not first order effect on the partition effect), and shall take as fundamental Campbell's mean-square theorem[1], and some modified version of the Thompson-North theory of space-charge reduction of shot effect.[2]

Campbell's theorem states that if a quantity $y(t)$ is the resultant of a large number of randomly-occurring events of identical nature $f(t)$, generally assumed zero for negative values of time t, which have a mean rate of occurrence ν, then the mean square variation of $y(t)$ is given by

$$\overline{(y-\bar{y})^2} = \nu \int_0^\infty |f(t)|^2 dt \qquad .. \quad .. \quad (1)$$

If, further,

$$g(\omega) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t)\exp(-j\omega t)dt \qquad .. \quad (2)$$

is the Fourier transform of $f(t)$, then by a theorem of F— transformation,

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |g(\omega)|^2 d\omega = 2\int_0^\infty |g(\omega)|^2 d\omega \quad (3)$$

Hence

$$\overline{dy^2} = 2\nu|g(\omega)|^2 d\omega \qquad .. \quad .. \quad .. \quad (4)$$

The variable ω may be identified as $2\pi \times$ frequency, so that the process of F— transformation is equivalent to the derivation of frequency spectra. Equ. (4) then gives the magnitude of the mean-square contribution to the fluctuations in the frequency range $\frac{1}{2\pi}(\omega, \omega + d\omega)$.

In the following work the events (current pulses) are not identical, but depend on parameters λ and z. We have to replace $f(t)$ by $f(t, \lambda, z)$ and $g(\omega)$ by $g(\omega, \lambda, z)$. The mean rate ν is no longer constant, but becomes a two-dimensional distribution density $\nu\ (\lambda, z)$, the number of events happening in unit time with values of z in the range $z, z + dz$ and values of λ in the range $\lambda, \lambda + d\lambda$ being $\nu(\lambda, z)\, d\lambda\, dz$.

The equation (4) then takes the form

$$d^3\overline{y^2} = 2\nu(\lambda, z)|g(\omega\lambda z)|^2\, d\omega\, d\lambda\, dz \qquad .. \quad (5)$$

i.e.,

$$\frac{d\overline{y^2}}{d\omega} = 2\int_\lambda d\lambda \int_z \nu(\lambda z)|g(\omega\lambda z)|^2 dz \qquad .. \quad (6)$$

This is a consequence of the statistical central-limit theorem. Since λ and z in the following are quite independent, $\nu(\lambda z)$ can be split into two independent distributions

$$\nu(\lambda z) = \nu_1(\lambda)\ .\ \nu_2(z)$$

and (6) may be written

$$\frac{d\overline{y^2}}{d\omega} = 2\int_\lambda \nu_1(\lambda)d\lambda \int_z \nu_2(z)|g(\omega\lambda z)|^2 dz \qquad .. \quad (7)$$

This indicates that, having deduced the quantity $|g(\omega\lambda z)|^2$, it may first be averaged with respect to z, to give

$$\overline{|g(\omega\lambda z)|^2} = \overline{|g^2|}(\omega\lambda)$$

and this average value may then be integrated over values of λ to give the complete energy-density frequency spectrum of the fluctuations.

According to the elements of the Thompson-North theory, a negative charge e of initial energy λ, in transit in the interelectrode space, depresses

MS accepted by the Editor, March 1947.

the potential minimum, so that some charge which normally would have reached the anode is reflected back to the cathode, or vice versa, the net (low-frequency) result being equivalent to the transit of a charge $e \cdot \gamma(\lambda)$ from cathode to anode.

In the case of an electron which reaches the anode, it is convenient to consider the equivalent charge γe as comprising the actual point charge e and a "deficiency charge" $-e(1-\gamma)$ which, because of the above smoothing effect of the space charge, will be distributed over appreciable area perpendicular to the general direction of motion. Thus it is reasonable to suppose that the deficiency charge almost invariably induces currents in the electrodes that are in proportion to the mean currents induced by the point charges. This may be clearer from the analysis of Section 3.

Such a theory cannot rigorously be applied at transit-time frequencies, but is probably capable of a generalization which will reduce to the theory as its stands at present when transit-time effects become negligible and it will be used in that general sense, precision upon this point being unimportant.

It is also implicit in the analysis that electron movements parallel to the electrodes are assumed independent and uncorrelated. While it is not easy to demonstrate the accuracy or otherwise of this assumption, it is a great convenience in that it permits a direct application of the Campbell mean-square theorem, and will suffice in a preliminary discussion of this nature.

3. Analysis

It appears simplest to treat the problem as one of deviation from the ideal. Let λ be the normal component of initial electron energy at the cathode, in a triode in which the field configuration is constant in the transverse direction. The passage of unit charge will give rise to electrode currents depending on both initial energy and time t:

$$i_g = i_1(t, \lambda)$$
$$i_a = i_2(t, \lambda) \qquad \ldots \qquad (8)$$

Then in the case of a coarse grid and a charge e, it will be sufficiently general to write

$$i_g = e \cdot \alpha_g(t\lambda z) \cdot i_1(t\lambda)$$
$$i_a = e \cdot i_2(t\lambda) + e \cdot \alpha_a(t\lambda z) \cdot i_1(t\lambda) \qquad \ldots \qquad (9)$$

the quantity z representing position in the plane of the electrodes, since the field is no longer constant in this plane. We neglect the angles at which the electrons are emitted.

The parameters α_g and α_a which we have used to take account of the field irregularities will have mean values

$$\alpha_g(t\lambda z) = \overline{\alpha_g}(t\lambda) \; ; \quad \alpha_a(t\lambda z) = \overline{\alpha_a}(t\lambda) \qquad \ldots \qquad (10)$$

taken over the trajectories.

Since in a triode, spacings are small compared with the wavelength of oscillations handled, and electron velocities are much less than that of e.m. propagation, cathode currents are completely determined by those flowing to grid and anode, so that we may restrict attention to these without losing sight of the possible effects of cathode fluctuations. In the presence of space charge, recalling the postulations of the previous section, we can write for the effective electrode currents due to an electron reaching the anode

$$i_g = e\alpha_g i_1 - e(1-\gamma)\overline{\alpha_g} i_1$$
$$i_a = e(i_2 + \alpha_a i_1) - e(1-\gamma)(i_2 + \overline{\alpha_a} i_1) \qquad (11)$$

Now let $\psi(\omega\lambda)$ be the Fourier transform (or frequency spectrum) of the current $i_1(t\lambda)$; i.e.,

$$\psi(\omega\lambda) = F\{i_1(t\lambda)\} \qquad \ldots \qquad \ldots \qquad \ldots \qquad (12)$$

and similarly let

$$\phi(\omega\lambda) = F\{i_2(t\lambda)\}$$
$$\beta_g(\omega\lambda z)\psi(\omega\lambda) = F\{\alpha_g(t\lambda z) i_1(t\lambda)\}$$
$$\beta_a(\omega\lambda z)\psi(\omega\lambda) = F\{\alpha_a(t\lambda z) i_1(t\lambda)\}$$
$$\overline{\beta_g}(\omega\lambda)\psi(\omega\lambda) = F\{\overline{\alpha_g}(t\lambda) i_1(t\lambda)\}$$
$$\overline{\beta_a}(\omega\lambda)\psi(\omega\lambda) = F\{\overline{\alpha_a}(t\lambda) i_1(t\lambda)\} \qquad \ldots \qquad \ldots \qquad (13)$$

then

$$\overline{\beta_g(\omega\lambda z)} = \overline{\beta_g}(\omega\lambda)$$
$$\overline{\beta_a(\omega\lambda z)} = \overline{\beta_a}(\omega\lambda) \qquad \ldots \qquad \ldots \qquad \ldots \qquad (14)$$

If $G_g(\omega)$ and $G_a(\omega)$ are the transfer functions (e.g., impedance) relevant between grid and anode currents and output quantity, $y(t)$, the frequency spectrum of this output quantity is clearly given from (11) by

$$g_1(\omega\lambda z) = eG_g\{\beta_g\psi - (1-\gamma)\overline{\beta_g}\psi\} + eG_a\{\phi + \beta_a\psi - (1-\gamma)(\phi + \overline{\beta_a}\psi)\} \qquad \ldots \qquad (15)$$

or, rearranging,

$$g_1(\omega\lambda z) = eG_g\{(\beta_g - \overline{\beta_g})\psi + \gamma\overline{\beta_g}\psi\} + eG_a\{(\beta_a - \overline{\beta_a})\psi + \gamma(\phi + \overline{\beta_a}\psi)\} \qquad \ldots \qquad (16)$$

This holds for values of emission energy $\lambda_m < \lambda \leqslant \infty$, λ_m being the value of λ for which electrons just fail to pass the potential minimum. Similarly, for electrons which return to the cathode,

$$i_g = \gamma e\overline{\alpha_g} i_1, \; i_a = \gamma e(i_2 + \overline{\alpha_a} i_1) \qquad \ldots \qquad (17)$$

and

$$g_2(\omega\lambda) = e\{G_g\overline{\beta_g}\psi + G_a(\phi + \overline{\beta_a}\psi)\} \qquad \ldots \qquad (18)$$

The expressions (16) and (18) may be regarded as special cases of a more general function, defined by (18) for $0 \leqslant \lambda \leqslant \lambda_m$ and by (16) for $\lambda_m < \lambda \leqslant \infty$.

This function is our $g(\omega\lambda z)$ of Section 2. To find $|g(\omega\lambda z)|^2$, it is easiest to take the sum of the squares of real and imaginary parts, assuming each quantity in $g(\omega\lambda z)$ to have real and imaginary parts.

ratio is $Z_1/(Z_1 + Z_3)$. The condition for oscillation is clearly that the product of these complex ratios shall be equal to $-1/\mu$, a condition which may equally well be deduced by simply writing $(Z_1 + Z_3)$ large compared to Z_2 in equation (1) above.

If the amplification factor μ is positive, it is clearly not possible to satisfy this condition using normal circuit components unless phase-reversing means be added (e.g., mutual inductance, half-wave lines, etc.). With a negative amplification factor, however, it is simply a matter of choosing two voltage-divider circuits one of which displaces the phase and the second of which restores it to the original value. This can be done in a very large number of ways, the four simplest being shown in Fig. 3. Of these four, circuit 3(a) is well known, while the others are possibly novel.

Oscillators using positive-μ valves may be derived from these circuits by using the phase-reversing properties of mutual inductance. This

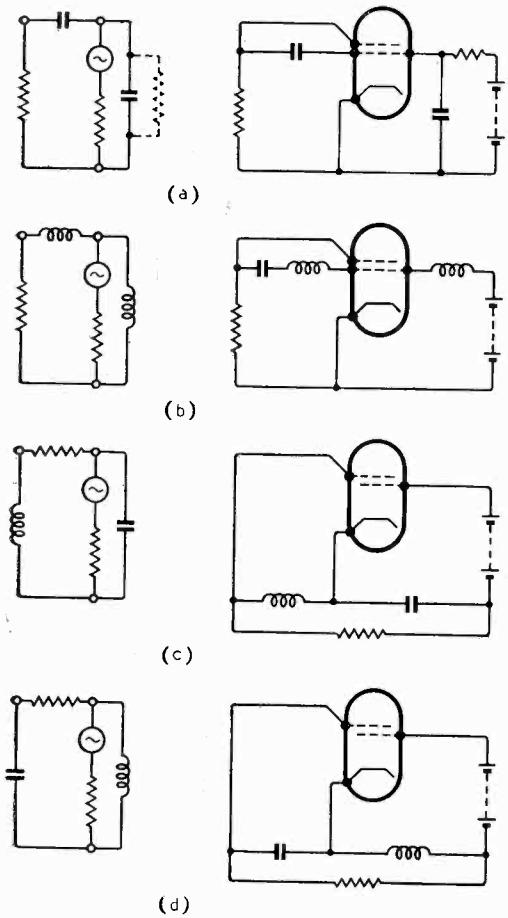


Fig. 3. *Phase-restored circuits with negative-μ; only cathode, g_2 and g_3 are shown for the transitrons.*

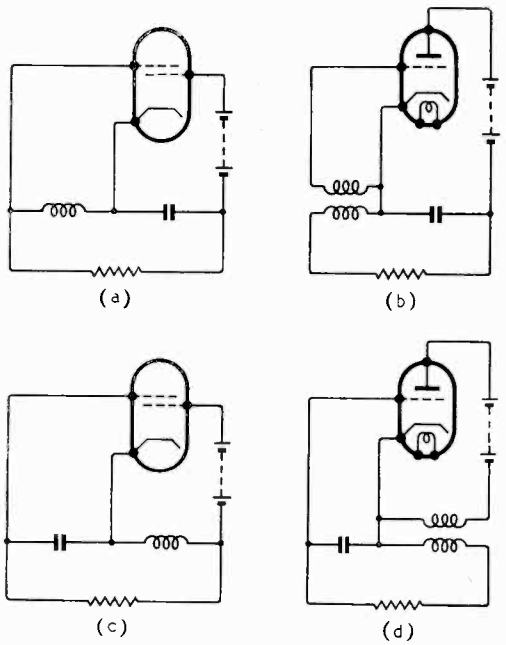


Fig. 4. *These diagrams illustrate the derivation of positive-μ from negative-μ circuits.*

is possible, of course, only where the corresponding negative-μ circuit incorporates a self-inductance. For example, circuit 3(c) has a self-inductance between cathode and control-grid. To convert this circuit to the corresponding positive-μ oscillator, it is simply necessary to make this self-inductance the primary of a transformer and to connect cathode and grid to the secondary instead of to the primary. The result is shown in Fig. 4 (a) and (b), and will be recognized as the well known tuned-anode oscillator. Similarly, as shown in Fig. 4(c) and (d), conversion of circuit 3(d) for use with positive-μ valves gives the well-known tuned-grid oscillator.

4. Zero Phase-shift Circuits

If, at one particular frequency, the three impedances (Z_1, Z_2 and Z_3) of Fig. 1 are all pure resistances, then at this frequency there is zero phase-shift between the e.m.f. μV_g and the p.d. across Z_1. This is the required condition for oscillators using negative-μ valves. The circuit may thus have resistors for two of the three impedances, Z_1, Z_2, Z_3, and the third impedance may consist of a circuit which is resistive at one particular frequency. Restricting ourselves to the use of simple tuned circuits (series or parallel) for the third impedance, this gives the six possibilities shown in Fig. 5, of which the first two, (a) and (b) are well known, and the last three (d), (e) and (f), do not give sinusoidal

oscillation, as explained later (see Relaxation Oscillation). The use of more complicated tuned circuits, piezo-electric crystals and lines makes possible a large number of further oscillator circuits of this type.

Oscillator circuits suitable for positive-μ valves may again be derived from the above by using the phase-reversing properties of mutual inductance. Though this may be an amusing exercise, the results are not very useful. It may be mentioned that the application of this procedure to circuit 5(a) gives a version of the tuned-grid oscillator, while circuit 5(b) becomes the parallel-fed version of the tuned-anode oscillator.

5. Tapped Resonant Circuits

If the three impedances of the basic circuit be made as far as possible pure reactances, and not all of the same sign, then at one frequency the sum of these reactances will be zero and $Z_1\ Z_2\ Z_3$ will constitute a closed resonant circuit on to which are tapped the anode, grid and cathode (or, to be more general, the output electrode, input or control electrode, and cathode). In this case the equation giving the condition for sustained sinusoidal oscillation may be derived by writing $(Z_1 + Z_2 + Z_3)$ equal to zero in equation (1), giving the well known result that

$$Z_2 = \mu Z_1$$

Thus for circuits of this type using positive-μ valves, Z_1 and Z_2 must be reactances of like sign (i.e., both inductances or both capacitances) and Z_3 must be a reactance of the opposite sign. This gives only two possible circuits[1] (viz., the Hartley oscillator and the Colpitts oscillator), unless more complicated reactance networks be used for Z_1 and Z_2, giving such well-known circuits as the tuned-anode tuned-grid oscillator and the Pierce crystal oscillator.

For circuits of this type using negative-μ valves, however, Z_1 and Z_2 must be reactances of unlike sign and there are therefore four basic possibilities as shown in Fig. 6, and correspondingly more possibilities if more complicated reactance networks be used. These circuits are possibly novel.

6. Relaxation Oscillation

Lest the outlook here presented should seem to encourage the rash devising of new oscillator circuits, it must be said that some oscillators evolved by the methods given are inherently not capable of sinusoidal oscillation, but function as relaxation oscillators. Whether this will be so in a particular case can often be predicted by deriving the differential equation for the oscillator in question in the form

$$ap^2 + bp + c = 0$$

For sustained oscillation, of course, the coefficient b must be zero or negative. But if b is negative *and* b^2 is greater than $4ac$, relaxation oscillation will occur.

In almost all feedback oscillators, relaxation

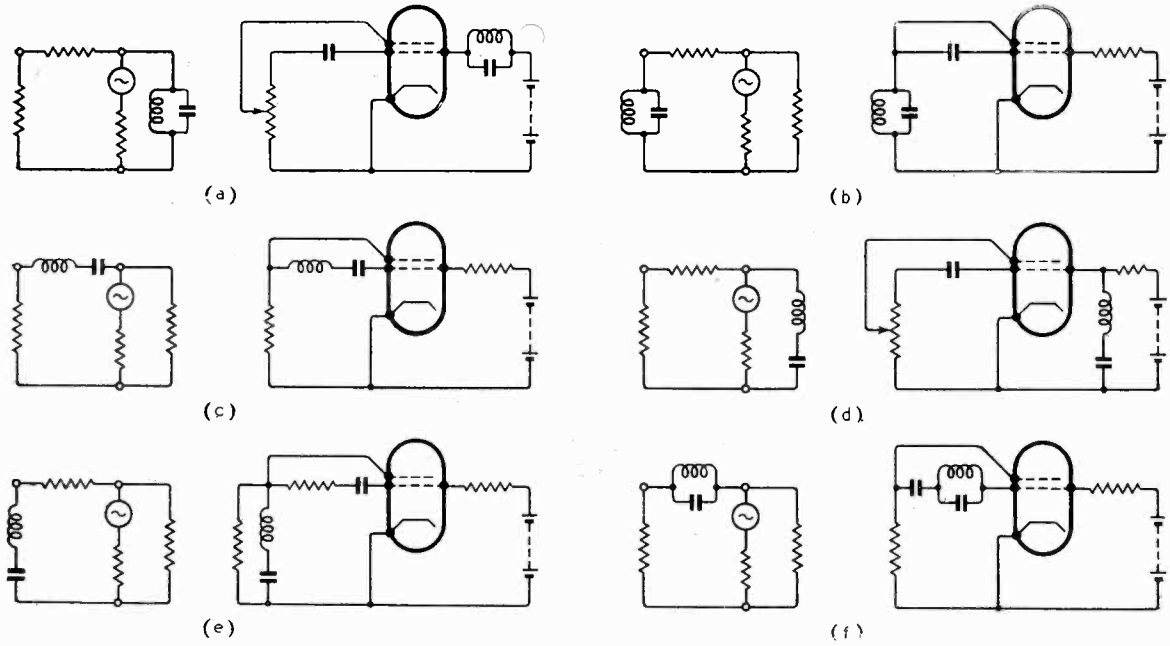


Fig. 5. *Zero phase-shift circuits with negative-μ valve.*

amount of valuable theoretical and experimental material and the accuracy of the treatment is never in question.

It is good to note that early British developments, both in triodes and in reflex oscillators, are duly acknowledged. In view of the volume of research and development subsequently achieved in U.S.A., it would have been easy for the authors to forget the CV90 and the Heil tube.

Only one error in the book proved irritating. That was the reiteration of the phrase 'data is (shown)'. Surely Massachusetts must recognise the plural of 'datum'.

J. T.

Microwave Receivers

Edited by S. N. VAN NOORHIS (M.I.T. Radiation Laboratory Series, Vol. 23.) Pp. 618 + xviii, with 420 illustrations. McGraw-Hill Book Company, Aldwych House, London, W.C.2. Price 48s. (in U.K.).

This book contains a great wealth of information on the theory, design and testing of microwave receivers. Most of the circuits considered find their main application in radar, but an attempt has been made to extend and generalize the treatment to cover other applications; and the design criteria established are sufficiently general to form the basis for receiver design in most fields where signal frequencies above 100 Mc/s, wideband i.f. amplifiers and automatic-control techniques are involved.

The treatment throughout is encyclopedic rather than didactic, and the book will therefore be more useful to the engineer and designer working actively in the field than to the student approaching it for the first time—perhaps a natural consequence of the multiple authorship. Familiarity with the subject is assumed tacitly.

The arrangement of the book falls into three natural divisions. The first twelve chapters are devoted to the units making up a complete receiver: duplexers, r.f. amplifiers, local oscillators, mixers, i.f. amplifiers, detectors, video amplifiers, a.g.c. and a.f.c. circuits; and there are also separate discussions on mechanical construction and on test gear. Chapters 13 to 17 discuss in detail the design and construction of specific receivers, starting with a general-purpose 10,000-Mc/s radar receiver, then an airborne 3,000-Mc/s radar receiver with anti-clutter circuits, a shipborne 10,000-Mc/s radar receiver with circuits for auto-following, a 200-Mc/s receiver for spot-frequency pulse reception, and an f.m. relay receiver. The remaining chapters 18 to 21 return to a more general discussion of special receiver types, including receivers with extra-wide pass-bands, crystal video receivers, and super-regenerative receivers.

The early chapters are badly arranged, a.f.c. circuits are treated immediately after a chapter on the practical design of mixers and local oscillators; then follows the chapter on i.f. input circuit theory, followed by one on the theory of v.h.f. amplifiers, mixers and oscillators, which is succeeded by the chapter on i.f.amplifier design.

The most glaring omission is the lack of any real treatment of receivers with wide tuning ranges, and the circuits and valves developed for them during the past few years. Altogether, the 'microwave' aspect is dismissed rather too briefly in favour of a much more detailed treatment of the i.f. and video circuits, and particularly the various automatic-control circuits. Local oscillators are treated far too skimpily and a good deal more could usefully have been included on mixers; but there is the excuse that a whole volume of the series has been devoted to mixers separately. The section on mixer theory on pp. 138–147 is particularly inadequate.

The design of v.h.f. amplifiers and i.f. input circuits for optimum noise factor is treated very well and represents a most useful summary of the scattered published material. Noise data are given for a large number of valves. An error was noticed in the formula for the noise factor of the cathode-input amplifier with transit-time loading (eqn. 14, p. 137). This should read:

$$F = 1 + \frac{5G_1}{G'_a} + \frac{2.5g_m}{(g_m + 1/r_p^2)}\left(G'_a + 2G_1 + \frac{G_1^2}{G'_a}\right),$$

and would also be clearer if the substitution $R_{eq} = 2.5/g_m$ had not been carried out at this point.

The treatment of i.f., video, and automatic-control circuits is very thorough, and many curves are given as well as full design formulæ.

The numerous illustrations are a most valuable feature of the book. Constructional points and circuit layout are fully covered at every phase with photographs and line drawings; and detailed circuit diagrams are given for all the receivers discussed.

With the limitations mentioned, the book can be thoroughly recommended to the receiver designer.

P. E. T.

Microwave Duplexers

Edited by LOUIS D. SMULLIN and CAROL G. MONTGOMERY. Pp. 437 + xiv, with 393 illustrations. (Vol. 14, M.I.T. Radiation Laboratory Series). McGraw-Hill Publishing Co., Ltd., Aldwych House, London, W.C.2. Price 39s. (in U.K.).

Of the twenty-seven volumes in the series this must have been one of the most difficult to write, combining as it does, a mass of important empirical data, a valiant attempt at a theory of the gaseous discharge, and much detailed circuit analysis. The authors are to be congratulated on having so far disentangled the various phenomena encountered in duplexers as to make a logical, if sometimes trying, study of the various aspects of the subject.

The greater part of the book is concerned with t.r. and a.t.r. switches. Since these devices were completely unknown in 1940, and indeed were 'discovered' in their modern form at Telecommunications Research Establishment in 1941, the work described bears the imprint of war development. That is, the studies are short term, designed to meet the needs of the situation quickly, and without overmuch seeking for fundamental knowledge. Nevertheless, the reader cannot but be impressed by the mass of experimentally verified theory which has emerged during this period. In that respect Chapters 1 to 4 are remarkable. In Chapter 5, which is the longest in the book, Dr. Smullin does his best with the intractable subject of the gaseous discharge at ultra-high frequencies. Having spent several laborious years attempting to elucidate similar phenomena at lower frequencies, the reviewer can sympathise with the author. Suffice it to say that enough emerged from the M.I.T. wartime research to improve considerably the quality of the switches employed. The reader is, however, left with a feeling of surprise that the devices described ever worked at all. One of the most important features of this chapter and the succeeding one is that they draw attention in some detail to the very wide field for research which is still lying fallow.

Chapters 7 and 8 return the reader to more orthodox radio engineering. The section on balanced duplexers could have been expanded to advantage. The possible importance of such devices for communication purposes warrants a closer study. There is a particularly good account of measurement techniques in the last chapter.

As in all books of this series there are remarkably few typographical errors. The practical design data should prove valuable to all who work in this field. Indeed, it is as a handbook for the latter that the volume will find its maximum utility.

J. T.

CORRESPONDENCE

Letters to the Editor on technical subjects are always welcome. In publishing such communications the Editors do not necessarily endorse any technical or general statements which they may contain.

Simplified Magnetic Amplifier

SIR,—Many of the applications of magnetic amplifiers have been concerned with computing devices where proportionality between output and input is important. There are, however, cases where a magnetic amplifier would be useful simply as a form of relay interposed between a low-power source and a robust electro-mechanical relay. In such cases proportionality is not important and one looks for means of simplifying the magnetic amplifier to something less than the conventional doubly-balanced system of four coil units which is shown in Fig. 1.

In Fig. 1 it is necessary to use two coils on each side of the system in order to balance out mutual inductance between the a.c. circuit and control circuit, and it is necessary to use two halves of the system in push-pull in order to balance out the standing current in the load in the absence of current in the control windings. It is clear that the circuit could be somewhat simplified by omitting the control winding from one-half of the circuit. This still uses four magnetic cores, but with the number of windings slightly reduced, and the polarizing windings can be omitted from both sides of the amplifier.

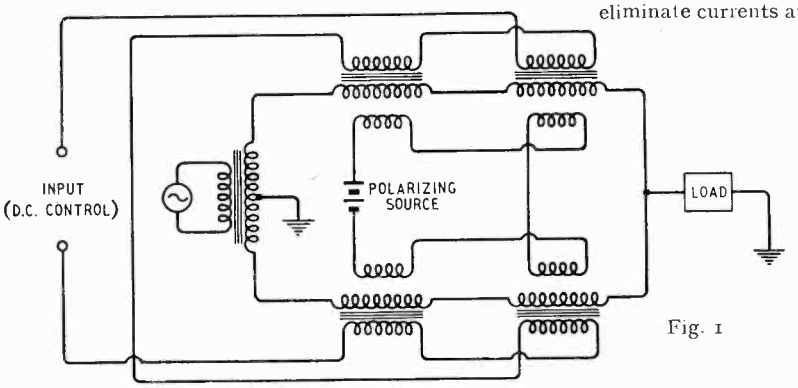


Fig. 1

In normal use, both sides of the amplifier are biased by polarizing current to about the mid-point of the characteristic of the core so that application of a control current causes the alternating current to increase in one side of the amplifier and decrease in the other.

But by using one side merely to balance the standing current for zero signal and reducing the impedance of the other side by the application of a signal, one can eliminate the whole of the polarizing system and two of the control windings. (Fig. 2.)*

The next step is to endeavour to eliminate the pair of inductances which are now used simply for balancing purposes in the absence of a signal. The obvious step is to replace these by a capacitance fed in the opposite phase (i.e., in the same phase as the controlled windings) and by doing so one also eliminates the need for a balanced input supply. There are still two coils so that the mutual inductance between control and a.c. windings can still be balanced out.

* This system has been recommended by A. S. Fitzgerald, *J. Franklin Inst.*, Oct. 1947, Vol. 244, p. 249.

Some further improvement can be effected by introducing a balance for the resistive component of the coils in the absence of signal. The circuit with capacitive balance can be converted to the familiar bridged-T circuit by the addition of one resistance as shown in

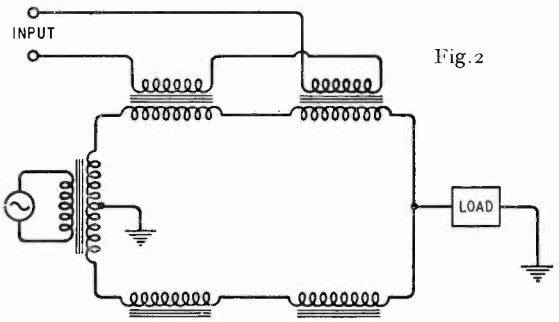


Fig. 2

Fig. 3. Such a filter will have infinite impedance only at a fixed frequency and at a fixed amplitude of the input voltage. In addition to these limitations it must be remembered that the bridged-T network will not eliminate currents at harmonic frequencies which may be generated by the non-linearity of the coils. The balance is therefore by no means perfect, but it does provide a simplified form of magnetic amplifier which can be used for operating an electro-mechanical relay. It is, of course, possible to apply feedback from rectified output to the control circuit or to separate feedback windings. A feature of such a circuit with feedback is that the output current has a constant polarity depending solely on the connection of the rectifier which is used to convert the a.c. output to d.c. Since

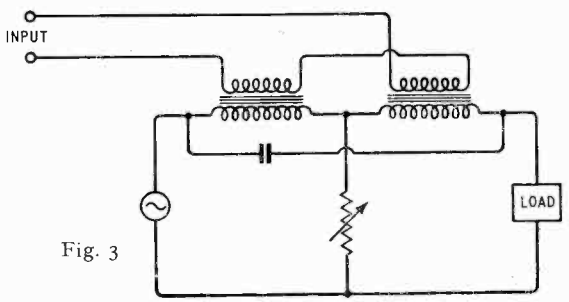


Fig. 3

there is no initial polarization of the cores, a direct current of either polarity in the control windings will lead to increased output. With one sign of input the feedback derived from the rectified output will aid the input, and with the other sign it will oppose it. Therefore

Hazeltine Corporation (*assignees of H. A. Wheeler*). *Convention date* (*U.S.A.*) *15th August*, 1944.

593 461.—Circuit for producing square or stepped pulses of predetermined duration, as used say for marking or strobing in radiolocation.

Standard Telephones and Cables Ltd. (*assignees of A. Priesman*). *Convention date* (*U.S.A.*) *27th February*, 1943.

593 467.—Phasing means for ensuring the correct operation of the transmitter and receiver circuits in a radiolocation set which is energized by an unrectified a.c. supply.

Standard Telephones and Cables Ltd. (*assignees of E. Labin*). *Convention date* (*U.S.A.*) *23rd January*, 1943.

593 491.—Time-base circuit for offsetting the distortion due to the variation of slant-range with height, when presenting plan-position indications in radiolocation.

F. C. Williams. Application date 21st February, 1945.

593 510.—Radiolocation set with a p.p.i. indicator in which provision is made to adjust the synchronization of the time base in order to prevent " centre " distortion.

F. C. Thompson. Application date 20th July, 1945.

593 539.—Variable delay device, arranged to simulate the echo-signals in radiolocation, and so serving to check the performance of the radio installation.

A. E. Solley and F. E. J. Girling. Application date 2nd June, 1945.

593 661.—Radiolocation receiver, having a wide pass-band, wherein a derived heterodyne component is utilized to eliminate the effect of strong interference.

Hazeltine Corpn. (*assignees of D. B. Hoisington*). *Convention date* (*U S.A.*) *15th August*, 1944.

593 662.—Radiolocation receiver for detecting pulsed signals which have 'bi-directional' characteristics; i.e., show positive or negative amplitudes relatively to a given zero.

Hazeltine Corpn. (*assignees of D. B. Hoisington*). *Convention date* (*U S.A.*) *15th August*, 1944.

RECEIVING CIRCUITS AND APPARATUS

(*See also under Television*)

592 628.—Receiver designed, under simplified switch-control, for the alternative detection of frequency-modulated or amplitude-modulated signals, on a number of wavebands.

Marconi's W.T. Co. Ltd. (*assignees of F. C. Everett*). *Convention date* (*U.S.A.*) *27th January*, 1944.

592 659.—Preparation and processing of crystal contacts, of the germanium type, for use as short-wave mixers or detectors.

The General Electric Co. Ltd., D. E. Jones and J. W. Ryde. Application date 16th July, 1941.

593 361.—Super-regenerative receiver in which the inherent variation in quenching frequency with signal amplitude is utilized to rectify frequency-modulated or amplitude-modulated signals.

Marconi's W.T. Co. Ltd. (*assignees of W. R. Koch*). *Convention date* (*U.S.A.*) *25th June*, 1942.

TELEVISION CIRCUITS AND APPARATUS

FOR TRANSMISSION AND RECEPTION

592 581.—Non-linear amplifier for controlling picture-contrast, or for separating the picture signals from the synchronizing pulses, in television.

Farnsworth Television and Radio Corpn. Convention date (*U.S.A.*) *29th March*, 1943.

592 741.—Receiver for a television system in which the sound signals are interposed between successive lines of the picture signals.

Standard Telephones and Cables Ltd. (*assignees of N. H. Young, Junr.*). *Convention date* (*U.S.A.*) *12th June*, 1944.

TRANSMITTING CIRCUITS AND APPARATUS

(*See also under Television*)

592 997.—Resistance-capacitance-coupled amplifier for generating a sinusoidal output, which is balanced with respect to the common supply point, and free from phase-drift.

Muirhead & Co. Ltd. and J. A. B. Davidson. Application date 5th June, 1945.

593 137.—Generating a continuous carrier-wave by first pulsing a valve-oscillator and then smoothing-out the pulse-frequency modulation.

Marconi's W.T. Co. Ltd. (*assignees of C. W. Hansell*). *Convention date* (*U.S.A.*) *29th December*, 1942.

593 20_2.—Mounting and controlling the probe or loop device used for coupling a waveguide or other h.f. resonant chamber to an external circuit.

Western Electric Co. Inc. Convention date (*U.S.A.*) *27th April*, 1944.

593 215.—Construction and mounting of the movable probe used for plotting and measuring the standing-wave pattern inside a waveguide.

The British Thomson-Houston Co. Ltd., L. W. Brown and H. B. Taylor. Application date 25th June, 1945.

SIGNALLING SYSTEMS OF DISTINCTIVE TYPE

592 796.—System in which separate groups of signals are transmitted with a fixed time interval between each group, and represent different symbols according to the number of elements contained in each group.

The National Cash Register Co. Convention date (*U.S.A.*) *16th September*, 1942.

592 804.—Receiver designed for a pulsed signalling system and arranged to deliver an output signal only in response to a predetermined coincidence of applied signals.

E. W. Anderson and T. J. McDermott. Application date 16th February, 1945.

592 835.—Receiver for detecting the amplitude variations, and also any reversals of phase, or changes in sense, of the signal voltage applied to a carrier wave.

E. A. Johnson. Application date 2nd June, 1945.

593 101.—Time-modulating a train of pulses by passing them through a retardation network comprising inductances, the permeability of which is varied by the signal voltage.

Standard Telephones and Cables Ltd. (*assignees of E. Labin*). *Convention date* (*U.S.A.*) *24th July*, 1944.

593 482.—Time-delay device, comprising a pair of gas-filled valves, connected to a relaxation capacitor for phase or time modulating a series of signalling pulses.

Standard Telephones and Cables Ltd., P. K. Chatterjea and L. W. Houghton. Application date 3rd January, 1945.

593 511.—Means for reducing interference when receiving pulsed signals which are modulated either in time or phasing.

Standard Telephones and Cables Ltd. (*assignees of D. D. Grieg*). *Convention date* (*U.S.A.*) *23rd November*, 1943.