THE RADIO AND ELECTRONIC ENGINEER

The Journal of the British Institution of Radio Engineers

FOUNDED 1925 INCORPORATED BY ROYAL CHARTER 1961

"To promote the advancement of radio, electronics and kindred subjects by the exchange of information in these branches of engineering."

VOLUME 25 APRIL 1963 NUMBER 4

NEGATIVE FREQUENCY

In its less abstract contexts, frequency is always regarded as a quantity devoid of sign; crudely, one might say that a sine-wave looks much the same whether the phase increases at ω rad/s or at $-\omega$ rad/s. One might alternatively say that as $\cos(-\omega t) = \cos \omega t$ and $\sin(-\omega t) = -\sin \omega t$, the distinction between positive and negative frequency is trivial, amounting at the most to a difference of phase. Or one might think of the vector representation of $\cos \omega t$ or $\sin \omega t$ as

$$\tfrac{1}{2}[\exp(+j\omega t) \pm \exp(-j\omega t)]$$

and say that both positive and negative frequencies exist together in any wave. Ignoring thus the sign of a frequency is undoubtedly justified in the vast majority of practical circumstances.

Practical problems: polyphase modulation. There are practical circumstances where some use can be made of what at first sight looks like a distinction between positive and negative frequencies. In a heterodyne receiver, for example, the intermediate-frequency signal has a frequency (ω_2) equal to the difference between the frequencies of the incoming signal and the local oscillator, ω_1 and ω_0 respectively. Thus $\omega_2 = \omega_1 - \omega_0$ say. But a signal in the "image channel" at frequency $\omega_0 - \omega_2$ also gives an i.f. of ω_2. Of course, $(\omega_0 - \omega_2) - \omega_0 = -\omega_2$, but normally no advantage can be taken of the idea that the i.f. is of "opposite sign" for the wanted signal and the image channel respectively. This conception is indeed meaningless in a normal circuit.

If the heterodyne process is carried out in a polyphase modulator, however, it is easily shown[1] that the sequence of the polyphase i.f. signal is opposite for the wanted and the image channels; and since sequence discriminators are easily made, the image channel can be rejected on this basis. It is tempting to think of this as discriminating between negative and positive frequency, and to confuse the direction of sequence with the sign of frequency; although there may be some practical justification for this, there is no theoretical justification.[2]

Theoretical problems: Circuits with time-varying parameters. In the analysis of linear circuits with time-varying parameters[3]—forming the basis of modulators and parametric amplifiers—there is considerable advantage in writing a modulation product frequency as $\omega_q + m\omega_p$, where ω_q is the signal frequency, ω_p is the local "carrier" or "pump" frequency, and m is an integer which may be positive or negative or zero. With this notation, the frequency obviously becomes negative in most likely circumstances whenever negative values of m are taken. The more obvious alternative notation, $n\omega_p \pm \omega_q$, where n is a positive integer or zero, has the small advantage that the frequency is positive in almost all likely circumstances, but has the very serious disadvantage that phase angles associated with the current or voltage at a particular frequency become reversed whenever a difference-frequency interaction product is formed. This leads to complication in solving the equations.[4]

Since in the analysis, a negative frequency $(\omega_q - n\omega_p)$ will be assumed to have a corresponding circuit impedance Z_{-n}, it is important to consider what is the relation of this impedance to the physical or measurable impedance Z_{n-} which the practical circuit has at the scalar (or positive) frequency $(n\omega_p - \omega_q)$. Alternatively, the latter may be specified in a particular problem, and we need to know what to insert in the equations.

It is known from Foster's Reactance Theorem[5] that reactance may always be expressed as an odd function of frequency, so that we may write $Z_{-n} = R_{-n} + jX_{-n} = R_{-n} + j\omega_{-n} f(\omega^2_{-n})$ where ω_{-n} is

a shortened notation for $(\omega_q - n\omega_p)$ and $f(\omega^2_{-n}) = f(\omega^2_{n-})$. Thus $Z_{-n} = R_{n-} - j\omega_{n-} f(\omega^2_{n-}) = Z_{n-}{}^*$ so that it is merely necessary to use the conjugate of the practical impedance in the theoretical analysis.

The band-pass/low-pass transformation. The idea that a low-pass filter of cut-off frequency ω_c behaves exactly as a band-pass filter of pass-band extending from a negative frequency $(-\omega_c)$ to a positive frequency ω_c is now widely used. It leads, for example, to the very useful result [6] that the response of an "ideal" filter (i.e. one with zero loss between cut-off frequencies ω_1 and ω_2 and infinite loss elsewhere, and with constant delay-time τ in the pass-band) to a suddenly-applied excitation at frequency $\frac{1}{2}(\omega_1 + \omega_2)$—i.e. a unit step envelope—is

$$\left\{\frac{1}{2} + \frac{1}{\pi}\mathrm{Si}[\tfrac{1}{2}(\omega_2 - \omega_1)(t-\tau)]\right\}\cos \tfrac{1}{2}(\omega_1 + \omega_2)t$$

where Si is the sine-integral function. When ω_1 and ω_2 are both positive frequencies, the filter is band-pass. When $\omega_1 = -\omega_2 = -\omega_c$, the filter is low-pass, $\frac{1}{2}(\omega_1 + \omega_2) = 0$, and the applied excitation is unidirectional (i.e. d.c.); the output is then

$$\frac{1}{2} + \frac{1}{\pi}\mathrm{Si}[\omega_c(t-\tau)].$$

To obtain this useful relationship, we have made use once more of the concept of negative frequency, and imagine a kind of image frequency-band below zero frequency. But again it is an abstraction; no measurements can be made in such a band.

Complex frequencies. It is well-known that the transfer operational function (T) of a two-terminal-pair network (in its most general form) may be written as the ratio of two polynomials in the operator p (according to the Heaviside operational calculus), e.g. $T = A(p)/B(p)$. (A similar expression in the variable s is used in the Laplace calculus.) The response of the network to any given excitation is determined by the values of the roots of $A(p) = 0$ and $B(p) = 0$, i.e. by the zeros and poles of the function T.

Now, in general, the roots of the polynomials are of the form $(\alpha + j\omega)$ or $(\alpha - j\omega)$. For example, in a circuit involving R, L and C in series, we obtain the roots of $B(p)$ as

$$p_1 = -\frac{R}{2L} + j\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \qquad \text{and} \qquad p_2 = -\frac{R}{2L} - j\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}.$$

Of course, in the network response each of these values of p occurs in the index of an exponential, and the $\exp(+j\omega t)$ and $\exp(-j\omega t)$ usually combine to give $\cos \omega t$ or $\sin \omega t$, so that the $\exp(\alpha t)$ represents the decay of a transient. But in advanced network theory it is found analytically convenient to regard the $(\alpha \pm j\omega)$ as corresponding to complex frequencies $(\pm\omega - j\alpha)$, and the location of these roots is plotted on what is called the "complex frequency plane". Thus, not only are negative as well as positive real frequencies used, but also imaginary frequencies. But this is an extreme type of mathematical abstraction, and the concept of a decay envelope as the waveform of an imaginary frequency is not likely to be of much value in visualizing circuit performance.

Conclusions. The concept of negative frequency (and even of complex frequency) has wide theoretical application. It must be accepted as an abstract concept, however, although in certain circumstances the distinction of sequence in a polyphase system may appear to correspond to a distinction of sign in the frequency.

D. G. TUCKER.

References

1. I. F. Macdiarmid and D. G. Tucker, "Polyphase modulation as a solution of certain filtration problems in telecommunication", *Proc. Instn Elect. Engrs*, 97, Part III, p. 349, 1950.
2. See published discussion by G. B. Madella and N. F. Barber, *Wireless Engr*, 24, p. 310, 1947, and 25, p. 98, 1948.
3. D. G. Tucker, "Circuits with time-varying parameters", *J. Brit.I.R.E.*, 25, No. 3, p. 63, March 1963.
4. See, for example, D. G. Tucker, "Modulators and Frequency-Changers", p. 201 (Macdonald, London, 1953).
5. See, for example, E. A. Guillemin, "Communication Networks", Vol. 2, p. 184 (Wiley, New York, 1935).
6. E. A. Guillemin, *loc. cit.*, and many other books.

U.D.C. 621.375.43

Transistor Stabilization Matrix

By

W. D. RYAN, Ph.D.†

Summary: In the design of transistor networks where the bias circuits may involve several direct-coupled stages, it is sometimes difficult to obtain a simple indication of the effect of the various elements involved on the bias stability of the system. The stabilization matrix is proposed as an extension of the concept of the stability factor and examples are given of applications to several simple circuits. A procedure for calculating the stability matrices is suggested in Appendices.

1. Introduction

Satisfactory operation of the transistor as a linear amplifier depends on the maintenance of the desired operating point. This is a function not only of the supply voltages and the circuit element values but also of the internal transistor parameters. The primary source of these parameter variations is usually changes in junction temperature, due in turn to changes in dissipation or ambient temperature. Since the effect of junction temperature on all the parameters with the exception of the collector saturation current I_{co} and the base-emitter voltage V_{be} is irregular and difficult to predict, it is considered satisfactory to describe the change in operating point in terms of I_{co} and V_{be} only. Indeed, in many cases where an indication only is required, I_{co} alone need be considered. It is also customary to neglect the transistor ohmic resistances, although if necessary they may be included in series with each of the electrode terminals. Nevertheless, it is still possible to predict with reasonable accuracy the effect of junction temperature on the operating point of a single transistor stage.

The effect of I_{co} and V_{be} on the operating point of a transistor may be described by the series of stability factors shown in Table 1. It can be seen that each stability factor merely relates the change in the transistor current to the change in the variable.

Table 1

Single-stage Transistor Current Stability Factors

I_{co}-derived	
Emitter current stability factor	$S_e = \partial I_e/\partial I_{co}$
Base current stability factor	$S_b = \partial I_b/\partial I_{co}$
Collector current stability factor	$S_c = \partial I_c/\partial I_{co}$
V_{be}-derived	
Emitter current stability factor	$S'_e = \partial I_e/\partial V_{be}$
Base current stability factor	$S'_b = \partial I_b/\partial V_{be}$
Collector current stability factor	$S'_c = \partial I_c/\partial V_{be}$

A similar series describing changes in the transistor voltages may be defined, but will not be considered here. Such factors may be found using the techniques of linear network analysis. A method of calculating the stability factors, applicable to most isolated transistor stages, is given in Appendix 1.

From Table 1 it can be seen that if the change in I_{co} is ΔI_{co} and the change in V_{be} is ΔV_{be}, then the emitter current changes by

$$\Delta I_e = S_e \,.\, \Delta I_{co} + S'_e \,.\, \Delta V_{be} \qquad \ldots\ldots(1)$$

If the changes in I_{co} and V_{be} are due to temperature, the change in I_e may be estimated by replacing ΔI_{co} and ΔV_{be} by $(\partial I_{co}/\partial T).\Delta T$ and $(\partial V_{be}/\partial T).\Delta T$ respectively, where the rates of change of I_{co} and V_{be} are evaluated at a suitable point in the temperature range under consideration. This may be done theoretically, V_{be} falling linearly and I_{co} increasing exponentially with increase with temperature. For approximate calculations these may be taken to be $-2{\cdot}1$ mV/deg C and $+(0{\cdot}10)I_{co}$/deg C for germanium and $-1{\cdot}8$ mV/deg C and $+(0{\cdot}16)I_{co}$/deg C for silicon.

2. Stabilization Matrix

There are many advantages to be gained by using direct coupling in transistor amplifiers even when a response is not required at zero frequency. There may be a considerable economy in components and low-frequency instability problems may be reduced. There is also the possibility of obtaining bias stabilization by the application of strong d.c. negative feedback. This may not always be as effective as it may seem at first and some method of assessing the overall bias stability of a direct-coupled system would appear to be required. This may be done by introducing a factor to take account of neighbouring transistors.‡ However, such a method may have limitations. It may be necessary to assume that all junctions suffer the same change in temperature. A measure of the influence of each stage on the others may be difficult to achieve. It would appear to be better to introduce in these cases an extension of the concept of the stability factor which would relate all the changes in I_{co} and

† Department of Light Electrical Engineering, The Queen's University of Belfast, Northern Ireland.

‡ R. B. Hurley, "Junction Transistor Electronics", Chapter 5. (Wiley, New York, 1958).

V_{be} to each of the transistor currents in turn. In this case each stability factor would become a stability matrix, since the changes in, say, the emitter currents would now be given by,

$$\begin{aligned}\Delta I_{e_1} = &\ S_{e_{11}} \cdot \Delta I_{co_1} + S_{e_{12}} \cdot \Delta I_{co_2} + \ \ldots . \\ &+ S'_{e_{11}} \cdot \Delta V_{be_1} + S'_{e_{12}} \cdot \Delta V_{be_2} + \ \ldots . \\ \Delta I_{e_2} = &\ S_{e_{21}} \cdot \Delta I_{co_1} + S_{e_{22}} \cdot \Delta I_{co_2} + \ \ldots . \\ &+ S'_{e_{21}} \cdot \Delta V_{be_1} + S'_{e_{22}} \cdot \Delta V_{be_2} + \ \ldots . \\ &\text{etc.} \end{aligned} \qquad \ldots\ldots(2)$$

where $S_{e_{nm}}$ and $S'_{e_{nm}}$ relate the change in I_{co_m} and V_{be_m} respectively of transistor m to the change in the emitter current of transistor n. In matrix form this becomes,

$$\begin{bmatrix} \Delta I_{e_1} \\ \Delta I_{e_2} \\ . \end{bmatrix} = \begin{bmatrix} S_{e_{11}} & S_{e_{12}} & . \\ S_{e_{21}} & S_{e_{22}} & . \\ . & . & . \end{bmatrix} \times \begin{bmatrix} \Delta I_{co_1} \\ \Delta I_{co_2} \\ . \end{bmatrix} + \begin{bmatrix} S'_{e_{11}} & S'_{e_{12}} & . \\ S'_{e_{21}} & S'_{e_{22}} & . \\ . & . & . \end{bmatrix} \times \begin{bmatrix} \Delta V_{be_1} \\ \Delta V_{be_2} \\ . \end{bmatrix} \qquad \ldots\ldots(3)$$

or,

$$[\Delta I_e] = [S_e] \times [\Delta I_{co}] + [S'_e] \times [\Delta V_{be}] \qquad \ldots\ldots(4)$$

which has exactly the same form as for a single isolated stage. The determination of the stability matrices is not much more difficult than the stability factors. Appendix 2 indicates one way of doing so which is applicable to many circuits.

Having derived the stability matrices for one of the transistor currents, in this case the emitter currents I_{e_n}, the matrices for the other currents I_{c_n} and I_{b_n} may be obtained. As the following analysis shows, the expressions again have the form as for the single stage. For one stage we have

$$I_c = \alpha I_e + I_{co}$$

$$\frac{\partial I_c}{\partial I_{co}} = \alpha \frac{\partial I_e}{\partial I_{co}} + 1$$

$$S_c = \alpha S_e + 1 \qquad \ldots\ldots(5)$$

Similarly,

$$\frac{\partial I_c}{\partial V_{be}} = \alpha \frac{\partial I_e}{\partial V_{be}}$$

$$S'_c = \alpha S'_e \qquad \ldots\ldots(6)$$

Also,

$$I_b = (1-\alpha) I_e - I_{co}$$

$$\frac{\partial I_b}{\partial I_{co}} = (1-\alpha) \frac{\partial I_e}{\partial I_{co}} - 1$$

$$S_b = (1-\alpha) S_e - 1 \qquad \ldots\ldots(7)$$

and

$$\frac{\partial I_b}{\partial V_{be}} = (1-\alpha) \frac{\partial I_e}{\partial V_{be}}$$

$$S'_b = (1-\alpha) S'_e \qquad \ldots\ldots(8)$$

With direct coupled stages

$$I_{c_n} = \alpha_n I_{e_n} + I_{co_n}$$

$$\frac{\partial I_{c_n}}{\partial I_{co_m}} = \alpha_n \frac{\partial I_{e_n}}{\partial I_{co_m}} + \delta_{nm}$$

$$S_{c_{nm}} = \alpha_n S_{e_{nm}} + \delta_{nm} \qquad \ldots\ldots(9)$$

where $\delta_{nm} = 0$ if $n \neq m$

and $\delta_{nm} = 1$ if $n = m$.

Thus the stabilization matrix for the collector currents becomes

$$\begin{bmatrix} S_{c_{11}} & S_{c_{12}} & . \\ S_{c_{21}} & S_{c_{22}} & . \\ . & . & . \end{bmatrix} = \begin{bmatrix} \alpha_1 & 0 & . \\ 0 & \alpha_2 & . \\ . & . & . \end{bmatrix} \times \begin{bmatrix} S_{e_{11}} & S_{e_{12}} & . \\ S_{e_{21}} & S_{e_{22}} & . \\ . & . & . \end{bmatrix} + \begin{bmatrix} 1 & 0 & . \\ 0 & 1 & . \\ . & . & . \end{bmatrix}$$

or

$$[S_c] = [\alpha] \times [S_e] + [1] \qquad \ldots\ldots(10)$$

where $[\alpha]$ represents the diagonal α-matrix and [1] the unit matrix.

Similarly,

$$[S'_c] = [\alpha] \times [S'_e] \qquad \ldots\ldots(11)$$

$$[S_b] = [1-\alpha] \times [S_e] - [1] \qquad \ldots\ldots(12)$$

$$[S'_b] = [1-\alpha] \times [S'_e] \qquad \ldots\ldots(13)$$

where $[1-\alpha]$ represents the diagonal matrix,

$$\begin{bmatrix} 1-\alpha_1 & 0 & 0 \\ 0 & 1-\alpha_2 & 0 \\ 0 & 0 & 1-\alpha_3 \end{bmatrix}$$

3. Examples

As a simple example, consider the feedback pair circuit shown in Fig.1.† With the values given, the stability matrices, derived by the method of Appendix 2, become

$$[S_e] = \begin{bmatrix} +3{\cdot}52 & +0{\cdot}55 \\ -6{\cdot}66 & +0{\cdot}68 \end{bmatrix}$$

$$[S'_e] = \begin{bmatrix} -0{\cdot}188 & -0{\cdot}037 \\ +0{\cdot}268 & -0{\cdot}045 \end{bmatrix} \text{ mA/V}$$

$$[S_c] = \begin{bmatrix} +4{\cdot}45 & +0{\cdot}54 \\ -6{\cdot}53 & +1{\cdot}60 \end{bmatrix}$$

† W. D. Ryan, "Tuned transistor audio amplifier", *Electronic Engineering*, **31**, p. 103, February 1959 and **33**, p. 675, October 1961.

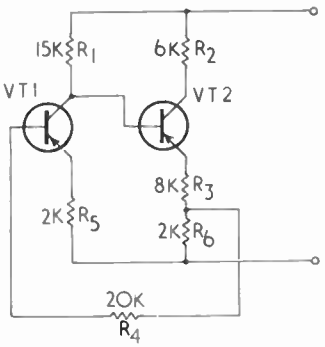


Fig. 1. Transistor feedback pair.

$$[S_c'] = \begin{bmatrix} -0{\cdot}184 & -0{\cdot}036 \\ +0{\cdot}263 & -0{\cdot}044 \end{bmatrix} \text{mA/V}$$

$$[S_b] = \begin{bmatrix} -0{\cdot}925 & +0{\cdot}012 \\ -0{\cdot}133 & -0{\cdot}986 \end{bmatrix}$$

$$[S_b'] = \begin{bmatrix} -4{\cdot}00 & -0{\cdot}97 \\ +5{\cdot}36 & -0{\cdot}90 \end{bmatrix} \mu\text{A/V}$$

Using measured values of α and I_{co}, the change in collector current with temperature of each transistor may be predicted. This has been shown in Fig. 2, where it has been assumed that I_{co} increases by a factor of eight times for each 20 deg C increase in temperature, a fair approximation to the theoretical value. Measured values are shown for comparison. If both junctions experience the same change in temperature, inspection of the stability matrices shows that it is parameter changes in transistor VT1 which have the major influence on the change in the operating points of both transistors. The matrices indicate that a considerable improvement may be obtained by replacing VT1 by a silicon unit. This can be done because of the small contribution from the changes in V_{be} of VT1. If this were large, little advantage would be gained since the V_{be} changes are of comparable magnitude for both silicon and germanium.

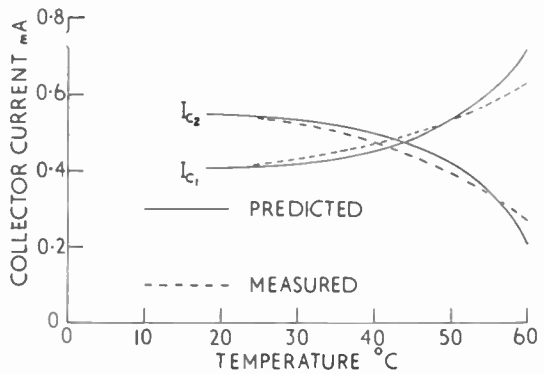


Fig. 2. Temperature stability of feedback pair.

The principal value in the use of the stability matrices would appear to be derived from the additional insight they permit into the bias interrelationships in multi-stage direct-coupled transistor amplifiers. Consider the circuit shown in Fig. 3.

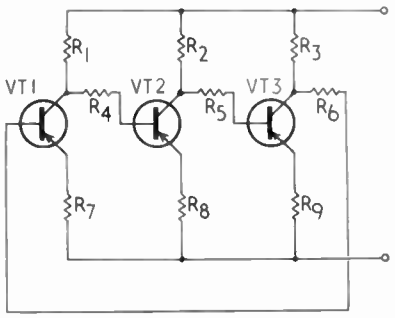


Fig. 3. Transistor ring-of-three.

This shows a d.c. negative feedback "ring-of-three" transistor cascade. If $R_1 = R_2 = R_3 = 5000\ \Omega$ and $R_4 = R_5 = R_6 = 250\ \text{k}\Omega$ and $R_7 = R_8 = R_9 = 2000\ \Omega$, the stability factors of the symmetrical circuit become,

$$[S_e] = \begin{bmatrix} +26{\cdot}8 & +13{\cdot}2 & -19{\cdot}2 \\ -19{\cdot}2 & +26{\cdot}8 & +13{\cdot}2 \\ +13{\cdot}2 & -19{\cdot}2 & +26{\cdot}8 \end{bmatrix}$$

and

$$[S_e'] = \begin{bmatrix} -0{\cdot}106 & -0{\cdot}050 & +0{\cdot}073 \\ +0{\cdot}073 & -0{\cdot}106 & -0{\cdot}050 \\ -0{\cdot}050 & +0{\cdot}073 & -0{\cdot}106 \end{bmatrix} \text{mA/V}$$

for the emitter currents, where it has been assumed that $\alpha = 0{\cdot}98$ for each transistor. It should be noted that the circuit element values, which have been chosen for this and the following example, do not necessarily represent a good practical design. Here again it can be seen that the main contribution is from the collector leakage currents. The matrices also show the tendency of d.c.-coupled common emitter stages to compensate for thermal drift if all transistors have the same polarity. For example, if transistor VT1 is kept at constant temperature, then the changes to be expected in the emitter currents of transistors VT1 and VT3 are relatively small while the emitter current of VT2 increases very rapidly with increase in junction temperature.

Finally, if R_4 and R_5 are both made zero, Fig. 3 represents a three-transistor common emitter cascade with the bases directly connected to the preceding collectors except for the single d.c. feedback resistor. Assuming that $\alpha = 0{\cdot}98$ for each transistor and

$R_1 = R_9 = 7500\,\Omega$, $R_2 = R_8 = 4500\,\Omega$, $R_3 = R_7 = 1500\,\Omega$ and $R_6 = 500\,\text{k}\Omega$, the stability matrices for the emitter currents become

$$[S_e] = \begin{bmatrix} +38{\cdot}8 & +0{\cdot}175 & -0{\cdot}184 \\ -63{\cdot}0 & +1{\cdot}34 & +0{\cdot}291 \\ +36{\cdot}6 & -1{\cdot}37 & +0{\cdot}424 \end{bmatrix}$$

and

$$[S'_e] = \begin{bmatrix} -0{\cdot}078 & -0{\cdot}014 & +0{\cdot}015 \\ +0{\cdot}123 & -0{\cdot}193 & -0{\cdot}024 \\ -0{\cdot}071 & +0{\cdot}122 & -0{\cdot}188 \end{bmatrix} \text{ mA/V}$$

Here the main contribution is due to the I_{co} changes in VT1, all the others being relatively small. In the experimental verification of the performance of this circuit, VT2 approached a cut-off condition for junction temperatures only slightly above ambient and the circuit must be considered impractical unless VT1 has a very small value of I_{co}.

4. Conclusion

The bias stabilization matrix is presented as a possible technique in the study and design of bias systems for direct-coupled transistor amplifier stages. It should at least enable an assessment to be made of any proposed bias arrangement and may permit modifications to be made in existing circuits by indicating possible sources of change in the transistor operating points.

5. Appendix 1

Calculation of Stability Factors for a D.C.-Isolated Stage

The stability factors relate changes in the variables I_{co} and V_{be} of a transistor to the electrode currents and voltages. The change in I_{co} is assumed to be independent of the voltage between collector and base or emitter and the change in V_{be} is assumed to be independent of the base or emitter currents. Thus the system of equations defining the stability factors is linear in I_{co} and V_{be}.

It is necessary to calculate initially only one of the electrode currents since the others are related by α and I_{co} only. Similarly, the inter-electrode voltages may be found later since they are assumed to be functions of the supply voltages, the circuit element values, the electrode currents and V_{be}. Hence for a single stage only one basic equation is required. This should include V_{be} as the inter-electrode voltage, which leads to the following basic rule: "Choose a current path between fixed potential points, or a closed loop, which crosses only the base-emitter junction of the transistor". A possible current path is shown in Fig. 4. The analysis is really a branch analysis and all branch currents are shown and then expressed in terms of the chosen electrode current, in this example I_e. This gives a linear equation containing I_e, I_{co} and V_{be} only

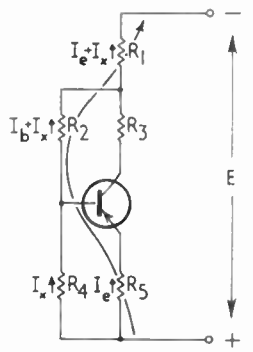


Fig. 4.
Single-stage stabilization network.

Thus, from Fig. 4, we have

$$I_eR_5+V_{be}+(I_b+I_x)R_2+(I_e+I_x)R_1 = E$$

Substituting for I_b and I_x in terms of I_e gives

$$I_e\left[R_1+(1-\alpha)R_2+R_5\left(1+\frac{R_1+R_2}{R_4}\right)\right]+$$
$$+V_{be}\left[1+\frac{R_1+R_2}{R_4}\right]-I_{co}[R_2] = E$$

from which S_e and S'_e may be obtained directly.

6. Appendix 2

Calculation of Stabilization Matrices for Direct-Coupled Transistor Stages

This is an extension of the procedure described in Appendix 1. As before, the equations are linear in I_{co} and V_{be}, and only one electrode current per transistor need be considered initially. Thus the number of equations required will generally be equal to the number of directly-coupled transistors, each equation being obtained by summing the voltage drops along an independent current path which crosses transistor base-emitter junctions only. For example, in Fig. 1, two possible paths could be R_5-R_4-R_3-R_1 and R_5-R_4-R_6, leading to the two equations

$$I_{e_1}R_5+V_{be_1}+I_{b_1}R_4+I_{e_2}R_3+V_{be_2}+(I_{c_1}+I_{b_2})R_1 = E$$

and

$$I_{e_1}R_5+V_{be_1}+I_{b_1}R_4+(I_{b_1}-I_{e_2})R_6 = 0$$

which, after substitution for the base and collector currents in terms of the corresponding emitter currents, yield two equations linear in I_{e_1}, I_{e_2}, V_{be_1}, V_{be_2}, I_{co_1} and I_{co_2}. These may then be differentiated partially with respect to I_{co_1}, I_{co_2}, V_{be_1}, V_{be_2} in turn and the equations so obtained may then be solved for $S_{e_{11}}$ and $S_{e_{21}}$, $S_{e_{12}}$ and $S_{e_{22}}$, $S'_{e_{11}}$ and $S'_{e_{21}}$, $S'_{e_{12}}$ and $S'_{e_{22}}$ respectively.

Manuscript received by the Institution on 11th October 1962 (Contribution No. 62).

U.D.C. 534.88: 621.396.677

Investigation of an Interaction Effect between Sound Projectors mounted in an Array

By

J. S. M. RUSBY, Ph.D.†

Presented at the Symposium on "Sonar Systems" in Birmingham on 9th–11th July, 1963.

Summary: Measurements on acoustically small, low-loss projectors in arrays in water have shown that near resonance violent changes in their electrical admittance take place. This behaviour has been investigated by computing theoretical admittance diagrams for a cruciform array and comparing these with the anomalous measured diagrams. The shape of the two sets of diagrams are in agreement so that it can be shown from the analysis that the behaviour is caused by the dominant mutual radiation impedance terms. Near resonance, where the mechanical impedance of the projectors is low, these terms control the phase and amplitude of motion through the water coupling.

1. Introduction

In 1958 an experimental investigation of the total radiation impedance seen by rigid piston sound sources in arrays was started at the Admiralty Research Laboratory. The projectors used in this investigation were acoustically small, the bulk of the measurements being made with projectors having a square head $\lambda/6$ wide, where λ is the wavelength of sound in water at their resonant frequency of 18 kc/s. In order to measure the radiation loading a simple basic projector design was used with only two components, an ADP crystal plus a flint-glass head, so that the equivalent circuit of the projector was accurately known. From this knowledge of the equivalent circuit an electrical simulator was constructed which was used to calibrate the projectors by measuring the change in electrical admittance of the simulator near resonance when equivalent resistive and reactive loads were placed in the head circuit. It is not intended in this paper to describe the results of the overall investigation but only to give a brief description of the most important part.

In November 1959, when measurements of the radiation impedance of these projectors in arrays was being made, it was found that the electrical admittance diagrams obtained from individual elements were not circular but varied violently with frequency near resonance. At certain frequencies negative conductance values were measured indicating energy absorption from the loading medium. The diagrams were only circular, corresponding to a constant radiation load, when there was a unique value of radiation impedance independent of the position of a projector in the array, i.e. a square array containing four projectors. These effects are illustrated in Fig. 1 which shows the anomalous diagram from a projector in a square 36-element array which can be compared with the circle diagram for a similar projector in a square 4-element array.

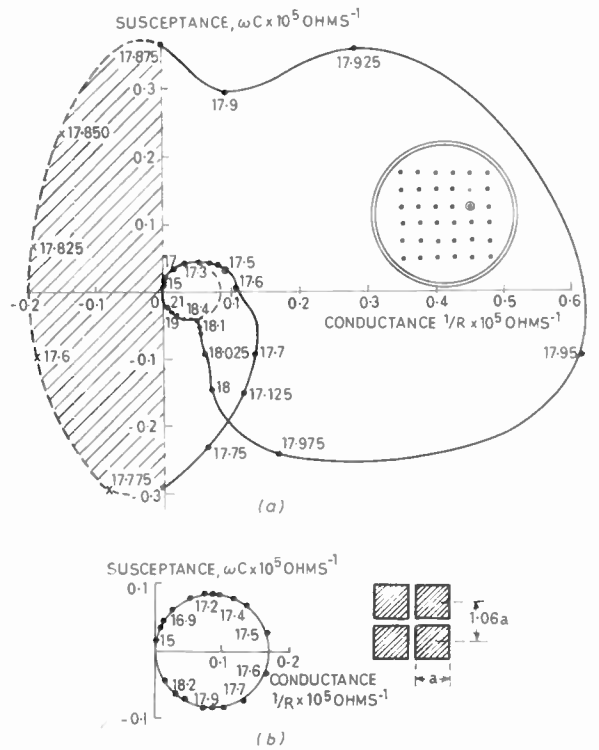


Fig. 1 (*a*). The electrical admittance diagram of one projector in a square 36-element array when the distance between the projector centres is $19a/16$. The position of the projector in the array is shown in the inset.

(*b*) The electrical admittance circle of one projector in a square 4-element array when the distance between the projector centres is $1{\cdot}06a$.

† Admiralty Research Laboratory, Teddington, Middlesex.

At the time the author put forward the hypothesis that the behaviour was due to large variations in the phase and amplitude of the radiating face velocities of neighbouring projectors near resonance.† In acoustically small projectors of low mechanical resistance the mutual impedance values would dominate and could control the phase of the projectors near resonance where their own mechanical impedance was approaching zero. If this hypothesis is correct then it should be possible to compute electrical admittance diagrams with the same violent behaviour near resonance as those measured using accurate radiation impedance data and the known equivalent circuit of the projectors. If the theoretical and measured diagrams are similar then it is an easy matter to investigate the behaviour further since all the required parameters will already have been computed to produce the electrical admittance diagram.

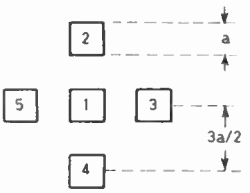


Fig. 2. Five-projector cruciform array.

2. Investigation of the Anomaly

In order to test the above hypothesis it was decided to use a single 5-projector cruciform array. (Fig. 2.) This would provide a high mutual radiation impedance on the centre projector so that the anomalous behaviour should be clearly seen in the measured electrical admittance diagram. Furthermore, due to the symmetry involved, only two discrete radiating face velocities were involved in the computation for the theoretical electrical admittance diagram.

The theoretical electrical admittance diagram for the ith projector in the array was computed from the known equivalent circuit terminated by a total radiation impedance Z_i given by the first order radiation impedance equation

$$Z_i = Z_{ii} + Z_{ij} \cdot \frac{v_j}{v_i} + \ldots\ldots + Z_{in} \cdot \frac{v_n}{v_i}$$

where Z_{ii} is the self radiation impedance of the ith projector, Z_{ij} is the mutual radiation impedance coefficient of the ith projector due to the presence of the jth projector, and v_i and v_j are the complex radiating face velocities of the ith and jth projector respectively. For the case of the cruciform array this gives two equations:

for the centre projector

$$Z_1 = Z_{11} + 4Z_{12} \cdot \frac{v_2}{v_1}$$

and for the outside projector

$$Z_2 = Z_{12} \cdot \frac{v_1}{v_2} + (Z_{11} + 2Z_{23} + Z_{24})$$

It was assumed that the static presence of the remaining projectors would not greatly affect the value of the mutual radiation impedance coefficients between two of them so that values for the self radiation impedance Z_{11} and the mutual radiation impedance coefficients Z_{12}, Z_{23} and Z_{24} were simply taken from previous measurements made on two projectors as a function of separation distance. Since the above two equations are dependent on the velocities v_1 and v_2 the derived equations using the equivalent circuit to connect the driving voltage to the radiation impedance will also be dependent on these velocities. These two complex linear equations were solved for v_1 and v_2 by a *Pegasus* computer for 80 discrete frequencies between 14 and 21 kc/s. Having solved the velocities it was then possible to compute the electrical admittance seen at the driving terminals of the centre and outside projectors.

The electrical admittance diagrams computed in this way are given in Figs. 3 and 4. They clearly show violent changes in admittance near resonance and in the case of the central projector large negative conductance values are evident. As a contrast the circle loci corresponding to a constant radiation load when v_1 equals v_2 are included, only at off-resonant frequencies does the computed admittance lie on these circles. The equivalent measured electrical admittance diagrams made on a cruciform array of the ADP–glass projectors driven at a constant voltage are

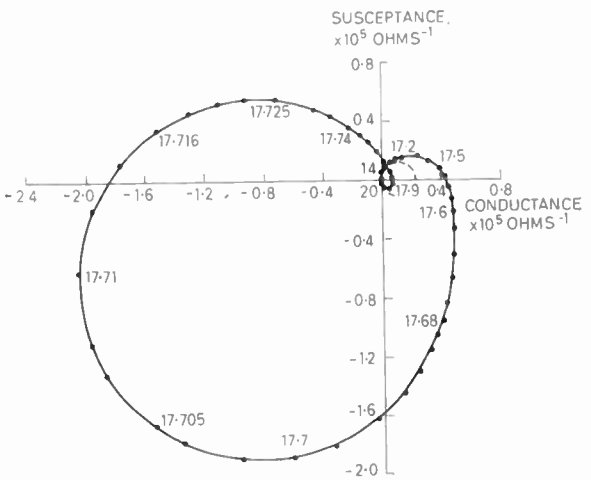


Fig. 3. Computed electrical admittance diagram for the central projector in the 5-element cruciform array.

† J. S. M. Rusby, "Measurements of the total acoustic radiation impedance of rigid pistons in an array", *Nature*, **186**, pp. 144–5, 9th April 1960.

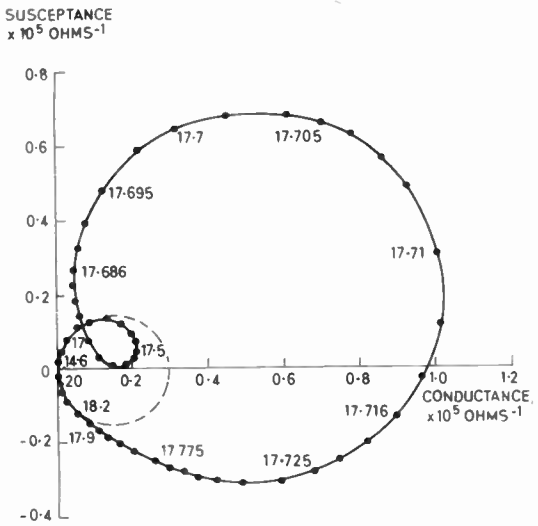


Fig. 4. Computed electrical admittance diagram for an outside projector in the 5-element cruciform array.

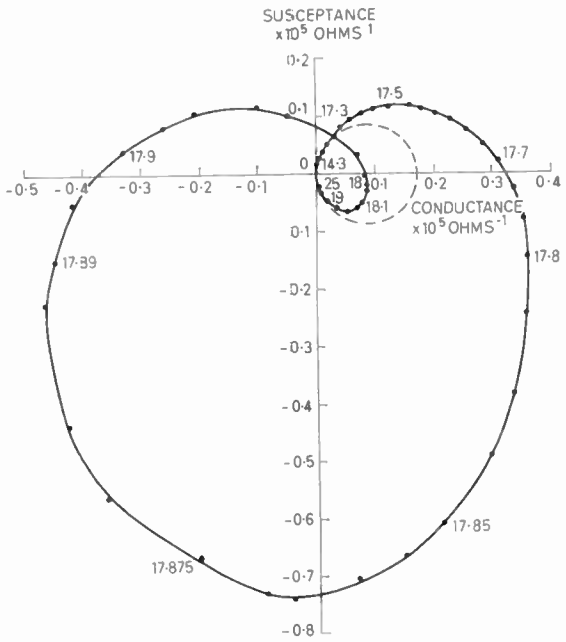


Fig. 5. Measured admittance diagram for the central projector in the 5-element cruciform array.

given in Figs. 5 and 6. They are similar in shape to the computed diagrams although there is a difference in the magnitude of admittance. This error is due to two factors: mainly to inaccuracies in the value of the transformation coefficient used, ϕ, where the electrical admittance is proportional to ϕ^2 times the mechanical admittance, but there is also a small error introduced because the computer is dealing with a lossless equivalent circuit rather than the finite Q projector ($Q \simeq 500$).

Since there is good agreement in form between the computed and measured admittance diagrams it is an easy matter to investigate the anomaly further since all the necessary parameters such as radiating face velocity and total radiation impedance have already been computed for the comparison of the admittance diagrams.

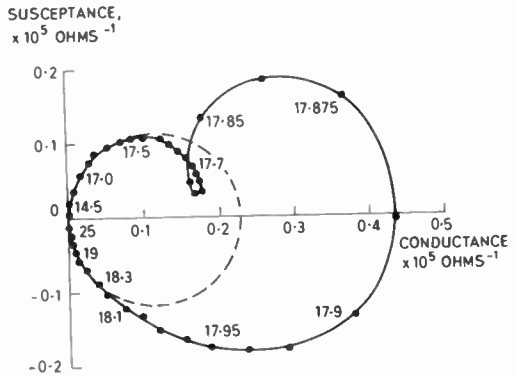


Fig. 6. Measured admittance diagram for an outside projector in the 5-element cruciform array.

The parameters are given in Fig. 7 for both the central and an outside projector. In this figure it can be seen that at off-resonant frequencies, where the phase and amplitude of the radiating face velocities are controlled by the mechanical impedance of the projectors, the two velocities are equal. But near resonance, where the mechanical impedance is small, the phase and amplitude of the velocities will be largely determined by the total radiation impedances which are not the same for the central and outside projectors. When the projectors have only a small

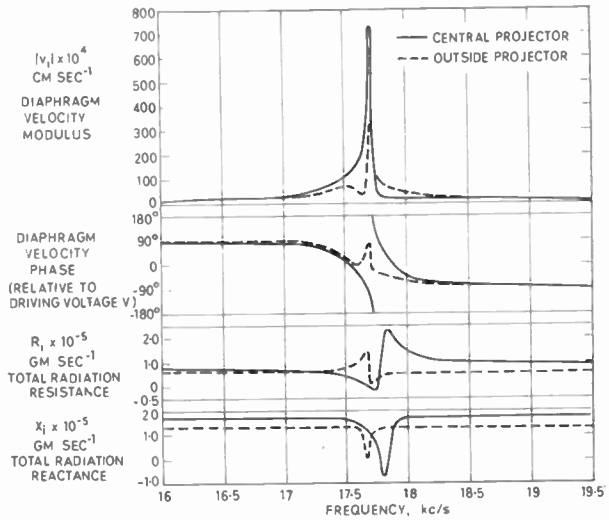


Fig. 7. Computer results for the investigation of the 5-projector array anomaly.

mechanical loss, insufficient to control phase, then the high mutual impedance values can produce phase differences in the steady state in excess of 90 deg. When this happens it is possible for the total radiation resistance of a projector to become negative and so absorb energy from the water rather than radiate it. This has occurred at 17·7 kc/s in Fig. 7, where the large phase difference between v_1 and v_2 has caused the real part of the mutual radiation impedance on the central projector to become negative, that is $\mathscr{R}4Z_{12}.\frac{v_2}{v_1}$. Because there is only a slight difference between this negative term and the positive self-radiation resistance the loading on the projector is very small and excessive velocity amplitudes are recorded on the figure near 17·7 kc/s.

3. Conclusions

Certain qualitative conclusions were drawn from the investigation. It is clear that the severity of the anomalous behaviour will depend on a number of factors, namely:

1. The mechanical resistance of the projectors. If this is low then there is little damping left in the circuit at resonance to control phase.
2. The acoustic size of the radiating faces of the projectors. If this is small then the mutual radiation impedance terms will predominate over the self-impedance and so largely control the radiating face velocity near resonance.
3. The spacing between the projectors in an array. The smaller this is the higher the mutual radiation impedance will be.

It is obvious that this anomalous behaviour has important practical repercussions. The output power, efficiency and bandwidth of an array will all be detrimentally affected since it reduces the total radiation resistance on projectors in the array near resonance to extremely low values. Furthermore the absorption of energy from the loading medium rather than radiation into it also presents problems which have to be solved.

Consideration may be given to the remedies which can be employed to overcome these unwanted effects:

1. The projectors in an array could be decoupled by moving them further apart.
2. Only symmetrical array configurations could be used, i.e. where there is a unique value of projector total radiation impedance independent of position.
3. The comparatively narrow band near the resonant frequency could be avoided where the anomaly is found, and
4. Additional mechanical resistance could be added to the projector to control the velocity phase near resonance.

Most of these suggestions are straightforward and require little amplification. If (1) is adopted the off-resonance radiation resistance will fall so that radiated power, efficiency and bandwidth will be reduced. It is possible to drive a limited number of projectors in a symmetrical array, as suggested in (2), but it would be difficult to accommodate large numbers. Avoiding the area near the resonant frequency where the anomaly occurs, as in (3), is a practical possibility, if the reduction in useful bandwidth can be tolerated. The fourth suggestion would lead to a loss of efficiency, which in certain designs may be acceptable.

Manuscript received by the Institution on 4th June 1962 (Paper No. 802/SS19).

U.D.C. 534.88 : 639.2.081.7

Measurements of the Target Strength of Fish

By

D. H. CUSHING, †

F. R. HARDEN JONES, †

R. B. MITSON, †

G. H. ELLIS, ‡

AND

G. PEARCE ‡

Presented at the Symposium on "Sonar Systems" in Birmingham on 9th–11th July 1962.

Summary: Measurements have been made at 30 kc/s of the target strength of dead cod, herring, plaice and perch. Four independent series of measurements have been made between 1955 and 1959. In three of the series, false Onazote swimbladders were placed in the body of the fish to simulate the acoustic properties of the normal swimbladder. The results are tabulated and presented in a graph, the reference level being the target strength of a sphere of 2 m radius. The results are discussed in relation to the frequencies which are most likely to be suitable for echo-sounding on fish targets when quantitative data are required.

1. Introduction

Standard echo-sounders have been used to estimate the catch of fish visually from a cathode-ray tube[1] and automatically from a counting system,[2] the estimates being confirmed by the quantity of fish caught in the trawl. The principle used was to count or record a measure of the total echo signal in microvolts during the trawling period and multiply this by the mean depth squared. A relationship was shown between the total echo signal and the fish caught.

The target strength of single fish is an important factor in attempts to measure their abundance by echo-sounding methods. Under normal echo sounding conditions the dorsal surface of the fish is presented to the acoustic signal and it is the target strength of this aspect of the fish which forms the subject of the four series of experiments described in this paper.

2. Methods

Most of the measurements were made with cod (*Gadus morhua* L.) and perch (*Perca fluviatilis* L.). Single measurements were made with herring (*Clupea harengus* L.) and plaice (*Pleuronectes platessa* L.).

2.1. *Outline of Procedure and Accuracy of Measurements*

The experiments were carried out with particular regard to the detection of fish by normal echo-sounding equipment and the results given here refer only to the target strength of the dorsal aspect of the fish used.

The fish were either suspended in their normal swimming position below a stationary ship or held vertically with the projector-hydrophone system operating horizontally. Use was made of a reference sphere target, usually an 8 in. diameter (20 cm) air-filled steel trawl float suspended close to or in the same position as the fish target.

The basic elements of the measuring apparatus were units of standard echo-sounding equipment connected to a cathode ray tube display. The received echoes were measured with a continuous wave or pulse signal generator. The measurements were all made at a frequency of 30 kc/s.

In general, the fish targets were between 15 and 100 cm in length, (3–20λ at 30 kc/s). For the larger targets, in order to make measurements in the far sound field, it is necessary to use large distances. For example, in order to obtain an echo signal measurement of 2% accuracy from a 90 cm cod the hydrophone should be a distance of 21·3 m from the target. In order that a reasonable amount of control could be maintained, and considering the facilities available, it was not possible to make measurements over a distance greater than 10·5 m, except at sea. The accuracy of the echo pressure measurement is then reduced to about 6% (0·5 dB). This is considered quite acceptable when the other sources of error in the experiments are considered. These other errors arise from several causes. For example, errors could occur because the fish were never similar in shape, and in some cases—particularly where the fish could not be seen when it was below the water surface—it was difficult to measure the range with great accuracy, or to be certain that the fish was correctly orientated with respect to the acoustic axis of the transducer system. It also seems possible that biological variations in the fish targets could give rise to variations in target strength. It would have been ideal to use freshly killed fish and this was done whenever possible. When an artificial swimbladder was used, this was made of Onazote, a rigid foamed ebonite material, modelled to simulate a real swimbladder in shape and in volume.

Throughout the experiments the usual precautions were taken when using a pulsed echo measuring system, and every care was taken to avoid trapping

† Ministry of Agriculture, Fisheries and Food, Fisheries Laboratory, Lowestoft.

‡ Kelvin Hughes Division of S. Smith & Sons (England) Ltd., New North Road, Barkingside, Essex.

air in the target fish when they were lowered into the water.

2.2. *The Experiments*

Four independent series of measurements were made.

Series 1. Barents Sea measurements on cod (1955).

The measurements were made on freshly killed air-free cod suspended beneath a stationary ship. These experiments have been fully described elsewhere.[1]

Series 2. Indoor tank measurements on perch (1957).

This work has been fully described.[3] The measurements were made under carefully controlled conditions in an indoor freshwater tank 6 m long, 3 m wide and 3 m deep. The perch were freshly killed and had gas filled swimbladders occupying about 7% of the volume of the fish.

Series 3. Burnham-on-Crouch measurements on cod (1958).

In this series the cod had been preserved in formaldehyde for several weeks and were fitted with an artificial swimbladder made of Onazote. The measurements were made from a barge moored in the river. The projector-hydrophone system was fitted to a training shaft in the barge, and was arranged to transmit and receive horizontally propagated signals. The fish target and a reference sphere target were suspended over the side of the barge by nylon strands. These were arranged to keep the fish with its correct aspect facing towards the transducer system. It was only possible to make measurements at slack water, but even then some movement of the target occurred. To eliminate this error, successive echo pulses were recorded on moving film, and a value for the received signal strength was obtained from this record. These measurements were made over a period of about eight weeks. The variation in received echo signals is due to the change in transmitter output and to the changing water conditions throughout this time.

Series 4. Dry dock measurements on cod, herring and plaice (1959).

A dry dock at Lowestoft was used for this work. The dock was 75·5 m long, 15 m wide and filled with seawater to a depth of 4·3 m. With the projector-hydrophone horizontal and the fish target suspended vertically, as in the Burnham-on-Crouch and Indoor Tank experiments, it was possible to use a working range of 10·5 m. The fish had been kept in cold storage and were thawed out before use. The cod and herring were fitted with Onazote swimbladders occupying 5% of their volume. The suspension system for the fish was similar to that used by Jones and Pearce[3] for the Indoor Tank experiments on perch.

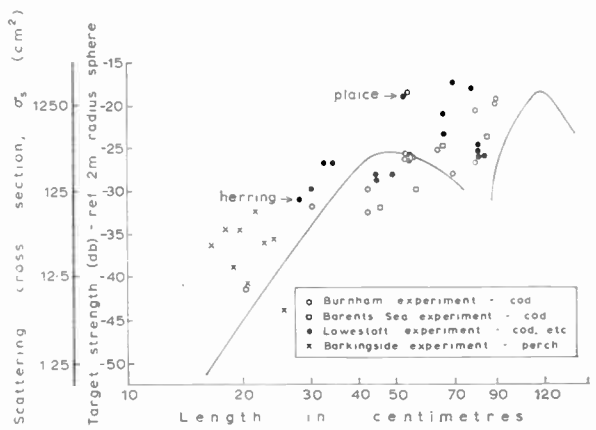


Fig. 1. Scattering cross-section and target strength for fish of different lengths. Experimental observations and Haslett's curve.

3. Results

The results are summarized in Table 1 and shown graphically in Fig. 1 where the target strength and scattering cross-section are plotted against fish length on a logarithmic scale. The Barents Sea cod measurements (Series 1) were made on air-free cod. As the swimbladder contributes about 50% of the echo returned by a fish[3] 6 dB should be added to the target strengths of the Barents Sea results to make them comparable with those obtained in the other series of measurements.

There is considerable variability in target strength of fish of similar size, and this deserves some comment. Jones and Pearce[3] discussed the point in connexion with the indoor tank measurements on perch (Series 2). While these experiments were made under carefully controlled conditions, there were unexpected differences in the target strengths in dorsal aspect which did not appear to be related to the length of the fish. No satisfactory explanation could be found to account for these results. The variability observed in the Dry Dock and Burnham experiments may have been partly due to method of suspension which would allow the fish to swing slightly in the sound beam. Misalignment may also have been a source of variability, as Cushing's plasticine model experiments[4] showed that the polar diagram can be highly directive in dorsal aspect.

4. Discussion

Haslett[5, 6] has adapted the theory of the back-scattering of electro-magnetic waves to similar situations involving acoustic waves. In this way, it is possible to obtain a relationship between fish echo strength and fish length. Fundamentally, this treatment may be said to consider the fish as regular solid bodies, suitable allowance being made for the acoustic properties of this equivalent body. The

Table 1

Measurement of the target strength of fish

Series	Fish	Length in cm	Received signal across 30 ohm tuned hydrophone millivolts	Range in m	Target strength referred to a 2 m radius sphere dB	Scattering cross-section cm^2
1. Barents Sea measurements (1955) on air-free cod made from a ship hove-to.	Cod	85	0·25	50	− 23·9	510
		63	0·212	50	− 25	398
		55	0·125	50	− 30	126·0
		44	0·100	50	− 32	78·9
		20	0·031	50	− 41·7	8·5
Note: The impedance of the tuned hydrophone was unknown in this series.						
2. Indoor Tank measurements (1957) with perch with swimbladders occupying 7% of the volume of the fish.	Perch	25·0	92·8	3·3	− 44·5	4·97
		24·0	37·1	3·3	− 36·1	3·1
		22·5	33·6	3·3	− 36·4	28·7
		21·5	53·0	3·3	− 32·5	71·0
		20·5	19·8	3·3	− 41·1	9·83
		19·5	40·3	3·3	− 34·7	42·5
		19·0	24·8	3·3	− 39·2·	15·0
		18·0	40·3	3·3	− 34·8	42·2
		16·5	33·9	3·3	− 36·3	29·2
3. Burnham-on-Crouch measurements (1958) made from a moored barge on cod fitted with artificial swimbladders ocupying 5% of the volume of the fish.	Cod	90	13·4	7·7	− 20·1	1229
			20·9	7·7	− 19·8	1332
		80	13·4	7·3	− 20·7	1068
			8·6	10·7	− 27·1	246
		70	7·5	7·3	− 28·2	191
		64	13·4	7·4	− 25·4	364
		53	6·7	7·9	− 26·2	330
			8·4	8·6	− 26·5	282
			35·9	6·0	− 18·5	2180
		42	7·5	7·6	− 30·0	110
			5·0	7·5	− 32·7	103·8
		30	5·9	8·6	− 32·0	78·9
4. Dry Dock measurements at Lowestoft (1959). The cod and herring were fitted with artificial swimbladders occupying 5% of the volume of the fish.	Cod	84	7·73	10·5	− 26·5	279
		81	9·48	10·5	− 24·7	423
			8·85	10·5	− 25·3	380
			7·10	10·5	− 18·5	1780
		79	19·4	10·5	− 18·5	1780
		70	22·0	10·5	− 17·8	2100
		66	14·8	10·5	− 20·9	1030
			8·16	10·5	− 26·0	314
		54	8·0	10·5	− 26·2	302
		53	7·7	10·5	− 26·5	279
			8·16	10·5	− 26·0	314
		48	6·28	10·5	− 28·3	188
		44	5·75	10·5	− 29·0	156
			6·4	10·5	− 28·2	192
		34	7·21	10·5	− 27·06	247
		32	7·21	10·5	− 27·06	247
		30	5·1	10·5	− 30·1	124
	Herring	28	4·4	10·5	− 31·4	98
	Plaice	51	19·4	10·5	− 18·5	1780

curve obtained from these measurements (Fig. 9 in ref. 6) is shown superimposed on the graph.

Apart from the experiments which are here described, all measurements on large fish which have been made correspond to the straight portion of Haslett's curve. The results, in this case, cover a much wider range and it is of interest to compare these with the values obtained from the curve. Visual examination shows that there appears to be some correlation between the two, but so far no attempt has been made to determine this correlation statistically. The most careful measurements occur on the straight part of the curve. They do not, however, straddle the curve. The results which occur at the first dip, that is, at a fish length of about 80 cm, are of greatest interest but are, unfortunately, the most inaccurate. Of the sixteen experimental values of target strength occurring in this region, none occur at the minimum of Haslett's curve. However, it is quite obvious that many more measurements must be made at the critical points.

When a target is small compared with the wavelength of incident sound, the back scattering will follow Rayleigh's Law ($\sigma_s \propto 1/\lambda^4$). Figure 1 shows that the scattering cross-section is more or less proportional to $1/\lambda^4$ up to fish lengths of 40 cm ($L = 8\lambda$ at 30 kc/s). With bigger fish the scattering cross-sections vary considerably and recent experimental work by Haslett leads one to expect a series of maxima and minima. Nevertheless, a line joining the two peaks of Haslett's curve shown on Fig. 1 is a reasonable fit to the experimental data when it is remembered that the Barents Sea experiments, which lie well below the line, were made on air-free cod. For fish targets within the length range of 40–120 cm (8–24λ at 30 kc/s) the data would appear to fit the relationship $\sigma_s \propto 1/\lambda$. With bigger fish, when the target is large compared with the incident wavelength, the relation between echo signal and target size will approach that expected from a plane surface: reflection will be geometric, $\sigma_s \propto 1/\lambda^2$.

Let us suppose that geometric reflection occurs with fish lengths greater than 500 cm (100λ). The relation between scattering cross-section and length of fish can then be summarized as shown in Fig. 2. Up to 40 cm (8λ), $\sigma_s \propto 1/\lambda^4$; from 40–500 cm, $\sigma_s \propto 1/\lambda$; above 500 cm (100$\lambda$), $\sigma_s \propto 1/\lambda^2$. In the middle zone (fish length 40–500 cm, 8–100λ) the relation between scattering cross-section and length will be complex and a series of maxima and minima are to be expected, differing by a factor of 10.

In Fig. 3 the relationships shown in Fig. 2 have been set out for a band of frequencies between 1 kc/s and 10 Mc/s. This figure is of interest in connexion with the results of some earlier experiments carried out by Cushing and Richardson.[7]

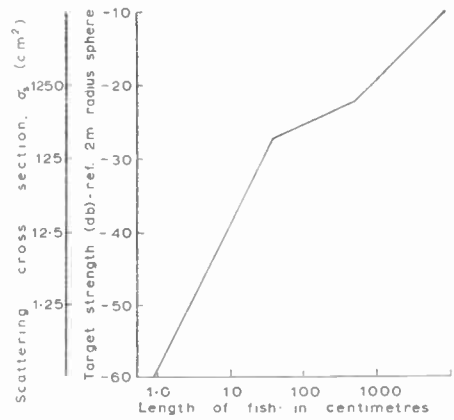


Fig. 2. The relation between scattering cross-section and the length of the fish at 30 kc/s.

Table 2 gives the ratios of the amplitudes of echo signals returned by cod and herring at three frequencies, the original having been made at sea.[7]

Table 2

Ratios of the amplitudes of the echo signals returned by cod and herring (Cushing & Richardson[7])

	Frequency		
	10 kc/s	15 kc/s	30 kc/s
Cod	1	5·4	0·6
Herring 1	1	2·9	9·2
Herring 2	1	2·7	7·0

From Fig. 3 it will be seen that herring of 20–30 cm lie in the Rayleigh scattering zone with frequencies of 10–30 kc/s. The scattering cross-section ratios for herring from Fig. 3 are:

10 kc/s	15 kc/s	30 kc/s
1	2·5	14·0

For cod (of 60–70 cm, the lengths on which the observations in Table 2 are based), the ratios from Fig. 3 are:

10 kc/s	15 kc/s	30 kc/s
1	2·5	6·0

The cod do not lie fully in the zone of Rayleigh scattering. The 30 kc/s observation is in the middle zone (where $\sigma_s \propto 1/\lambda$) and Haslett shows that the variation between maxima and minima is of the order of ten times. So the anomalous observation on cod in Table 2 at 30 kc/s may have fallen on a minimum.

In the original work on frequency dependence, Cushing and Richardson[7] suggested that the best results were obtained when $H/\lambda \simeq 1$, where H is the dorsiventral dimension of the fish, i.e. the height from

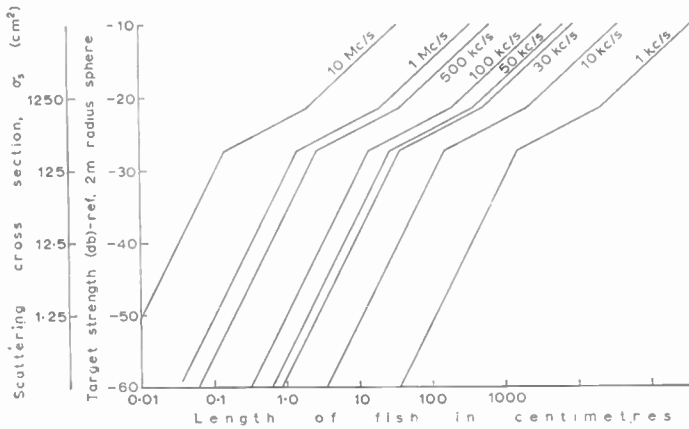


Fig. 3. The relation between scattering cross-section and the length of the fish at different frequencies.

the base of the dorsal fin to belly. For cod and herring when $H/\lambda \simeq 1$, $L/\lambda \simeq 5$. We have taken the value of $L/\lambda = 8$ as the lower limit of the middle zone. The rationale of the original empirical rule now becomes a little clearer. At $H/\lambda \simeq 1$, or $L/\lambda \simeq 5$, the highest point in the Rayleigh scattering zone was being used without crossing into the middle zone of oscillating maxima and minima.

The transition from Rayleigh scattering to geometrical reflexion probably depends on the dimension of some part of the fish approaching a critical value in relation to the wavelength. The changeover appears to take place when $H/\lambda = 1$ and $L/\lambda = 8$. But it is not known what part or dimension of the fish is important acoustically during the transition. Haslett's work[9] on the standard fish structure provides the basis from which further studies can be made.

If echo sounders are to be used quantitatively for fish counting, the middle zone should be avoided. It might be possible to describe the maxima and minima theoretically for a standard fish shape, but the variation in signal due to size differences and aspect differences would be high. Cushing and Richardson[8] interpreted the variability of the echo signal as a biological character, such as variation in shoaling pattern. It was highest for cod at 30 kc/s, in the middle zone: further, in this zone the polar diagrams of scattering are complex and very variable.[7] Thus it is possible that the anomalous variability was physical in origin.

The avoidance of the middle zone means that the frequency bands can be specified. For example, for herring of 20–30 cm, the best frequencies below the middle zone are below 50 kc/s; above the middle zone, anything above a megacycle can be used. For cod of 70–120 cm in length, the best frequencies below the middle zone are below 10 kc/s and the best frequencies above the middle zone are above 100 kc/s.

Obviously other considerations play a part in the choice of frequency. But for quantitative work we may make three points:

1. We can make sure that the targets being measured are in the Rayleigh zone or in the geometric zone;
2. We should work as near the upper end of the Rayleigh zone as possible, to obtain maximum signal;
3. In the geometric zone, σ_s increases with frequency regularly; this is also true in the other zones, but we should aim at the maximum in the Rayleigh zone and avoid the middle zone.

5. References

1. I. D. Richardson, D. H. Cushing, F. R. Harden Jones, R. J. H. Beverton and R. W. Blacker, "Echo sounding experiments in the Barents Sea", *Fish Invest. Lond.*, Ser. 2, **22**, No. 9, 1959.
2. R. B. Mitson and R. J. Wood, "An automatic method of counting fish echoes", *J. Cons. Int. Explor. Mer*, **26**, pp. 282–91, 1961.
3. F. R. Harden Jones and G. Pearce, "Acoustic reflexion experiments with Perch (*Perca fluviatilis* Linn.) to determine the proportion of the echo returned from the swim bladder", *J. Exp. Biol.*, **35**, pp. 437–50, 1958.
4. D. H. Cushing, "Some echo sounding experiments on fish", *J. Cons. Int. Explor. Mer*, **20**, No. 3, pp. 266–76, 1955.
5. R. W. G. Haslett, Ph.D. Thesis, University of London, 1960.
6. R. W. G. Haslett, "Determination of acoustic backscattering patterns and cross sections of fish", *Brit. J. Appl. Phys.* **12**, 1962 (in the press), Fig. 9.
7. D. H. Cushing and I. D. Richardson, "Echo sounding experiments on fish", *Fish. Invest. Lond.*, Ser. 2, **18**, No. 4, 1955.
8. D. H. Cushing and I. D. Richardson, "A triple frequency sounder", *Fish. Invest. Lond.*, Ser. 2, **20**, No. 1, 1955.
9. R. W. G. Haslett, "Measurement of the dimensions of fish to facilitate calculations of echo-strength in acoustic fish detection", *J. Cons. Int. Explor. Mer*, **27**, pp. 261–9, 1962.

Manuscript received by the Institution on 27th June 1962 (Paper No. 803/SS12).

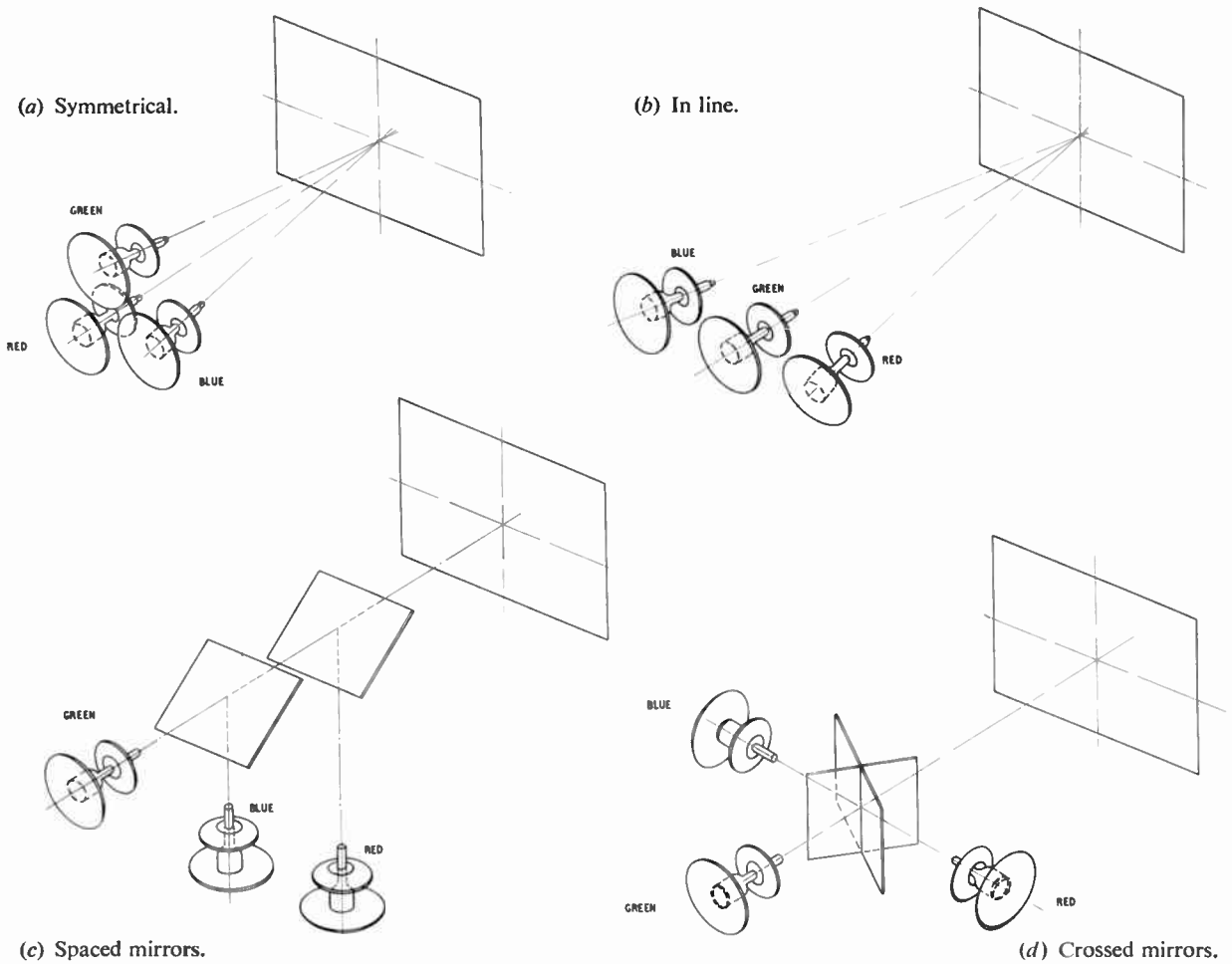


Fig. 1. Projector layouts.

Keystone error may be avoided by the arrangement shown in Fig. 1(*c*) where dichroic mirrors are used to give a common projection plane. The mirror adjacent to the red tube reflects red light but transmits blue and green, while the second mirror reflects blue light and transmits green. The necessity for equal path lengths gives rise to a rather cumbersome assembly. By using crossed dichroic mirrors as in Fig. 1(*d*) the required volume may be reduced.

At each colour selective layer losses of between 10 and 20% occur, dependent upon the angle of incidence. Since light passing through the dichroic mirrors is divergent, then colour shading will be present in one direction and it is desirable to arrange that the intersection of the mirrors is horizontal in order to confine picture shading to the vertical axis which subtends a smaller projection angle than that of the line axis.

After considering the relative merits of the various systems, especially in regard to minimum projector height and freedom from unwanted reflections—often artificially increased by deposits of dust and smoke tar—the scheme shown in Fig. 1(*b*) was adopted.

3. Requirements

Many conflicting factors are encountered when specifying from first principles the optimum parameters for the design of a projector, and experiences of previous projection systems is of great value in enabling one to estimate, by scaling, the effects of changing various components.

The screen brightness, to which there is a minimum acceptable value, will depend upon the brightness of the projection cathode-ray tube, the optical characteristics of the Schmidt system and the screen efficiency.

3.1. *Cathode-ray Tubes*

The light output which may be obtained from a cathode-ray tube is a function of accelerating anode potential, luminous efficiency of the phosphor, and the screen current density.

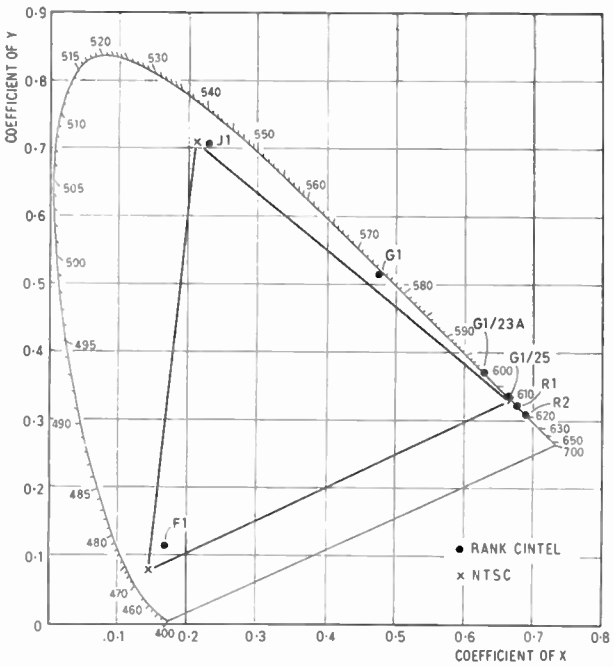


Fig. 2. C.I.E. colour diagram.

Practical limitations exist; for tubes which may readily be accommodated in optical assemblies, the range of envelope sizes is such that difficulty may be experienced in handling the high voltage anode supply and preventing flash-over; further, at the required values of beam current, the power dissipated in the face-plate may be excessive.

The range of phosphors available for use above 20 kV is limited and Table 1 lists a representative selection, together with their chromaticity co-ordinates which are shown on the C.I.E. colour diagram (Fig. 2). It will be seen that good approximations to the N.T.S.C. primaries are available, with the exception of the zinc beryllium silicate G_1, which has an orange appearance but may be corrected by the use of colour correcting filters.

From the chromaticity co-ordinates the relative brightness of the red, green and blue contributions necessary to match Illuminant C may be calculated and these have been tabulated as percentages in columns 7–10 of Table 1.

The efficiency of the green phosphor is high compared with that of the red and blue phosphors. From a knowledge of the screen efficiencies in lumens per watt, the ratio between the powers which have to be dissipated via the tube face-plates may be found. These ratios are shown in columns 11–14 for unit power dissipation in the green tube, from which it can be seen that the red tube must dissipate the greatest

Table 1. Phosphor Data

Colour and Code	Composition	Emission Peak Å	C.I.E. Co-ordinates		Luminous efficiency lumens/watts	Illuminant C match							
						% Luminance				Dissipation ratio			
			x	y									
1	2	3	4	5	6	7	8	9	10	11	12	13	14
Red R_1	Zinc phosphate (Mn)	6300	0·676	0·323	5·5	26				2·7			
Red R_2	Zinc cadmium magnesium silicate (Mn)	6500	·690	·310	2		24				6·7		
Red G_1	Zinc beryllium silicate (Mn)	5500	·476	·514	12								
G_1/23A	Zinc beryllium silicate (Mn) with Wratten 23 A	6000	·625	·375	5			34				3·5	
G_1/25	Zinc beryllium silicate (Mn) with Wratten 25	6150	·67	·330	2·5				27				6·3
Green J_1	Zinc orthosilicate (Mn)	5300	·228	·702	32	56	58	48	55	1	1	1	1
Blue F_1	Calcium magnesium silicate (Ti)	4150	·162	·113	4	18	18	18	18	2·5	2·4	2·9	2·5

The elements shown in parentheses refer to the phosphor activator.

power when projecting scenes of normal colour balance.

Of the red phosphors, R_1 requires least dissipation for white balance but discolours rapidly in use, while R_2, with a peak emission at 6500A enables highly saturated reds to be achieved but at the expense of face-plate dissipation. For less saturated reds, the G_1/23A combination is quite suitable but when corrected to the true N.T.S.C. red, with Wratten filter No. 25, it offers little advantage over R_2.

Unfortunately, owing to its toxic properties, the use of beryllium compounds introduces handling risks during manufacture and following breakages. This restricts its freedom of use and, for general applications, the R_2 phosphor is specified.

Heat may be removed from the face-plate by forced air cooling but the power that may be safely dissipated is limited by the permissible temperature gradient within the glass face, since this gives rise to differential stress within the bulb. This condition may be further aggravated when projecting pictures having a high peak-to-mean brightness.

The luminous efficiency of phosphors falls with rising temperature but the primary limit to dissipation within tubes having a high ratio of usable to overall screen diameter is set by glass stresses which, if exceeded, may give rise to failure of the envelope.

To preserve a good colour balance over the desired contrast range a linear relationship should exist between the light output and beam current of the three tubes. It is well-known that at high current densities this linearity is not maintained but at the peak currents required for maximum output little saturation exists. If, however, the equipment is seriously overdriven the picture assumes a reddish brown cast since both blue and green phosphors become more saturated than the red.

All these projection phosphors have a medium persistence with the exception of the blue which is short and, in consequence, can give rise to yellow trails at the rear of fast-moving images.

The cathode ray tubes used in this projector have a nominal 4-in diameter face, an accelerating potential of 30 kV, and are electro-magnetically deflected and focused. The limiting beam current for the red tube is about 1 mA, giving a screen dissipation of 30 watts.

Under these operating conditions X-rays are produced and ordinary glass discolours owing to electron bombardment. By fabricating the tube envelopes from a non-solarizing lead glass, discoloration is prevented and X-rays are adequately attenuated.

3.2. *Schmidt Optics*

The basic optical theory regarding the design of Schmidt systems has been described in the literature and will not be repeated here.[4] For the range of applications envisaged it was essential to keep the size and weight of the projector as low as possible and, at the same time, provide an acceptable level of screen brightness.

The optical system uses a 10½-in diameter mirror and an 8-in diameter glass corrector plate. A range of three plates are available and have been figured for nominal picture widths of 4 ft, 6 ft or 9 ft, the corresponding throw distances being 7 ft 3 in, 11 ft 6 in and 17 ft 6 in, respectively.

The picture size generally used is 6 ft × 4 ft 6 in and with a screen gain of 3 the maximum screen brightness is 10 ft-lamberts.

It is of interest to consider how elements may be scaled in an endeavour to increase the screen brightness for a given picture size.

The maximum apertures obtainable from modern Schmidt systems have *f* numbers between 0·7 and 1·0. Given a good original design, the brightness may be increased only by reducing the magnification of the system, i.e. by increasing the c.r.t. raster size. To maintain the same face-plate brightness the current density must remain the same, necessitating an increase in beam current and face-plate dissipation. Optical aberrations limit the angle of projection to approximately 40 deg, hence for a given picture size the throw distance cannot be materially reduced. The reduction in magnification necessitates an increase in focal length of the system, requiring the size of both mirror and corrector plate to increase. For a given deflection angle within the cathode ray tube, the larger raster necessitates a longer tube, but the increase in overall dimensions would permit a higher e.h.t. potential to be used. This would allow either a small reduction in screen loading for the same brightness, but it is more likely that the current density would be maintained. From the foregoing it can be appreciated that an increase in screen brightness can only be obtained at the expense of an increase in size, weight and complexity, or by degradation in the form of greater off-axis aberrations and vignetting.

3.3. *Keystone Correction*

The three Schmidt systems are so aligned that their optical axes intersect at the centre of the viewing screen. Since the outer blue and red assemblies are not normal to the screen, their magnifications are unequal in the left and right halves of the picture and the projected images, if not corrected, would exhibit keystone distortion.

These errors may be corrected electrically by modification of the deflecting fields. By increasing the velocity of the red line scan in the right half of the picture, the right-hand vertical edge of the red and

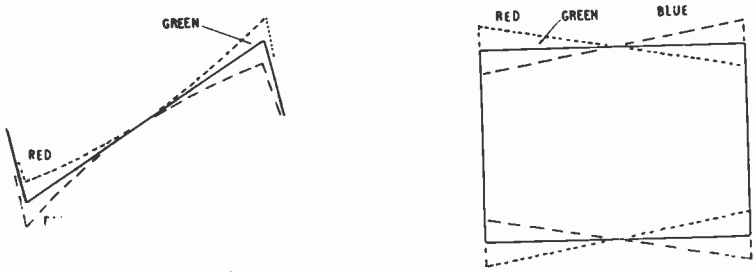


Fig. 3. Partial keystone correction.

(*a*) Line deflector currents.

(*b*) Superimposed images.

green images may be made coincident and a corresponding decrease in velocity in the left half of the red picture will align the left-hand verticals. Assuming a linear sawtooth deflection current in the green deflector coil the velocity correction for the red scan may be achieved by the addition of a parabolic component to the sawtooth current in the red deflector.

The correction necessary to obtain coincidence of the blue edges is similar but of opposite sign to that required for the red tube, that is, the velocity with respect to the centre has to be increased on the left and reduced on the right-hand side. Suitable waveforms for correction are shown, Fig. 3(*a*), and without further correction the imaged rasters would be as in Fig. 3(*b*). A pattern comprising vertical bars would appear to be in register but it can be seen that a progressive vertical error exists on either side of the horizontal centre line.

For full correction of the red raster the field amplitude has to be reduced in the left half and increased in the right half of the picture. This may be achieved by passing a line sawtooth current through an auxiliary field deflector, the sawtooth having maximum amplitude at the top of the picture, reducing to zero at the centre and increasing in the opposite polarity to a maximum amplitude at the bottom of the picture. This waveform is shown in Fig. 16(*e*).

3.4. *Rhombic Correction*

Close examination of the raster on a cathode-ray tube having electro-magnetic deflection and focusing may reveal a departure from exact quadrature between the line and field directions. This may be due to lack of symmetry of the deflector windings or the effects of induced current flowing in the focus or gun assembly. The angular error is usually less than 3 deg but when superimposing images it is necessary to provide adjustment to match the errors.

Correction of this rhombic, or skew, error may be made by introducing a fraction of the field deflection current into the line direction. For this purpose an auxiliary coil is wound on the same core as the keystone deflector and fed with a field sawtooth current.

By suitable adjustment of the amplitude and polarity of the correcting currents, the imaged rasters from the red and blue c.r.t.s may be correctly superimposed upon that of the green image.

3.5. *Pin-cushion Correction*

The radius of curvature of the cathode-ray tube face-plate is fixed by the optical design of the Schmidt system and is greater than the radius of deflection. In consequence, pin-cushion distortion is present. This may be corrected, by use of a multi-pole magnetic field adjacent to the face-plate to produce a rectangular raster, and the introduction of S distortion in the deflection currents to restore linearity. The extent of the pin-cushion error in this equipment is small and only velocity correction, by suitable modification to the scan generator drive waveforms, has been included.

3.6. *Amplitude and Linearity Matching*

Apart from keystone and rhombic requirements it might be thought that the use of three similar deflector yokes with the individual line and field windings connected in series, across their respective scan generators, would be sufficient to ensure registration. This is not the case, however; to compensate for mechanical and winding tolerances individual adjustments of about $\pm 2\frac{1}{2}\%$ must be provided for both amplitude and linearity of each deflecting field. The use of parallel rather than series connections enables the windings to be approximately balanced with respect to earth, thus minimizing undesirable coupling and also reducing the potential difference, and consequent insulation problems, between windings.

Residual mutual coupling between windings of a deflector can give rise to the top part of the picture being horizontally displaced. A coupling transformer is connected between the line and frame circuits and is so constructed that adjustment of a dust core enables the overall mutual coupling to be reduced to zero.

At line frequencies the deflector is mainly inductive and hence variable inductors and resistors are used for

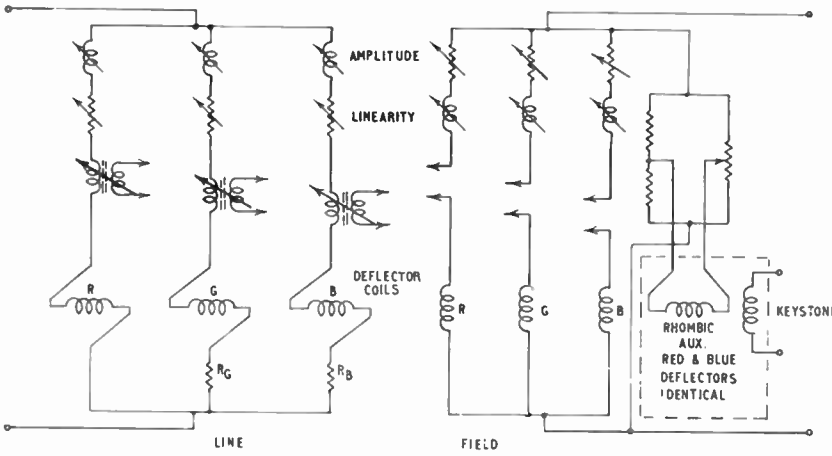


Fig. 4. Deflector arrangements.

amplitude and linearity trimming respectively, while at field frequency the components are used in the reverse order. A circuit diagram of the deflectors and associated components is shown in Fig. 4.

It will be recalled that parabolic deflection components are required in the line direction as part of the correction for keystone distortion. The voltage drive across the deflectors is arranged to have an excess of sawtooth which is as required for the red deflector. For the green and blue deflectors increasing amounts of series padding resistance R_G and R_B are added to obtain the desired correction.

Horizontal and vertical picture shift is obtained by passing direct current through shift coils contained within the focus assembly.

3.7. *Video Drive*

For white balance beam currents have to be in the ratio 6·7 : 1 : 2·4 for the red, green and blue tubes respectively, assuming that no appreciable saturation takes place. The beam current of a cathode-ray tube is related to the grid-cathode voltage by a power law having an exponent, gamma, between 2 and $2\frac{1}{2}$. Hence the drives should be in the ratio $2\frac{1}{2} : 1 : 1\frac{1}{2}$ (approximately).

Colour video signals, as normally distributed, have peak-to-peak values of up to 1 volt. In order to preserve colour balance throughout the cathode-ray tube drive range, the amplifiers must have similar linear characteristics, with individual balance controls, for delivering outputs of the required ratio up to peak values of 150 volts. Common contrast and black level controls are desirable for both setting up and operational use.

Fig. 5 shows in diagrammatic form the video system.

Each of the three incoming video cables is terminated by one section of a three gang, non-inductive, 75 ohm potentiometer which is used as a common contrast control.

A five-position switch is provided enabling, for setting-up purposes, one of the following to be displayed: red, green or blue only, composite colour or green input to all channels.

Adjustment of the black level may be obtained by altering the bias of each tube but it is necessary to

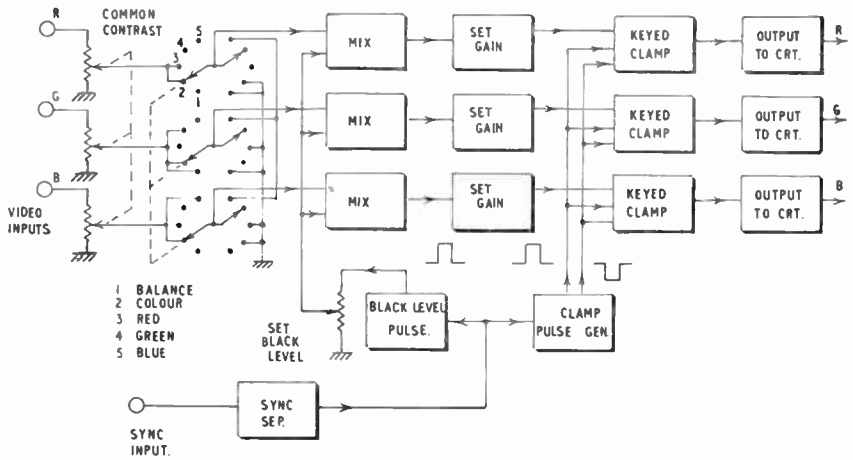


Fig. 5. Video control system.

ensure that such voltage changes are proportional to the relative drive ratios in order to preserve colour balance. This may be achieved by using suitable potentiometer networks, but is complicated due to the cathode-ray tubes having unequal values of cut-off bias.

An alternative arrangement used in this equipment enables a balanced black level change to be made using only one potentiometer. Pulses occurring during the back porch, i.e. the clamping period, are added to each signal in the video amplifiers prior to the individual gain controls. Each amplifier output stage has a driven clamp for establishing the black level at the start of every line and hence the d.c. potential of each video signal is dependent upon the amplitude of the added pulse. Since these pulses pass through the individual balance controls of the appropriate video amplifiers the combined effect of varying the amplitude of the common applied pulse is to change, in a balanced manner, the black level of the picture.

3.8. *Simplified Block Diagram*

A simplified block diagram of the equipment is shown in Fig. 6. The three positive-going red, green and blue video input signals pass via common gain controls and switching to the respective video amplifiers. The video outputs are fed to the cathodes of the respective cathode ray tubes.

Negative-going composite sync pulses are fed into the sync separator from which line sync, field sync, clamp pulses and a black level pulse are obtained. Line sync can be fed to the scan generator directly or, when the input is subject to interference, via a flywheel synchronizing unit. The main output from the line scan generator feeds the deflectors, while auxiliary outputs are distributed to the flywheel, keystone and protection units. The field scan generator accepts suitable sync pulses and provides current to the deflectors as well as auxiliary outputs to the keystone, rhombic and protection units. The protection unit detects the presence of the line and field outputs and, in the event of failure, interrupts both the focus and e.h.t. supplies.

4. Mechanical Description

In order to construct a projector that would be suitable for use in a wide variety of applications, it was necessary to anticipate certain operational requirements.

To exploit the capabilities of this display equipment, it is desirable to provide means for continually monitoring the total cathode-ray tube beam currents, especially when the amplitude range and colour content of the incoming signals is liable to vary. The most suitable position from which to monitor the picture is adjacent to the projector and it was decided to have a self-contained assembly with the three optical barrels on top and the electronic units mounted below, the overall height being such that the line of sight of the screen for persons seated behind the projector was not obscured. The external appearance had to be pleasantly styled and internal access was required for inspection and servicing. A photograph of the complete projector is shown in Fig. 7. Since the equipment was to be used in places of entertainment, it was necessary to provide protection for the public from the high voltage supplies—and protection of the equipment from the public.

The photograph (Fig. 8) shows the projector with covers removed; the main angle iron framework, of welded construction, carries the Schmidt barrels on cross members inclined at the required projection angle. Rubber-tyred castors are provided for normal movement and screw jacks enable the equipment to be

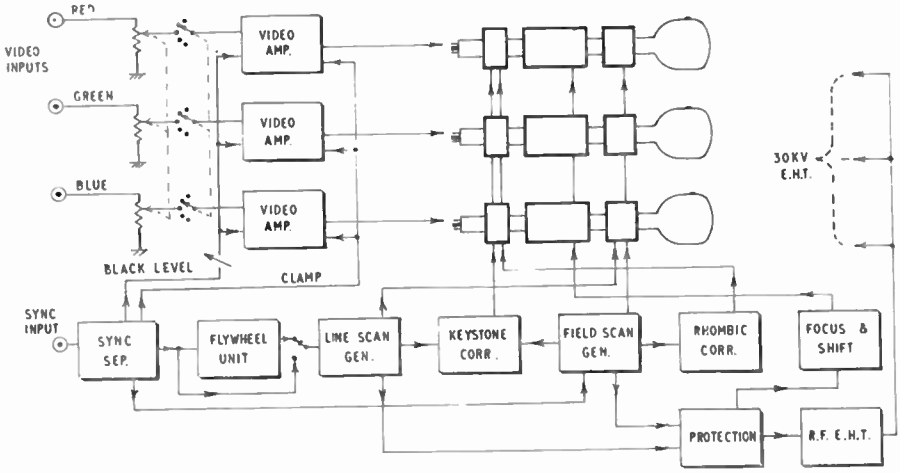


Fig. 6. Block diagram of the projector circuits.

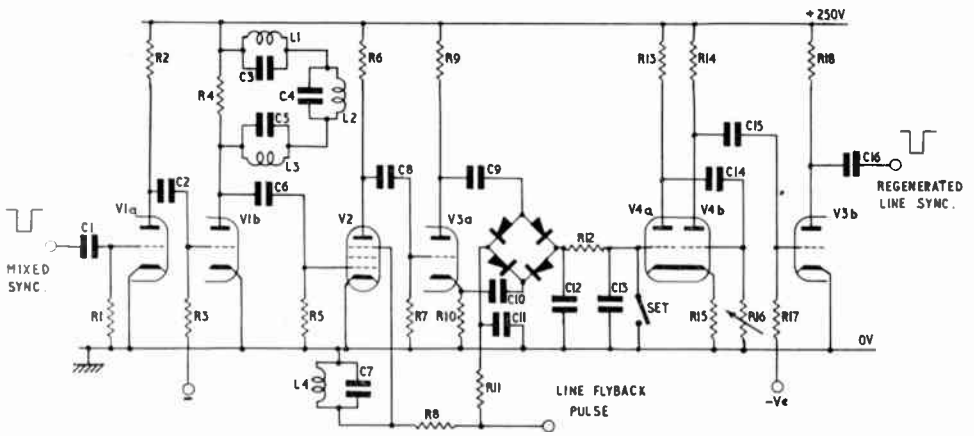


Fig. 11. Line flywheel unit.

The incoming negative pulses, (Fig. 12(*a*)), are amplified and inverted by V1a. The coupling network C2R3 differentiates the pulses, driving the grid of V1b with the waveform (*b*). Since R3 is returned to a negative potential, conduction only occurs during the positive spikes.

The anode load of V1b comprises the resistor R4 in parallel with a short-circuited artificial delay line L1C3, L2C4, L3C5. The anode waveform will consist of negative-going spikes followed by positive-going line reflections delayed by $3\frac{1}{2}$ μs, as shown in (*c*), and only the latter pulses can drive the following stage into conduction.

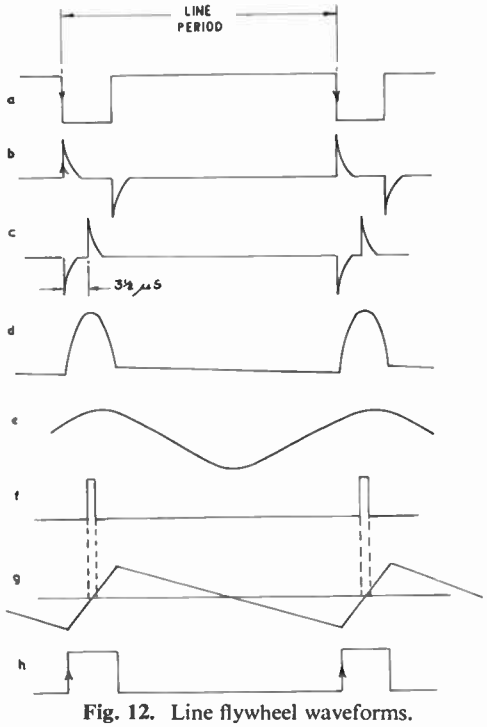


Fig. 12. Line flywheel waveforms.

Positive-going pulses (*d*) from the output of the line scan generator are fed back to this unit via R8 and sustain oscillation of the parallel tuned circuit L4C7, which provides a sinusoidal line gating waveform (*e*) to the suppressor grid of V2.

It will be recalled that the control grid waveform of V2 is derived from a composite sync signal and will therefore contain half line pulses during the field sync period. Since their presence would be detrimental to the correct operation of subsequent stages, the suppressor gating action permits current to flow in V3a only at line rate, when driven by the delayed positive control grid pulses. Hence the anode waveform comprises negative-going line pulses only, which drive V3a to cut off, producing equal and opposite switching pulses, one of which is shown at (*f*), to open the 4-diode gate.

During the conduction period of this bridge the right-hand junction will assume the potential of the left-hand junction. To the left-hand junction is applied an a.c. coupled line sawtooth (*g*) derived by integration in R11 C11. Assume for the moment that the sawtooth passes through zero during the gating period, with the result that C12 is at zero potential. V4 is a multivibrator working nominally at line rate but controllable over a small range by the variable resistor R16 and the grid potential of V4a. This oscillator is normally set to run at line rate by adjustment of R30, with the grid of V4a short-circuited to earth via the SET button. Upon releasing the button the oscillator continues to run at the desired rate and the regenerated sync pulses (*h*) are free from incoming interference. Should the incoming line frequency rise, the bridge will be opened earlier when the applied sawtooth is negative. A negative voltage will appear across C12 and be filtered by R12C13 taking the grid of V4a negative. A negative change applied to this grid reduces the period during which V4b is cut off and hence increases the

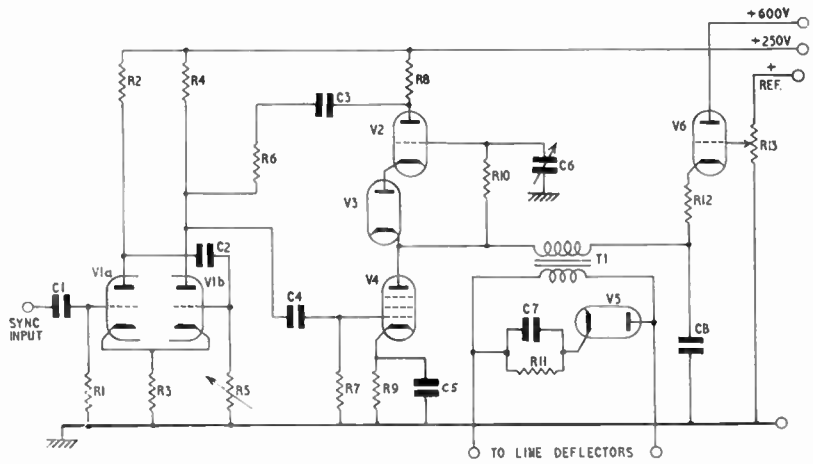


Fig. 13. Line scan generator.

oscillator frequency. Thus, the regenerated line sync pulses follow the mean incoming sync rate.

5.3. *Line Scan Generator*

The load presented to the output stage of the line scan generator is predominantly inductive and a sensibly linear current flows in the transformer primary during the scan period when connected to a constant voltage supply. The circuit is broken at the end of each line and the load current reverses via its shunt self-capacitance.

A simplified circuit diagram is shown in Fig. 13.

Valves V1a and V1b are connected as a cathode-coupled multivibrator, the free running speed of which can be adjusted by means of a variable resistor R5 in the grid circuit of V1b. Synchronizing pulses of negative polarity are applied to the grid of V1a following which V1b takes a pulse of current during retrace, producing an anode voltage waveform consisting of a negative pulse during flyback plus a positive-going sawtooth during scan. The negative-going pulse cuts off valve V4 and allows the output transformer primary voltage to perform a positive half cycle of oscillation, at the end of which it is caught by the diode V3 and the pentode V2, both of which conduct, developing across the anode load R8 a voltage drop which assists in keeping valve V4 cut off. Since the current in the deflectors after the flyback is steadily falling then the current through V2 must similarly decrease. A sawtooth waveform, which is produced by integrating the flyback voltage across the output transformer by R10C6, is fed to the grid of V2 causing its current to drop and anode potential to rise. This is communicated to the grid of V4 and conduction commences.

The circuit constants are such that V2 tends to a cut-off condition, the drive for V4 being maintained by the positive-going sawtooth from V1b.

The d.c. supply to the primary of the output transformer is fed from a stabilizer of variable voltage providing means for scan amplitude control. A parabolic waveform is impressed on this supply, by means of an integrating circuit R12C8, to correct in part for pincushion error.

The line deflector coils are driven in parallel from the secondary of the output transformer and a shunt connected damping diode prevents current ripple following flyback.

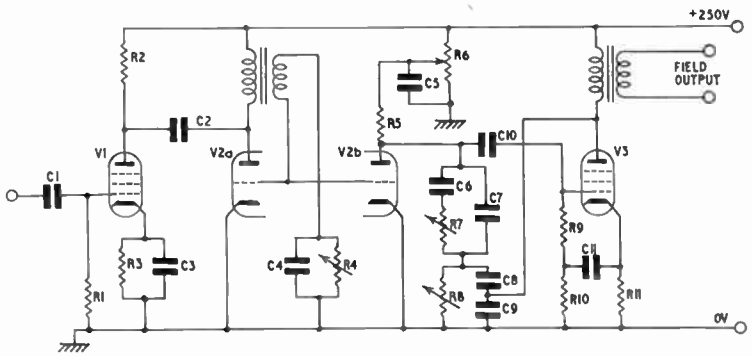


Fig. 14. Field scan generator.

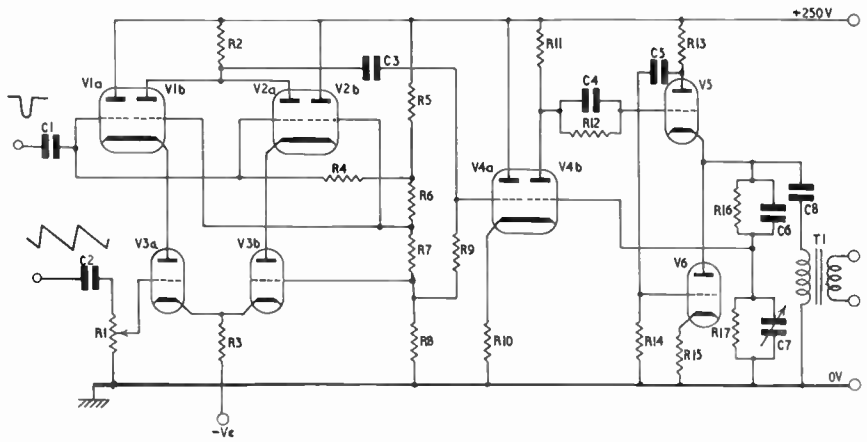


Fig. 15. Keystone corrector.

5.4. *Field Scan Generator*

This unit comprises a triggered blocking oscillator, sawtooth generator and output stage.

The circuit is shown in Fig. 14. The positive-going mixed sync applied to V1 initiates a rapid fall in anode potential of V1 which triggers the blocking oscillator via C2. By normal blocking oscillator action the grids of both halves of V2 are driven positively, conduction of V2b discharging the anode capacitance C7, C8, C9.

During the scan period this capacitance recharges via R5 to a value set by R6 and the positive-going sawtooth causes V3 to take an increasing current. The required deflector current is of sawtooth form and feedback from both anode and cathode to the grid circuit offsets non-linearity of the output stage.

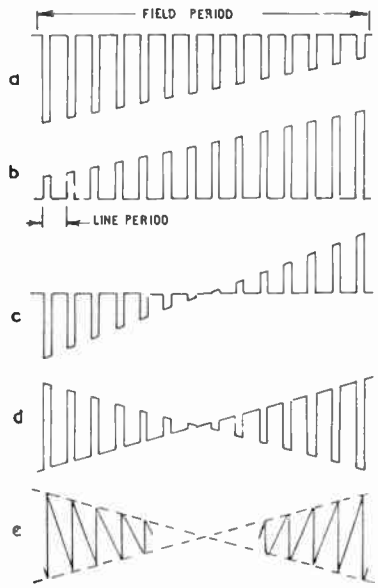


Fig. 16. Keystone corrector waveforms.

5.5. *Keystone Corrector*

The desired deflection current for keystone correction was discussed in Section 3 and a circuit[6] that has been used for generating this waveform is shown in Fig. 15.

Valves V1a and V1b form a cathode coupled pair with V3a as the cathode load. V1a is normally on. Negative-going line sync pulses are applied to the grid of V1a, causing negative-going pulses to be developed across R2. A negative-going field sawtooth is applied to the grid of V3a which, with V3b, forms a second cathode-coupled pair. At the start of the field sawtooth the current flowing in V3a is a maximum but decreases throughout the field. As a result the negative-going line pulses across R2 likewise decrease in amplitude throughout the field period, as shown at (*a*), Fig. 16.

The valves V2a, V2b and V3b are complementary to V1a, V1b and V3a with the result that line pulses flowing in V2a have a minimum amplitude at the start of the field and increase throughout the cycle (see waveform (*b*)). The anode load R2, however, is common to V1b and V2a, hence throughout a field there will be negative-going pulses of maximum amplitude decreasing to zero and then increasing in a positive direction to a maximum at the end of the field, as shown in (*c*). This waveform is amplified in the cathode coupled pair V4 and applied to a shunt-regulated cathode follower stage which feeds the output transformer T1. Linearization is obtained by applying voltage feedback from the output stage to V4 via R16 C6, R17 C7.

The transformer primary is a.c. coupled and the secondary voltage will be as shown in (*d*). The keystone deflectors, which are predominantly inductive,

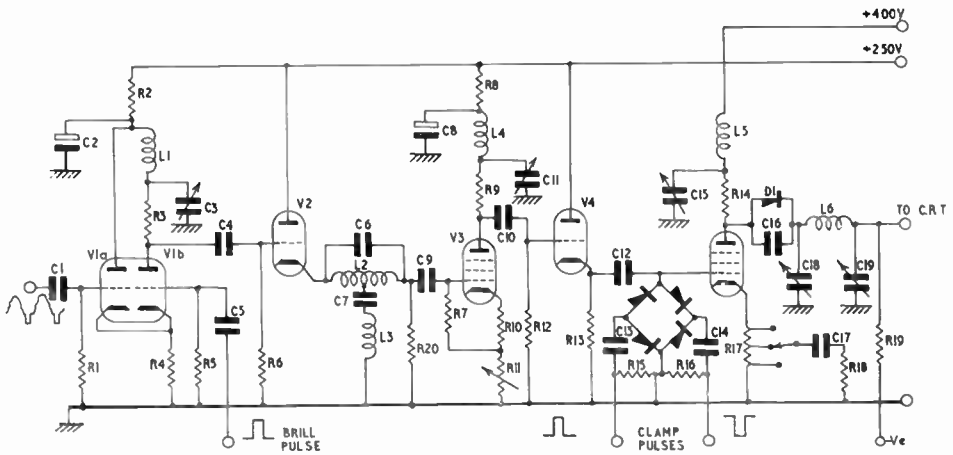


Fig. 17. Video amplifier.

are connected across the secondary and their current will be the integral of the applied waveform shown in (*e*), which is as required for keystone correction. The amplitude of the correcting current is adjusted by the potentiometer R1 which controls the amount of frame sawtooth applied to V3a.

5.6. *Video Amplifier*

Three similar video amplifiers are provided, each having means for gain adjustment and d.c. insertion as described in Section 3.7.

Simple anode correction is used in the early stages, and a four terminal filter network couples the output valve to the c.r.t. cathode. Phase error introduced by this network is corrected by a bridged-T section,[7] and the overall amplifier response is flat within $\pm\frac{1}{2}$ dB to $5\frac{1}{2}$ Mc/s, followed by a relatively sharp cut-off.

Coupling network time constants have been balanced to reduce l.f. error and a keyed diode clamp establishes the d.c. level of the signal at the last amplifying stage.

A simplified circuit of one amplifier is shown in Fig. 17.

The positive-going video signal, of between 0·7 and 1·0 volt amplitude, is amplified and inverted in the cathode-coupled pair V1a and V1b. The common black level pulse is applied to the grid of V1b with the result that the waveform at the anode of V1b comprises the video signal with the addition of the black level pulse during the back porch period. The cathode follower V2 provides a low impedance source for the phase correction network C6 L2 C7 L3. Further amplification is provided by V3, the variable cathode resistor R11 providing means for adjustment of the overall amplifier gain.

A second cathode-follower V4 provides a low driving point impedance for the four-diode clamp circuit, enabling the desired reference potential to be rapidly established during the clamping period. The output stage V7 has a load comprising a low-pass filter with a terminating half section having an m of 0·65[8].

The video output is fed to the cathode of its respective c.r.t. via the series limiter comprising diode D1, which has a shunt capacitor C16, and a resistor R19 through which flows the beam current. Under normal conditions the voltage developed across this resistor is such that the c.r.t. cathode potential is less than that of the anode of the output valve and the diode conducts. As the video drive increases the anode potential falls and the drop across R19 increases until a condition is established when the diode ceases to conduct, limiting the mean c.r.t. current. The network R17 C17 R18 in the cathode provides compensation for h.f. loss which may be present due to the use of long video feed cables.

5.7. *Stabilized R.F. E.H.T. Unit*

The 30 kV cathode-ray tube supply must be capable of delivering currents ranging from a few microamperes up to peak values of $2\frac{1}{2}$ mA, with a permissible voltage drop of less than 1%, if picture degradation due to defocusing is to be avoided.

Although they are more complex, r.f. e.h.t. units are generally used in preference to direct power-frequency supplies, for this and similar applications, since they are less bulky and inherently safer.

A block diagram of the unit that was developed for this equipment is shown in Fig. 18.

The sinusoidal output from a Hartley oscillator operating at approximately 75 kc/s is first shaped into a square wave of unity mark/space ratio. The square wave is used as the grid drive for a bank of power tetrodes having as an anode load a tuned circuit capacitively coupled to a step-up transformer.

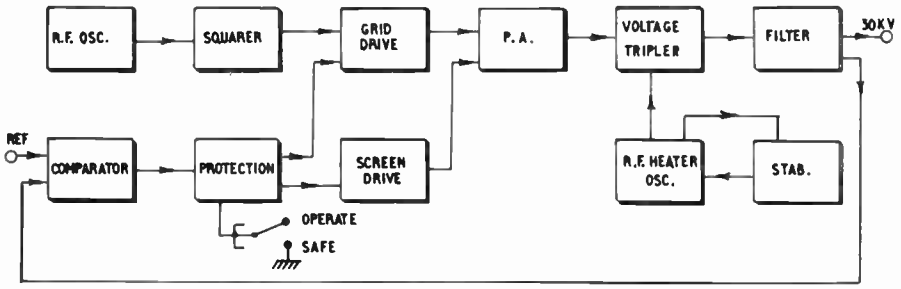


Fig. 18. R.f. e.h.t. unit.

Across the tuned secondary of this transformer a peak sinusoidal voltage of 10 kV is developed which is rectified by diodes connected in a voltage tripler circuit to produce 30 kV. Resistor-capacitor filtering is used and a fraction of the d.c. output is returned to a comparator unit where it is compared with a negative reference voltage. The comparator output provides, via a protection stage, the d.c. bias for the square wave drive to the power amplifier stage such that a fall in output voltage reduces the bias, thus compensating for the increased load. This action is augmented by a second path which controls the screen supply to the power amplifier valves.

When the projector is operated in damp conditions it is advisable, on first switching, to apply the e.h.t. gradually in order to avoid brush discharges. For this purpose the reference supply is applied to the comparator via a potentiometer and when its slider is at the earthy end the drive stage is cut off, resulting in zero output. The potentiometer may thus be used as a voltage control, maximum output being obtained when utilizing the full reference voltage.

For safety during maintenance it is desirable to provide means for preventing the accidental application of the e.h.t. and a line has been provided which when earthed, either by local or remote switches, renders the unit safe.

The cathodes of the e.h.t. rectifiers in the tripler unit are subject to high d.c. and a.c. potentials and their heater supplies must be adequately insulated. A separate stabilized r.f. oscillator working at approximately 1·2 Mc/s provides heater excitation via a suitable isolating transformer.

A series limiting resistor is connected within the unit in the output lead, limiting the instantaneous discharge current in the event of tube flash-over, to that which may be drawn from the energy stored in the co-axial connecting cables between the unit and c.r.t.s.

5.8. *Power Supplies*

Separate stabilized power supplies have been provided for the display and r.f. e.h.t units. Each is of conventional design, bridge-connected silicon rectifiers delivering current via a choke input filter to series connected stabilizer valves. A neon tube is used as a voltage reference source.

A separate stabilized current supply is provided for the three sets of focus and shift coils which are all serially connected.

The main focus control determines the supply current and for individual adjustment rheostats are connected in shunt with each focus coil. Current of either polarity is obtained for shift purposes by connecting each coil across a bridge network comprising two similar series resistors in shunt with a potentiometer.

5.9. *Switching*

When the e.h.t. is applied the current density of the focused cathode-ray tube beam is such that in the absence of either line or field deflection serious damage may be caused to both the phosphor and face-plate.

Protection has been provided against failure of either scan generator, loss of c.r.t. cooling air or premature application of e.h.t., either when switching on or following a temporary loss of incoming mains power.

Outputs from both the line and field scan generators are rectified, yielding voltages which, when both are present, allow a valve to conduct and operate a relay. When the relay is at rest both the focus current stabilizer and e.h.t. unit are rendered inoperative. Ideally the presence of scan fields should be detected at each of the deflectors but in this equipment it has been found sufficient to monitor the output stage of each main scan generator.

When the equipment is consistently used in a dusty atmosphere regular maintenance of the air filters is essential. Should this be overlooked the flow of cooling air to the c.r.t.s can be greatly reduced. A vane operated micro-switch is positioned in the main air trunking to the upper optical section and, in the absence of adequate air flow, earths the e.h.t. safety line.

A circuit diagram of the switching and protection system is shown in Fig. 19.

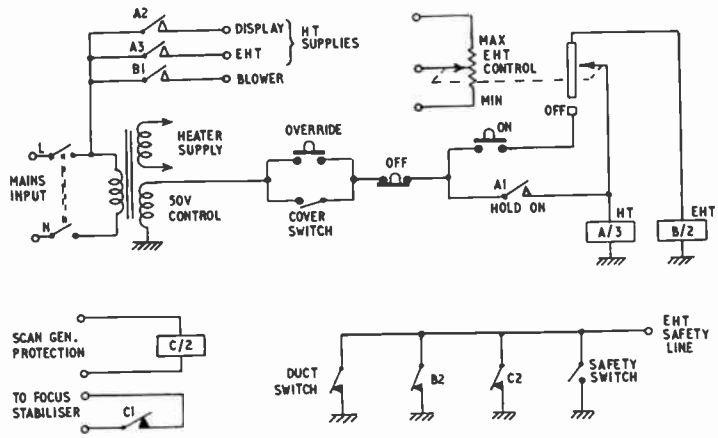


Fig. 19. Switching system.

After the heaters have been on for a short period the h.t. supply may be applied via relay A, providing the top cover switch is closed and the e.h.t. control is in the OFF position. Following a second short period the e.h.t. control may be advanced, energizing relay B and, providing sufficient c.r.t. cooling air is flowing, the e.h.t. is applied.

An over-ride push-button has been provided on the control panel to enable the cover to be removed without interrupting the operation of the projector.

The diagram also includes the scan generator protection relay and the e.h.t. safety rail which must be isolated from earth at each of four positions before the e.h.t. can be applied.

6. Performance

The three projected television images are formed upon a viewing screen, normally by front projection. The brightness distribution will depend upon the reflectance characteristics of the screen and the incident angles of projection of the three light sources. Compared with the central green projector, the red and blue projectors are angled with respect to the screen. A plain matt screen having a co-sinusoidal intensity distribution will give a picture free from colour shading but for many applications is inefficient since it reflects light over a greater angle than is normally used for viewing.[9] Materials are available which, for incident light normal to the screen, concentrate the reflections over an area in front of the screen at the expense of peripheral viewing brightness. The ratio between normally reflected light for such a screen and an ideal diffusing surface is defined as the screen gain.

Generally a narrow reflectance polar diagram is required in the vertical direction, but the horizontal polar diagram must be sufficiently wide to avoid undesirable shading across the picture due to the non-coincidence of light sources. The screen material normally used has lenticular embossing on a plastic base and is sprayed with a reflective compound, giving a forward gain of approximately 2:1 with an acceptable brightness distribution. Beam trimming plates have been provided adjacent to the red and blue barrels in order to correct for residual side to side shading.

When the majority of viewers are likely to be well behind the projector the screen may, with advantage, be tilted back to obtain a more suitable light distribution and, within limits, optical focus can be corrected by off-setting the tube alignment controls.

For certain applications back projection is desirable and again, care must be taken to select a screen material having suitable transmission and front reflection characteristics. Besides reversing the direction of line scan when used in this manner the correction required for keystone error of the red and blue images has to be completely reversed.

The projector should be positioned so that the green optical axis is normal to the centre of the screen. The throw distance between the corrector plate and the screen should be as specified for the plate in use but variations of up to $\pm 10\%$ may be permitted for a slight loss in picture quality.

The toe-in of the outer barrels is normally fixed but may be checked by optical siting through the c.r.t. support tubes.

Optical adjustments, apart from checking focus, are limited to the alignment of the corrector plates with respect to the mirrors subsequent to their removal, and the horizontal and vertical pivoting about the c.r.t. face following a tube change.

The raster may be centralized on the tube face by pivoting the focus coil upon its mounting tube, and the beam aligned within the focus coil by adjustment of a ring magnet, enabling changes of focus to be made without introducing picture shift. The green

image is obtained from the central Schmidt system, consequently suffering the least geometrical distortion, and the picture size and linearity may be adjusted with this one display only.

Positional alignment is best achieved by using a grid pattern and adjusting separately the red and blue images to that of the green. Coarse adjustment of amplitude and linearity may be made with the images displaced slightly by the shift controls. All images are finally superimposed by slight adjustments of shift, rhombic and keystone controls.

When so aligned the positional accuracy is better than 2% within a circle of diameter equal to the picture width, and the registration error is within one picture point inside a circle of diameter equal to the picture height, falling to a maximum error of two picture points outside the area encompassed by a circle having a diameter equal to the picture width. The equipment has been designed for use with standard inputs of either 405 lines 50 fields, 625 lines 50 fields or 525 lines 60 fields, 2 : 1 interlaced, either composite or non-composite 0·7–1·0 V positive-going video and 2 V negative-going mixed syncs, the latter being looped or terminated in 75 ohms.

The single phase a.c. mains input may be in the range 110–120 V or 200–250 V, 50 to 60 cycles with a permissible supply voltage variation of ±6%.

The video amplifier response is flat up to $5\frac{1}{2}$ Mc/s and there is no significant degradation of 625-line C.C.I.R. pictures.

7. Applications

In addition to their use for entertainment, projectors of this type are in service in a number of specialized installations. For demonstrations to student audiences, subjects ranging from operative surgery to advanced cookery are regularly being shown.

Having developed this equipment specifically for colour a number are now being used for the display of monochrome pictures using three white tubes!

In the field of air traffic control and radar plotting, there is a need for large area displays of aircraft and target movements, and colour television projection is being increasingly used.

One interesting application is in connection with flight simulators which are being used by a number of aircraft operators for crew training. To practise the approach and take-off phases of a flight, added realism is achieved by presenting to the pilot a view of the area over which he is flying. The required picture is obtained by flying a television camera over a model of an airfield and its surrounding country and the image is projected on to a screen in front of the cockpit. Monochrome and colour displays are being used but the latter is preferable for night training where the correct interpretation of coloured lights is most important. Since it is necessary to mount the viewing screen close to the cockpit window the picture fidelity must be good and the line structure indiscernible.

For both general and specialized applications development is continuing in order that equipment size and weight may be reduced while picture size and brightness are increased.

8. Acknowledgments

The author is indebted to many colleagues who have contributed towards the success of this project, especially Mr. A. R. Tingley, who was responsible for much of the original development work, and to the Directors of the Rank Cintel Division of the Rank Organization for permission to publish this paper.

9. References

1. L. L. Evans and R. V. Little, "Large-screen colour-television projection", *J. Soc. Mot. Pict. Telev. Engrs*, **64**, pp. 169–73, April 1955.
2. T. Poorter and F. W. de Vrijer, "The projection of colour-television pictures", *Philips Tech. Rev.*, **19**, pp. 339–55, No. 12, 1957/8.
3. S. L. Bendell and W. J. Neely, "Medium screen colour television projection" *J. Soc. Mot. Pict. Telev. Engrs*, 67, pp. 166–8, March 1958.
4. P. C. Jansen, "Television optics", *Elect. Applic. Bull.*, **11**, No. 5, 1950/51.
5. J. E. B. Jacob, "High performance television monitors", *J. Brit.I.R.E.*, **10**, pp. 158–75, April 1950.
6. J. L. E. Baldwin, British Patent No. 859,238.
7. T. C. Nutall, "Some aspects of television circuit techniques: phase correction and gamma correction", *J. Television Soc.*, **5**, No. 9, pp. 257–65, March 1949.
8. H. A. Wheeler, "Wide-band amplifiers for television", *Proc. Inst. Radio Engrs*, **27**, pp. 429–38, July 1939.
9. E. W. D'Arcy and G. Lessman, "Objective evaluation of projection screens", *J. Soc. Mot. Pict. Telev. Engrs*, **61**, No. 6, pp. 702–20, December 1953.

Manuscript first received by the Institution on 12th March 1962 and in final form on 8th June 1962. (Paper No. 804/T20.)

U.D.C. 621.372.852.22

Loss Properties of Cylindrical Waveguides Containing Gyromagnetic Media

By

R. A. WALDRON, M.A., (*Member*) †

AND

Mrs. D. J. BOWE, B.Sc. †

Summary: The system considered is a waveguide of radius a containing a concentric rod of gyromagnetic material of radius b. The attenuation constant is given by

$$\beta'' = \frac{\partial\beta}{\partial\varepsilon}\varepsilon'' + \frac{\partial\beta}{\partial\mu}\mu'' + \frac{\partial\beta}{\partial\alpha}\alpha''$$

where ε, μ, and α are the relative permittivity and the relative permeability elements of the rod, and ε'', μ'', α'' are their imaginary parts. The quantities $\partial\beta/\partial\varepsilon$, $\partial\beta/\partial\mu$, $\partial\beta/\partial\alpha$, have been calculated for a number of values of ε, μ, α, b/a, a/λ_0, λ_0 being the wavelength of a plane wave in the material occupying the part of the guide not occupied by gyromagnetic material. Comparisons are made with experimental losses measured by other authors; good agreement is found. The optimum design for Faraday-rotation devices is discussed.

List of Symbols

- a Radius of waveguide
- b Radius of gyromagnetic material
- ε Relative permittivity of the gyromagnetic material
- μ, α Diagonal and off-diagonal elements, respectively, of the relative permeability tensor
- ε_0, μ_0 Permittivity and permeability of the material in the guide other than the gyromagnetic material
- λ_0 Wavelength of a plane wave in a material of permittivity ε_0 and permeability μ_0. Thus $\lambda_0 = 2\pi/\omega\sqrt{\varepsilon_0\mu_0}$
- β Phase constant
- $\bar{\beta}$ Normalized phase constant $= \lambda_0\beta/2\pi$

1. Introduction

The system we are concerned with in this paper is that of a waveguide of circular cross-section, of radius a, containing a concentric rod of gyromagnetic material, of radius b, the space in the waveguide outside the rod being filled with an isotropic material which may, in particular, be vacuum or air. We shall write ε_0 and μ_0 for the permittivity and permeability of this outer medium, and λ_0 for the wavelength, at the working frequency, of an electromagnetic wave propagating in an infinite extent of the outer medium, thus

$$\lambda_0 = 2\pi/\omega\sqrt{\varepsilon_0\mu_0} \qquad \ldots\ldots(1)$$

It is convenient to normalize the guide radius with respect to λ_0, i.e. a/λ_0 is taken as a parameter rather than a. The rod radius might also be normalized in the same way, but it is more convenient to think in terms of b/a, which remains constant as λ_0 or ω varies.

The gyromagnetic medium is characterized by its relative permittivity, ε, and its tensor relative permeability,

$$[\mu] = \begin{bmatrix} \mu & -j\alpha & 0 \\ j\alpha & \mu & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \ldots\ldots(2)$$

These quantities are expressed relative to the medium surrounding the gyromagnetic material, i.e. the absolute permittivity and permeability are $\varepsilon\varepsilon_0$ and $\mu_0[\mu]$.

The parameters of the system are thus a/λ_0, b/a, ε, μ, and α, and it is completely described by the values of these quantities. We shall also be concerned with the system of a waveguide containing a dielectric rod, which may be regarded as a special case of the guide with a gyromagnetic rod, with $\mu = 1$, $\alpha = 0$. Whenever the gyromagnetic material is referred to in this paper, the dielectric case is to be taken as also implied.

Waveguides containing gyromagnetic media as concentric rods or tubes have been extensively studied[1–18] theoretically, in the lossless case, i.e. for real values of the parameters. No theory appears to have been given, however, for such waveguides when losses are present, and such a study forms the subject of the present work.

2. Types of Loss

In a practical waveguide containing a ferrite rod, there are four kinds of loss—wall loss due to the resistance of the metallic wall to the currents flowing in it, dielectric loss in the material surrounding the ferrite, dielectric loss in the ferrite, and magnetic loss in the ferrite. The wall loss is likely to be small com-

† Marconi's Wireless Telegraph Co. Ltd., Research Division, Great Baddow, Essex.

pared with the loss in the ferrite, and we shall neglect it, treating the wall as perfectly conducting.

The other types of loss will give rise to imaginary parts of ε_0, $\varepsilon\varepsilon_0$, μ, and α; we accordingly write

$$\left.\begin{aligned}\varepsilon_0 &= \varepsilon_0' - j\varepsilon_0''\\ \mu &= \mu' - j\mu''\\ \alpha &= \alpha' - j\alpha''\end{aligned}\right\} \quad \ldots\ldots(3)$$

where it is assumed that $\varepsilon_0'' \ll \varepsilon_0'$, $\mu'' \ll \mu'$, $\alpha'' \ll \alpha'$. For ε, we write

$$\varepsilon\varepsilon_0 = (p - jq)\varepsilon_0 = (p - jq)(\varepsilon_0' - j\varepsilon_0'')$$
$$= \varepsilon_0'\left\{p - jq - j\frac{\varepsilon_0''}{\varepsilon_0'}p\right\}$$

neglecting the second-order small term $\varepsilon_0'' q$. We now write $p = \varepsilon'$, $q + \varepsilon_0'' p/\varepsilon_0' = \varepsilon''$. Then

$$\varepsilon\varepsilon_0 = \varepsilon_0'(\varepsilon' - j\varepsilon'') \quad \ldots\ldots(4)$$

The reason for adopting this notation will become clear in the next section. Notice that in general $\varepsilon \neq \varepsilon' - j\varepsilon''$. If ε_0 is real, or if $\varepsilon_0'' p/\varepsilon_0' \ll q$, then

$$\varepsilon = \varepsilon' - j\varepsilon'' \quad \ldots\ldots(5)$$

When ε_0 is complex, so is λ_0, and from equation (1) we obtain

$$\lambda_0' + j\lambda_0'' = \frac{2\pi}{\omega\sqrt{\varepsilon_0'\mu_0}}(1 + j\varepsilon_0''/2\varepsilon_0')$$

i.e.
$$\lambda_0' + j\lambda_0'' = \lambda_0'(1 + j\varepsilon_0''/2\varepsilon_0') \quad \ldots\ldots(6)$$

The parameter a/λ_0 now becomes

$$a/\lambda_0 = (a/\lambda_0')(1 - j\varepsilon_0''/2\varepsilon_0') \quad \ldots\ldots(7)$$

3. Method of Calculation and Attenuation

The method of calculation we shall follow here is that suggested in reference 5. First we write the characteristic equation in the form

$$\bar{\beta} = \bar{\beta}(a/\lambda_0, \varepsilon, \mu, \alpha, b/a) \quad \ldots\ldots(8)$$

where $\bar{\beta}$ is the normalized phase constant, $= \lambda_0\beta/2\pi = \lambda_0/\lambda\hat{g}$, $\lambda\hat{g}$ being the wavelength in the guide. The normalized attenuation constant may be regarded as an imaginary part of $\bar{\beta}$, and we write

$$\bar{\beta} = \bar{\beta}' - j\bar{\beta}'' \quad \ldots\ldots(9)$$

Our purpose is to calculate $\bar{\beta}''$ when some or all of the parameters become imaginary.

Differentiating equation (8), we obtain

$$\delta\bar{\beta} = \frac{\partial\bar{\beta}}{\partial(a/\lambda_0)}\delta(a/\lambda_0) + \frac{\partial\bar{\beta}}{\partial\varepsilon}\delta\varepsilon + \frac{\partial\bar{\beta}}{\partial\mu}\delta\mu + \frac{\partial\bar{\beta}}{\partial\alpha}\delta\alpha \ldots(10)$$

and letting $\delta(a/\lambda_0)$, etc., be the small imaginary parts given in equations (3), (4), and (7), this gives

$$\bar{\beta}'' = \frac{\partial\bar{\beta}}{\partial(a/\lambda_0)}\frac{\varepsilon_0''}{2\varepsilon_0'}\frac{a}{\lambda_0'} + \frac{\partial\bar{\beta}}{\partial\varepsilon}\varepsilon'' + \frac{\partial\bar{\beta}}{\partial\mu}\mu'' + \frac{\partial\bar{\beta}}{\partial\alpha}\alpha'' \ldots(11)$$

which holds as long as ε_0'', ε'', μ'', α'', and all the partial derivatives are sufficiently small.

The reason for defining ε'' as in equation (4) can now be seen. We are to calculate values of $\bar{\beta}''$ due to imaginary parts of the mathematical parameters involved in equation (8); since ε_0 is not a parameter of the system, losses due to the imaginary part of ε_0 must be expressed as factors of the parameters a/λ_0 and ε, and the imaginary parts of the mathematical parameters are then not quite the same things as the imaginary parts of the properties of the media present.

If the medium in the region $a > r > b$ is relatively lossless, as is likely to be the case, we may neglect the effect of ε_0''. This applies especially when the medium is vacuum, or, which is virtually the same thing, air. Equation (11) then becomes

$$\bar{\beta}'' = \frac{\partial\bar{\beta}}{\partial\varepsilon}\varepsilon'' + \frac{\partial\bar{\beta}}{\partial\mu}\mu'' + \frac{\partial\bar{\beta}}{\partial\alpha}\alpha'' \quad \ldots\ldots(12)$$

In equation (12), ε'' is given by equation (5). As a special case, the rod may be of dielectric, with $\mu = 1$ and $\alpha = 0$. Then $\mu'' = \alpha'' = 0$, and the attenuation is given by

$$\bar{\beta}'' = \frac{\partial\bar{\beta}}{\partial\varepsilon}\varepsilon'' \quad \ldots\ldots(13)$$

In order to calculate the attenuation in a waveguide, the quantities ε'', μ'', and α'' must be known. It may be assumed that they are known as a result of, for example, cavity measurements on a sample of the gyromagnetic material. Alternatively, if $\bar{\beta}''$ is measured and any two of ε'', μ'', and α'' are known, the third may be calculated. In either case, it is necessary to know $\partial\bar{\beta}/\partial\varepsilon$, $\partial\bar{\beta}/\partial\mu$, and $\partial\bar{\beta}/\partial\alpha$.

If $\mu'' = \alpha'' = 0$, the attenuation is given by equation (13), and we may say immediately that $\bar{\beta}''$ will show the same behaviour, as a function of b/a or of frequency, as does $\partial\bar{\beta}/\partial\varepsilon$, since ε'' is only a scale factor. The position is not so simple if μ'' and α'' are not zero.

The characteristic equation is extremely complicated, and to obtain the partial derivatives from it by differentiation would be a formidable task. Even if this were done, the result would be an even more complicated equation than the characteristic equation itself. However, if a number of values of $\bar{\beta}$ are available, as a function of the parameters, for the lossless case, it is possible to obtain approximate numerical values for the partial derivatives very easily. For example, to obtain $\partial\bar{\beta}/\partial\varepsilon$, a graph is plotted of $\bar{\beta}$ against ε, and the slope is read off by means of a transparent ruler. In this way, if $\bar{\beta}$ is known for enough values of ε, it is often possible to determine $\partial\bar{\beta}/\partial\varepsilon$ even for a value of ε for which $\bar{\beta}$ itself is not known.

A number of values of $\bar{\beta}$ were given in reference 5, and these have been used in the present calculations.

Occasional gaps were filled in by further computations of $\bar{\beta}$ by the method of reference 5, and a number of new computations, particularly in the neighbourhood of cut-off, were also undertaken.

4. Results

The results of numerical computations of $\partial\bar{\beta}/\partial\varepsilon$, $\partial\bar{\beta}/\partial\mu$, and $\partial\bar{\beta}/\partial\alpha$, are presented in Tables 1–11 for various values of the parameters. This information is not sufficient in itself to enable the graphs of Figs. 1–4 to be drawn, and additional information has been used which is deducible from the general forms of the curves of reference 5, although accurate quantitative results are not obtainable. It has also been helpful, when a series of curves are to be drawn for several values of a parameter, the other parameters being constant, to consider all the curves together, the general outline of one curve being drawn so as to conform with the general outlines of the others. For example, in Fig. 1 it is not always clear where the peak of the curve is to be situated. By drawing all the curves for a single value of α on the same graph, and drawing the locus of the peaks through those peaks which are sufficiently well defined, a guide has been obtained for drawing the peaks in the cases where they are not well defined by the numerical results alone.

The curves shown in Figs. 1–4 do not represent all the results. All the results have, however, been traced graphically, and those presented here were chosen as typical of the complete results. They are shown here to illustrate the general behaviour, and as a basis for discussion.

Figures 1–3 are drawn for $a/\lambda_0 = 0{\cdot}3$; for this value of a/λ_0, the empty guide supports propagation with a finite real value of $\bar{\beta}$. When a thin rod of ferrite or dielectric is introduced, the value of $\bar{\beta}$ rises, slowly at first, then more rapidly—very rapidly if ε is large—and then flattens off. Correspondingly, the values of $\partial\bar{\beta}/\partial\varepsilon$, $\partial\bar{\beta}/\partial\mu$, and $\partial\bar{\beta}/\partial\alpha$ rise slowly at first, then more rapidly. For sufficiently high values of ε, μ, and α, a peak is reached corresponding to the maximum slope of the $\bar{\beta}$ curve; if ε, μ, and α are not sufficiently high, an actual peak does not occur, but, as is seen, for example, in Fig. 1(*a*) for $\alpha = 0$, there is a suggestion of the peak, and the curve is clearly a development of those shown in Figs. 1(*b*)–(*d*).

It will be noticed, in Fig. 1, that as ε increases, for small values of b/a, the loss decreases. This phenomenon is presumably related to the fact, noted in reference 7, that for a very thin rod the power in the region occupied by the rod is actually less than the power in this region before introducing the rod. This is evident from Figs. 5 and 6 of reference 7. The same effect appears to occur for $\partial\bar{\beta}/\partial\mu$ and $\partial\bar{\beta}/\partial\alpha$, although it is not so marked as for $\partial\bar{\beta}/\partial\varepsilon$, and the computations are not sufficiently detailed nor accurate to demonstrate it clearly.

When a/λ_0 is less than 0·29303, the empty guide is cut off. For given values of a/λ_0, ε, μ, and α, a value of b/a can then be found for which $\bar{\beta} = 0$. A set of such values of the parameters is called a cut-off point, and a number of such sets of values are given in reference 5. It was found that for H modes in general, and in particular the H_{11} mode that we are now studying, a cut-off value of b/a depends only on a/λ_0 and ε, not on μ or α. Thus finite changes of μ or α cause no change of $\bar{\beta}$ at $\bar{\beta}=0$, so that $\partial\bar{\beta}/\partial\mu=\partial\bar{\beta}/\partial\alpha=0$ at cut-off. On the other hand, a finite change in ε causes a finite change in the cut-off value of b/a. It has been shown[16,19] that $\partial\bar{\beta}/\partial(b/a)$ is infinite at $\bar{\beta} = 0$ for the case $\mu = 1$, $\alpha = 0$. Writing

$$\frac{\mathrm{d}\bar{\beta}}{\mathrm{d}(b/a)} = \frac{\mathrm{d}\bar{\beta}}{\mathrm{d}\varepsilon}\frac{\mathrm{d}\varepsilon}{\mathrm{d}(b/a)} \qquad \text{.....(14)}$$

it follows that $\mathrm{d}\bar{\beta}/\mathrm{d}\varepsilon$ is also infinite at $\bar{\beta} = 0$. When $\mu \neq 1$ or $\alpha \neq 0$, it is difficult to see whether $\mathrm{d}\bar{\beta}/\mathrm{d}\varepsilon$ becomes infinite or not, but it will certainly rise to a

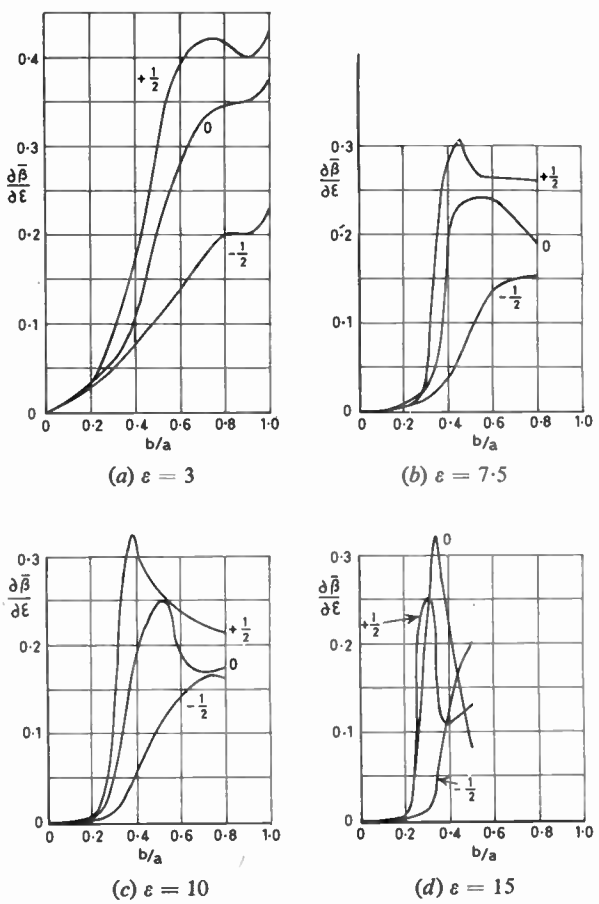


Fig. 1.

Curves of $\partial\bar{\beta}/\partial\varepsilon$ (ordinates) against b/a (abscissae) for $a/\lambda_0 = 0{\cdot}3$, $\mu = 1$, and various values of ε. The values of α are given with each curve.

high value. Even for $\mu = 1$ and $\alpha = 0$, $d\bar{\beta}/d\varepsilon$ only becomes infinite at cut-off when $\varepsilon'' = 0$, which will not be the case in practice. All we can say, then, is that $d\bar{\beta}/d\varepsilon$ becomes very large. In view of this, to decide whether or not $d\bar{\beta}/d\varepsilon$ becomes infinite at $\bar{\beta} = 0$ for $\mu \neq 1$ or $\alpha \neq 0$ is only of academic interest, and we shall not pursue the matter further. In Table 1, $d\bar{\beta}/d\varepsilon$ is given as infinite at cut-off; the sign ∞ is to be interpreted in the light of the foregoing discussion as meaning very large in a practical case.

For sufficiently high values of ε, μ, and α, it was found that the curve of $\bar{\beta}$ against b/a bulges into values of b/a below the cut-off value, giving another point of infinite $d\bar{\beta}/d(b/a)$, this time for a finite value of $\bar{\beta}$. A typical curve of this type is illustrated in Fig. 5, with $a/\lambda_0 = 0{\cdot}2$, $\varepsilon = 15$, $\mu = 1$, $\alpha = 0$. $d\bar{\beta}/d(b/a)$ is infinite at the cut-off point C and at M. The value of b/a at M is the minimum for which propagation can take place. Equation (14) shows that $d\bar{\beta}/d\varepsilon$ is again infinite at M, since $d\varepsilon/d(b/a)$ is finite. Also, at M, $d\alpha/d(b/a)$ and $d\mu/d(b/a)$ are finite, and by analogous equations to (14), clearly $d\bar{\beta}/d\alpha$ and $d\bar{\beta}/d\mu$ are infinite. These arguments hold whenever there is a point such as M for a value of b/a less than the cut-off value; $\partial\bar{\beta}/\partial\varepsilon$, $\partial\bar{\beta}/\partial\mu$, and $\partial\bar{\beta}/\partial\alpha$ then always become infinite (in the lossless case).

Between M and C, there is a point of inflection, indicated by E in Fig. 5. Hence $d\bar{\beta}/d(b/a)$ has a minimum. Now rewrite equation (14) in the form

$$\frac{d\bar{\beta}}{d\varepsilon} = \frac{d\bar{\beta}}{d(b/a)} \frac{d(b/a)}{d\varepsilon} \qquad \ldots\ldots(15)$$

If $d(b/a)/d\varepsilon$ is monotonic between M and C, there will be one and only one minimum of $d\bar{\beta}/d\varepsilon$ corresponding

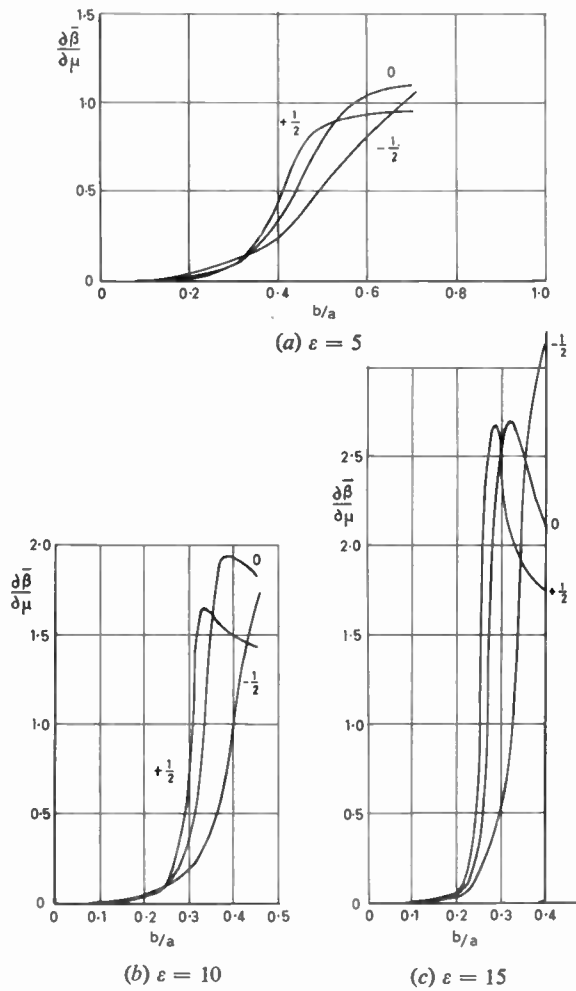


Fig. 2.

Curves of $\partial\bar{\beta}/\partial\mu$ (ordinates) against b/a (abscissae) for $a/\lambda_0 = 0{\cdot}3$, $\mu = 1$, and various values of ε. The values of α are given with each curve.

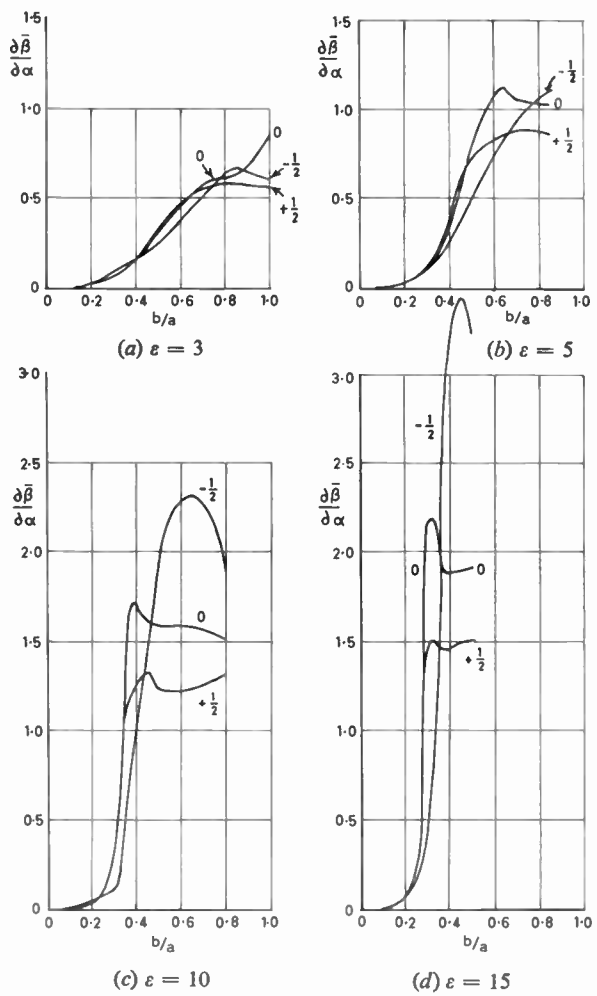


Fig. 3.

Curves of $\partial\bar{\beta}/\partial\alpha$ (ordinates) against b/a (abscissae) for $a/\lambda_0 = 0{\cdot}3$, $\mu = 1$, and various values of ε. The values of α are given with each curve.

to the minimum of $d\bar{\beta}/d(b/a)$, in the range of values of b/a between those at M and at C. It can be shown from the results of reference 5 that $d(b/a)/d\varepsilon$ is indeed monotonic, at least for $|\alpha| < |\mu|$, for ranges of values of the parameters that are of practical interest. In fact, $d(b/a)/d\varepsilon$ is always negative, with positive slope, and it follows that the minimum of $|d\bar{\beta}/d\varepsilon|$ occurs at a higher value of b/a (nearer to cut-off) than the minimum of $d\beta/d(b/a)$. It will be seen that the minimum of $|d\bar{\beta}/d\varepsilon|$ for $a/\lambda_0 = 0{\cdot}2$ in Fig. 4 occurs for a higher value of b/a than the point of inflection, E, in Fig. 5; the parameter values are the same for these two curves.

Since $|d\bar{\beta}/d\varepsilon|$ has a minimum, and is infinite at M, it must increase as b/a approaches the cut-off value from below. This is in keeping with the point made above that $d\beta/d\varepsilon$ becomes large at cut-off.

The relation between the curve of $d\bar{\beta}/d\varepsilon$ for $a/\lambda_0 = 0{\cdot}2$, with its two infinities, and the curves of $d\beta/d\varepsilon$ for $a/\lambda_0 > 0{\cdot}29303$ can be seen by plotting a number of curves, for various values of a/λ_0, as in Fig. 4. These curves are all plotted for the case $\varepsilon = 15$, $\mu = 1$, $\alpha = 0$, and the curve for $a/\lambda_0 = 0{\cdot}3$ is the same as the curve in Fig. 1(*d*). Consider first the curve for $a/\lambda_0 = 0{\cdot}29303$. For values of b/a greater than 0, it behaves in the same way as the curves for $a/\lambda_0 = 0{\cdot}3$ and $a/\lambda_0 = 0{\cdot}4$, staying small at first as b/a increases, then rising sharply to a peak, and falling off again for higher values of b/a. One difference, however, is that for $a/\lambda_0 = 0{\cdot}29303$ the curve comes in to the origin as b/a decreases, and then rises instantaneously to infinity, i.e. the line $b/a = 0$—the axis of $d\beta/d\varepsilon$—is also part of the curve. This behaviour may be seen as the limit of the process which is apparent in the curve for $a/\lambda_0 = 0{\cdot}275$; the curve appears to approach zero as b/a approaches the cut-off value, then turns upwards sharply and goes to infinity. As a/λ_0 increases from 0·25 through 0·275 to 0·29303, the approach to zero at cut-off becomes closer and closer, and zero is actually just reached at $a/\lambda_0 = 0{\cdot}29303$. With further increase in a/λ_0, the curves must terminate at the origin, for $d\beta/d\varepsilon$ must obviously be zero for an empty guide.

Another process is apparent in the curves of Fig. 4. As a/λ_0 decreases through values greater than 0·29303, the height of the peak decreases; this is also apparent from the tabulated results for other parameter values. As a/λ_0 decreases to 0·275 and 0·25, the height of the peak increases. The minimum occurs between $a/\lambda_0 = 0{\cdot}3$ and $a/\lambda_0 = 0{\cdot}275$; it is not unlikely that it occurs at $a/\lambda_0 = 0{\cdot}29303$, but it is not worth while investigating further, as its precise location is not important. What is interesting is the increase of peak

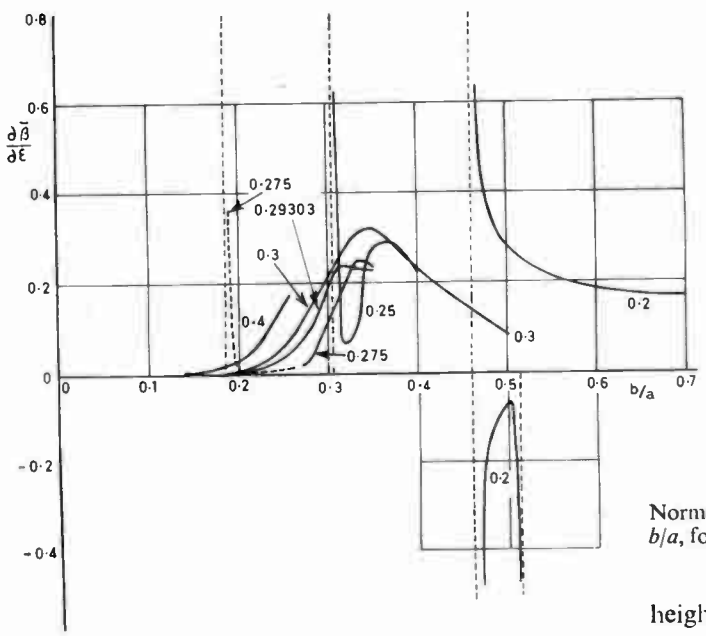


Fig. 4.

Curves of $\partial\bar{\beta}/\partial\varepsilon$ against b/a for $\varepsilon = 15$, $\mu = 1$, $\alpha = 0$; the values of a/λ_0 are given for each curve. The dotted curve is not accurately computed. Dashed lines are asymptotes. For $a/\lambda_0 < 0{\cdot}29303$, the curves go sharply to infinity at the cut-off value of b/a. For $a/\lambda_0 > 0{\cdot}29303$, the curves all terminate at the origin, lying very close to the b/a axis for small b/a. For $a/\lambda_0 = 0{\cdot}29303$, the curve reaches the origin, and then coincides with the $\partial\bar{\beta}/\partial\varepsilon$ axis.

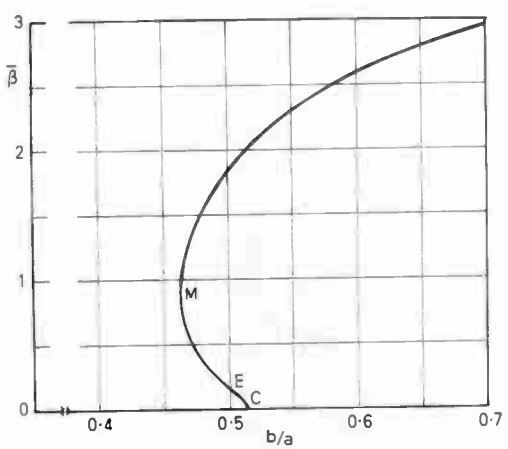


Fig. 5.

Normalized phase constant, $\bar{\beta}$, as a function of radius ratio, b/a, for $a/\lambda_0 = 0{\cdot}2$, $\varepsilon = 15$, $\mu = 1$, $\alpha = 0$. C is the cut-off point, and M the point of minimum radius ratio.

height as a/λ_0 decreases, and the fact that the peak moves towards higher values of b/a. The cut-off point also moves towards higher values of b/a as a/λ_0 decreases, and it moves more rapidly than the peak. There will be a value of a/λ_0, less than 0·25, where the cut-off point just catches up the peak. Here the peak will have risen steeply, just reaching infinity, so that as the infinity of $d\bar{\beta}/d(b/a)$ just appears (Fig. 5), with M and C having the same value of b/a, so the second

infinity of $d\bar{\beta}/d\varepsilon$ just appears, merged with the infinity at cut-off. As a/λ_0 decreases further, the cut-off infinity of $d\bar{\beta}/d\varepsilon$ overtakes the infinity corresponding to M (Fig. 4), and the curve takes the form shown for $a/\lambda_0 = 0{\cdot}2$. For still smaller values of a/λ_0, the M-type infinity will overtake the cut-off infinity, and then become a peak instead of an infinity. The curve will then be qualitatively similar to those for $a/\lambda_0 = 0{\cdot}25$ and $0{\cdot}275$. The maximum will eventually reach $b/a = 1$ and disappear, and later, at about $a/\lambda_0 = 0{\cdot}076$, the cut-off itself will reach $b/a = 1$ and disappear.

When there are two infinities, i.e. when there is a point M lying to the left of C (Fig. 5), $d\bar{\beta}/d\varepsilon$ is negative, as is seen in Fig. 4 for $a/\lambda_0 = 0{\cdot}2$. This appears to indicate a negative loss, i.e. a growth of the wave. In fact, it actually does indicate an increase in power as the wave progresses—progresses, that is, in the positive direction, with the wave front. But between M and C (Fig. 5) the negative slope implies a backward wave,[20] so that the direction of energy flow is opposite to that of the motion of the wavefront. The negative value of $d\bar{\beta}/d\varepsilon$ thus represents a decrease in power as the energy flows along the guide, which is what one expects.

The values of $\partial\bar{\beta}/\partial\mu$ and $\partial\bar{\beta}/\partial\alpha$ are also negative for the backward-wave region. No accurate numerical information has been obtained, but it is possible to see, qualitatively, how the curves behave. They are very similar to the curve shown for $\partial\bar{\beta}/\partial\varepsilon$, except close to cut-off, where, instead of turning sharply and going to minus infinity, they continue to increase (algebraically), becoming zero at the cut-off point ($b/a = 0{\cdot}514$ for the case $a/\lambda_0 = 0{\cdot}2$).

Before closing this section, a note on the accuracy of the tabulated results is in order. All values of $\partial\bar{\beta}/\partial\varepsilon$, $\partial\bar{\beta}/\partial\mu$, and $\partial\bar{\beta}/\partial\alpha$ are accurate to about ± 2 in the least significant figure given, unless this is enclosed in brackets, in which case the accuracy is about ± 4 in the least significant figure.

5. Comparison with Experimental Results

Before discussing the calculated results further, it is desirable to consider how far the theoretical results agree with practical results. Not many practical results have been given in the literature, but what there are support the theory.

Fox, Miller, and Weiss[21] give a curve of dielectric loss against radius ratio, b/a, for the parameter values $a/\lambda_0 = 0{\cdot}4$, $\varepsilon = 10$; the values of μ and α are not given. The loss is expressed in the form of loss per unit mass of ferrite, which is equivalent to loss divided by b^2. The actual loss is therefore given by multiplying the values on the Fox, Miller, and Weiss curve by b^2. The curve obtained in this way is very similar to the curve obtained from Table 7. A typical experimental curve is also given, which is of the same form, as far as it goes, as the theoretical curve. The peak is not observed. Fox, Miller, and Weiss point out that as b/a is increased, a point is reached when the E_{11} mode starts to propagate. For high values of ε, the E_{11} mode can propagate for values of b/a lower than that at which the peak occurs, so that the peak is difficult to observe. For lower values of ε, the peak can be observed.

Rizzi[15] gives a curve of loss, on a decibel scale, against ferrite rod diameter for a cylindrical guide containing a concentric rod of ferrite. The loss is seen to increase slowly at first, then rapidly as the ferrite rod radius becomes greater than about 0·3 of the guide radius. This behaviour is qualitatively similar to the results of Tables 4, 5 and 6, illustrated in Figs. 1, 2 and 3. Duncan and Swern[22] also give a curve of attenuation against ferrite diameter for a negative circularly polarized wave, showing the same steep rise above a certain value of rod diameter, followed by a peak.

6. Applications to the Design of Faraday-Rotation Devices

For values of $a/\lambda_0 > 0{\cdot}29303$, the curves of $\partial\bar{\beta}/\partial\varepsilon$, $\partial\bar{\beta}/\partial\mu$, and $\partial\bar{\beta}/\partial\alpha$ are similar in form. Because of this, equation (12) indicates that $\bar{\beta}''$, as a function of b/a, will be similar in form to the partial derivatives. It was not possible to foresee this before the partial derivatives were known. This fact accounts for the similarity noted in section 5 between practical results and computed values of the partial derivatives. Because of the similarity of the computed results, the form of $\bar{\beta}''$ will not be greatly dependent on the relative magnitudes of ε'', μ'', and α''. Therefore guides containing different ferrites, with the same values of ε', μ', and α', but different values of ε'', μ'', and α'', will exhibit the same sort of loss behaviour, although the absolute magnitude of $\bar{\beta}''$ will depend on ε'', μ'', and α''.

Comparison of the results of the present paper with those for the phase constants given in reference 5 shows that the loss peak is associated with the steep rise of the phase constant curve as b/a increases. It is in this region that most Faraday-rotation devices operate; the rotation increases more rapidly than the loss, so that the rotation per unit loss actually decreases as the peak loss is approached. Notice that $\partial\bar{\beta}/\partial\mu$ and $\partial\bar{\beta}/\partial\alpha$ are roughly equal, and considerably greater than $\partial\bar{\beta}/\partial\varepsilon$. Unless ε'' is considerably greater, therefore, than α'' and μ'', magnetic loss predominates over dielectric, and determines the general behaviour of the overall loss. The results of reference 5 show that $\bar{\beta}$ is largely linear with α, so that the specific Faraday rotation, equal to the difference of values of $\bar{\beta}$ for positive and negative α, is roughly proportional to $\partial\bar{\beta}/\partial\alpha$. If α'' were considerably greater than μ'', we should expect, therefore, that Faraday rotation and

loss would increase proportionately, so that there would be no optimum choice of ferrite radius giving a maximum rotation per unit loss. The fact that such a maximum is observed suggests that either μ'' is at least comparable with α'', or that $\varepsilon'' \gg \alpha''$, or both. This appears to agree with some published results of measurements of ε'', μ'', and α'',[23] but not with others.[24] However, the evidence is far from being conclusive, and this question requires further study.

It is shown in reference 5 that as the radius of the ferrite increases, the specific Faraday rotation increases rapidly when the phase constant is increasing rapidly, and then, for higher values of b/a, flattens off to a very flat maximum. At this maximum, the frequency dependence of Faraday rotation is small. This prediction has since been verified practically.[25] The maximum of Faraday rotation corresponds to a low loss, according to the present results. It thus seems that higher values of b/a should be used in Faraday-rotation devices than is customary; there would be advantages in increased bandwidth and decreased loss. However, for typical values of guide radius (a/λ_0 in the range 0·3 to 0·45) the E_{11} mode is able to propagate as well as the H_{11}, when b/a is sufficiently high for these features of the H_{11} mode to come into being. The difficulty can be overcome by using a value of a/λ_0 considerably below 0·29303, when the E_{11} mode is cut-off for all values of b/a, while the maximum of specific Faraday rotation occurs for a suitable value of b/a.

7. Conclusions

The behaviour of magnetic and dielectric loss, as functions of the geometry of the system and properties of the ferrite, have been studied extensively and intensively. The results indicate that optimum design of Faraday-rotation devices can be achieved using a guide of small radius, so small that it would be cut off when empty, and then choosing the ratio b/a, in relation to the ferrite properties, to give maximum specific rotation. A broad-band device would result, with no multimoding, as shown in reference 5, and the present results indicate that under those circumstances the loss, both magnetic and dielectric, would be minimized.

8. Acknowledgments

The authors wish to thank Mrs. E. M. Seabrook and Miss P. Ramm for assistance with the computations. Permission to publish this paper has been given by the Director of Research of Marconi's Wireless Telegraph Company.

9. References

1. Hideya Gamo, "The Faraday rotation of waves in a circular waveguide", *J. Phys. Soc. Japan*, **8**, pp. 176–82, 1952.
2. H. Suhl and L. R. Walker, "Topics in guided wave propagation through gyromagnetic media", *Bell. Syst. Tech. J.*, **33**, pp. 579–659, 939–86, 1133–94, 1954.
3. M. L. Kales, "Modes in waveguides containing ferrites", *J. Appl. Phys.*, **24**, pp. 604–8, 1953.
4. Ll. G. Chambers, "Propagation in a ferrite-filled waveguide", *Quart. J. Mech. Appl. Maths*, **8A**, pp. 435–47, 1955.
5. R. A. Waldron, "Electromagnetic wave propagation in cylindrical waveguides containing gyromagnetic media", *J. Brit.I.R.E.*, **18**, pp. 597–612, 677–90, 733–46, 1958.
6. R. A. Waldron, "Theory of the mode spectra of cylindrical waveguides containing gyromagnetic media", *J. Brit.I.R.E.*, **19**, pp. 347–56, 1959.
7. R. A. Waldron, "Feature of cylindrical waveguides containing gyromagnetic media", *J. Brit.I.R.E.*, **20**, pp. 695–706, 1960.
8. R. A. Waldron, "Field Components and Power in Cylindrical Waveguides Containing Ferrites". Paper presented at the 13th General Assembly of U.R.S.I., London, September 1960.
9. J. E. Tompkins, "Multimode propagation in gyromagnetic rods and its application to travelling-wave devices", *J. Appl. Phys., Suppl.*, **31**, No. 5, pp. 176S–177S, May 1960.
10. J. E. Tompkins, "Energy Distribution and Field Configuration in Ferrites", pp. 169–80 of "Solid State Physics in Electronics and Telecommunications", Vol. 3, edited by M. Desirant and J. L. Michiels, (Academic Press, New York, 1960).
11. A. P. van Gelder, A. M. de Graaf and R. Kronig, "New calculations on the Faraday effect in wave-guides", *Appl. Sci. Res.*, **7B**, pp. 441–8, 1959.
12. R. Kronig, "Het Faraday-effect in Golfgeleiders", *Ned. T. Natuurk*, **25**, pp. 217–26, 1959.
13. H. K. F. Severin, "Propagation constants of circular cylindrical waveguides containing ferrites", *Trans. Inst. Radio Engrs*, (*Microwave Theory and Techniques*), **MTT-7**, pp. 337–46, 1959.
14. A. J. Baden-Fuller, "Microwave propagation through round waveguide partially filled with ferrite", *Proc. Instn Elect. Engrs*, **108C**, pp. 339–48, September 1961 (I.E.E. Monograph No. 430E).
15. P. A. Rizzi, "High-power ferrite circulators", *Trans. Inst. Radio Engrs*, (*Microwave Theory and Techniques*), **MTT-5**, pp. 230–7, 1957.
16. P. J. B. Clarricoats, "Properties of Waveguides Containing Ferrites with Special Reference to Waveguides of Circular Cross-Section", Ph.D. Thesis, University of London, 1958.
17. P. J. B. Clarricoats, "A perturbation method for circular waveguides containing ferrites", *Proc. Instn Elect. Engrs*, **106B**, pp. 335–40, May 1959 (I.E.E. Paper No. 2796E).
18. R. A. Waldron, "Properties of ferrite-loaded cylindrical waveguides in the neighbourhood of cut-off", *Proc. Instn Elect. Engrs*, **109B**, Suppl. No. 21, pp. 90–94, 1962. (I.E.E. Paper, June 1961.)
19. P. J. B. Clarricoats, "Propagation along unbounded and bounded dielectric rods, part 2: propagation along a dielectric rod contained in a circular waveguide", *Proc. Instn Elect. Engrs*, **108C**, pp. 177–86, 1961. (I.E.E. Monograph No. 410E, October 1960.)
20. P. J. B. Clarricoats and R. A. Waldron, "Non-periodic slow-wave and backward-wave structures", *J. Electronics Control*, **8**, pp. 455–8, 1960.
21. A. G. Fox, S. E. Miller and M. T. Weiss, "Behaviour and applications of ferrites in the microwave region", *Bell Syst. Tech. J.*, **34**, pp. 5–103, 1955.

22. B. J. Duncan and L. Swern, "Effects of zero ferrite permeability on circularly polarized waves", *Proc. Inst. Radio Engrs*, **45**, pp. 647–55, 1957.

23. E. G. Spencer, R. C. LeCraw and F. Reggia, "Measurement of microwave dielectric constants and tensor permeabilities of ferrite spheres", *Convention Record of the I.R.E.*, **4**, Part 8, pp. 113–21, 1955.

24. C. M. Srivastava and J. Roberts, "Measurements of ferrite loss-factors at 10 Gc/s", *Proc. Instn Elect. Engrs*, **105B**, pp. 204–9, April 1958 (I.E.E. Paper No. 2518R.).

25. S. J. Lewandowski and J. Konopka, "On some problems in designing microwave Faraday-rotation devices", *Trans. Instn Radio Engrs*, (*Microwave Theory and Techniques*), **MTT-8**, pp. 249–51, 1960.

Table 1

Values of $\partial\bar{\beta}/\partial\varepsilon$ for $a/\lambda_0 = 0{\cdot}2$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\varepsilon$									b/a	$\partial\bar{\beta}/\partial\varepsilon$
	$\varepsilon = 3$			$\varepsilon = 5$			$\varepsilon = 10$				$\varepsilon = 15$
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$		$\alpha = 0$
0·546							∞	∞	∞	0·514	∞
0·55							0·17 (5)	0·46	0·64	0·50	− 0·07 (0)
0·56								0·4 (2)		0·49	− 0·11 (0)
0·57								0·3 (9)		0·48	− 0·15 (0)
0·58								0·3 (6)		0·47	− 0·44 (0)
0·59								0·3 (3)		0·462	∞
0·60							0·10 (0)	0·29	0·34 (1)	0·47	0·51 (5)
0·632				∞	∞	∞				0·48	0·36 (5)
0·65				0·34	0·44	0·85	0·09 (1)	0·25	0·29 (1)	0·49	0·31 (5)
0·70				0·21 (8)	0·31 (3)	0·53	0·09 (5)	0·23 (3)	0·26 (9)	0·50	0·28 (0)
0·75	∞	∞	∞	0·17 (9)	0·29 (6)	0·41 (3)	0·09 (8)	0·22 (5)	0·26 (5)	0·51	0·26 (0)
0·80	0·5 (9)	0·6 (5)	0·98	0·17 (3)	0·28 (2)	0·38 (2)	0·10 (6)	0·22 (3)	0·25 (5)	0·55	0·210
0·85	0·40	0·48	0·69	0·17 (4)	0·27 (3)	0·36 (1)	0·11 (2)	0·21 (5)	0·24 (9)	0·60	0·187
0·90	0·37	0·42	0·61	0·18 (0)	0·26 (1)	0·36 (5)	0·10 (5)	0·19 (3)	0·23 (3)	0·65	0·173
0·95	0·36	0·45	0·58	0·20 (6)	0·26 (5)	0·36 (6)	0·10 (0)	0·18 (3)	0·22 (0)	0·70	0·166
1·00	0·36	0·45 (1)	0·54	0·23 (2)	0·27 (4)	0·37 (8)		0·17 (6)	0·21 (3)		

Table 2

Values of $\partial\bar{\beta}/\partial\mu$ for $a/\lambda_0 = 0{\cdot}2$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\mu$								
	$\varepsilon = 5$			$\varepsilon = 10$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0·514							0	0	0
0·462								∞	
0·43									∞
0·51						∞			
0·543					∞				
0·546				0	0	0			
0·55							1·9 (1)	2·67	1·83
0·60				0·89	2·26	1·66	3·17	2·27	1·74
0·632	0	0	0						
0·65	0·15	0·21	0·40	1·36	2·00	1·51	3·26	2·26	1·68
0·70	0·32	0·61	0·72	1·58	1·88	1·42			
0·75	0·45	0·68	0·74	1·78	1·78	1·41			
0·80	0·61	0·74	0·73	1·77	1·7 (4)	1·34			
0·85	0·69	0·77	0·72						
0·90	0·82	0·78	0·71						
0·95	0·94	0·79	0·69						
1·00	1·16	0·85	0·71						

Table 3

Values of $\partial\bar{\beta}/\partial\alpha$ for $a/\lambda_0 = 0{\cdot}2$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\alpha$											
	$\varepsilon = 3$			$\varepsilon = 5$			$\varepsilon = 10$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = -\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0·514										0	0	0
0·462											∞	
0·43												∞
0·51									∞			
0·543								∞				
0·546							0	0	0			
0·55										1·8 (2)	2·17	1·50
0·60							0·76	2·11	1·28	3·6 (4)	2·10	1·50
0·632				0	0	0						
0·65				0·16 (0)	0·23 (6)	0·35 (4)	1·21	1·88	1·28	3·33	2·05	1·50
0·70				0·34 (0)	0·48 (0)	0·60 (8)	1·59	1·74	1·31	3·17	2·17	1·40
0·75	0	0	0	0·45 (6)	0·62 (0)	0·68 (0)	1·76	1·69	1·28			
0·80	0·15 (4)	0·12 (8)	0·13 (2)	0·58 (0)	0·68 (4)	0·73 (8)	1·86	1·64	1·26			
0·85	0·28 (8)	0·24 (0)	0·22 (4)	0·72 (2)	0·72 (6)	0·73 (0)	1·84	1·62	1·26			
0·90	0·39 (0)	0·30 (6)	0·28 (0)	0·82 (4)	0·74 (4)	0·71 (6)	1·78	1·55	1·26			
0·95	0·50 (8)	0·35 (4)	0·31 (8)	0·95 (8)	0·72 (0)	0·68 (6)	1·62	1·47	1·26			
1·00	0·63 (8)	0·40 (4)	0·32 (6)	1·18	0·70 (6)	0·62 (0)		1·20	1·08			

Table 4

Values of $\partial\bar{\beta}/\partial\varepsilon$ for $a/\lambda_0 = 0{\cdot}3$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\varepsilon$											
	$\varepsilon = 2$			$\varepsilon = 3$			$\varepsilon = 5$			$\varepsilon = 7{\cdot}5$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0	0	0	0	0	0	0	0	0	0	0	0	0
0·10	0·021	0·018	0·022 (3)	0·011 (5)	0·012 (6)	0·011 (8)	0·004 (0)	0·004 (1)	0·005 (0)	0·002 (8)	0·002 (4)	0·002 (8)
0·20	0·06	0·05 (5)	0·06	0·02 (9)	0·03 (8)	0·03 (1)	0·014 (0)	0·012	0·017 (8)	0·006 (8)	0·008 (3)	0·009 (3)
0·25												0·01 (0)
0·30	0·09	0·13	0·15	0·05	0·05 (7)	0·09 (3)	0·021	0·03 (3)	0·04 (5)	0·014 (8)	0·02 (6)	0·03 (8)
0·35									0·07 (3)			0·19
0·40	0·16	0·2 (0)	0·24	0·08	0·1 (1)	0·16 (8)	0·03 (4)	0·08	0·12 (0)	0·03 (8)	0·18 (3)	0·28 (3)
0·45									0·23 (0)			0·30 (8)
0·50	0·2 (0)	0·2 (7)	0·33 (5)	0·10 (5)	0·20 (5)	0·30 (5)	0·07 (3)	0·16 (3)	0·32 (0)	0·08 (8)	0·24	0·28 (3)
0·55									0·31 (8)			0·26 (5)
0·60	0·2 (6)	0·3 (4)	0·47	0·14	0·27 (8)	0·40 (0)	0·08 (5)	0·19 (0)	0·30 (5)	0·13 (8)	0·24	0·26 (5)
0·70	0·2 (9)	0·4 (0)	0·52	0·1 (6)	0·33	0·41 (8)	0·09 (4)	0·21 (6)	0·30 (5)	0·14 (8)	0·21 (5)	0·26 (3)
0·80	0·3 (3)	0·4 (5)	0·57	0·2 (0)	0·34 (4)	0·41 (8)	0·12 (5)	0·24 (5)	0·30 (3)	0·15 (3)	0·19 (0)	0·26 (0)
0·90	0·3 (5)	0·5 (1)	0·57	0·2 (0)	0·35 (0)	0·40						
1·00	0·3 (6)	0·5 (6)	0·62	0·2 (3)	0·37 (5)	0·43						

Table 4 (continued)

b/a	Values of $\partial\bar{\beta}/\partial\varepsilon$								
	$\varepsilon = 10$			$\varepsilon = 12{\cdot}5$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0	0	0	0	0	0	0	0	0	0
0·10	0·002 (3)	0·001 (6)	0·001 (0)	0·001 (0)	0·001 (2)	0·000 (7)	0·001 (0)	0·000 (4)	0·001 (0)
0·20	0·003 (3)	0·006 (0)	0·006 (8)	0·003 (5)	0·003 (8)	0·006 (3)	0·003 (0)	0·004 (3)	0·005 (5)
0·25			0·003 (3)			0·02 (8)		0·05 (8)	0·08
0·30	0·012 (8)	0·05 (3)	0·13	0·011 (0)	0·11	0·2 (9)	0·011 (5)	0·21	0·2 (5)
0·35		0·12	0·28 (5)	0·03 (5)		0·27	0·06 (3)	0·32	0·22
0·40	0·06 (0)	0·19 (3)	0·31 (8)	0·08 (3)	0·19 (8)	0·2 (5)	0·12 (0)	0·23 (3)	0·11
0·45		0·22 (3)	0·28 (0)	0·13 (0)	0·19 (0)	0·23	0·17 (5)	0·15 (3)	0·12
0·50	0·11 (3)	0·24 (8)	0·26 (5)	0·13 (7)	0·17	0·21 (8)	0·19 (8)	0·08 (3)	0·13
0·55			0·25 (0)						
0·60	0·13 (8)	0·19	0·24 (0)						
0·70	0·16 (5)	0·17	0·22 (3)						
0·80	0·16 (3)	0·17 (5)	0·21 (3)						

Table 5

Values of $\partial\bar{\beta}/\partial\mu$ for $a/\lambda_0 = 0{\cdot}3$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\mu$								
	$\varepsilon = 5$			$\varepsilon = 10$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0	0	0	0	0	0	0	0	0	0
0·10	0·005	0·005	0·003	0·005	0·003	0·003	0·008	0·005	0·003
0·20	0·035	0·023	0·018	0·045	0·033	0·028	0·028	0·048	0·05
0·25					0·100	0·11	0·16	0·34	0·66
0·30	0·12	0·09	0·09	0·18	0·35	0·68	0·55	2·54	2·48
0·35				0·41	1·49	1·62	2·34	2·53	1·90
0·40	0·24	0·34	0·45	0·96	1·93	1·50	3·15	2·11	1·75
0·45				1·68	1·83	1·44			
0·50	0·53	0·80	0·86						
0·60	0·80	1·04	0·93						
0·70	1·04	1·10	0·95						

Table 6

Values of $\partial\bar{\beta}/\partial\alpha$ for $a/\lambda_0 = 0{\cdot}3$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\alpha$								
	$\varepsilon = 1$			$\varepsilon = 3$			$\varepsilon = 5$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0	0	0	0	0	0	0	0	0	0
0·10	0·011	0·002	0·000	0·0020	0·0020	0·0020	0·0033	0·0033	0·0033
0·20	0·025	0·007	0·004	0·02 (2)	0·02 (2)	0·01 (4)	0·03 (8)	0·02 (4)	0·01 (4)
0·30	0·038	0·019	0·009	0·08 (6)	0·06 (4)	0·06 (0)	0·09 (8)	0·10 (8)	0·08 (6)
0·35									0·20 (2)
0·40	0·046	0·023	0·022	0·16 (4)	0·15 (0)	0·14 (4)	0·24 (6)	0·31 (4)	0·38 (0)
0·45									0·60 (2)
0·50	0·062	0·045	0·034	0·25 (2)	0·29 (0)	0·31 (6)	0·49	0·77 (6)	0·73
0·55								0·93 (4)	
0·60	0·075	0·067	0·037	0·39 (4)	0·45 (0)	0·47 (0)	0·72	1·06	0·82
0·65								1·12 (4)	
0·70	0·086	0·071	0·046	0·49 (2)	0·57 (8)	0·54 (0)	0·93	1·06 (2)	0·88
0·75								1·04 (4)	
0·80	0·100	0·084	0·066	0·63 (2)	0·61 (6)	0·58 (4)	1·05	1·03 (0)	0·87
0·85							1·10	1·02	0·86
0·90	0·109	0·096	0·055	0·64 (2)	0·67 (6)	0·57 (0)			
1·00	0·115	0·107	0·049	0·61	0·86 (0)	0·55 (6)			

Table 6 (continued)

b/a	Values of $\partial\bar{\beta}/\partial\alpha$					
	$\varepsilon = 10$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0	0	0	0	0	0	0
0·10	0·0040	0·0040	0·0040	0·0040	0·0040	0·0040
0·20	0·05	0·03	0·01	0·06	0·06	0·04
0·25		0·10	0·06	0·14	0·25	0·30
0·30	0·10	0·34	0·31	0·4 (5)	2·17	1·4 (3)
0·35	0·48	1·23	1·11	1·3 (3)	2·11	1·46
0·40	0·9 (8)	1·70	1·20	3·1 (6)	1·88	1·47
0·45	1·4 (3)	1·61	1·34	3·4 (4)	1·90	1·51
0·50	1·9 (6)	1·58	1·22	3·2 (4)	1·92	1·51
0·55	2·19	1·58				
0·60	2·28	1·58	1·25			
0·65	2·32	1·58				
0·70	2·26	1·57	1·24			
0·75	2·16	1·54				
0·80	1·92	1·52	1·31			

Table 7

Values of $\partial\bar{\beta}/\partial\varepsilon$ for $a/\lambda_0 = 0{\cdot}4$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\varepsilon$											
	$\varepsilon = 1$			$\varepsilon = 5$			$\varepsilon = 10$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0	0	0	0	0	0	0	0	0	0	0	0	0
0·05	0·00 (7)	0·01 (1)	0·00 (7)	0·000 (4)	0·000 (7)	0·000 (8)	0·000 (2)	0·000 (2)	0·000 (2)	0·000 (1)	0·000 (1)	0·000 (1)
0·10	0·016	0·018	0·016	0·002 (0)	0·002 (4)	0·002 (2)	0·000 (7)	0·000 (9)	0·000 (9)	0·000 (2)	0·000 (2)	0·000 (2)
0·15	0·027	0·033	0·032	0·004 (8)	0·006 (4)	0·008 (4)	0·002 (2)	0·003 (4)	0·004 (2)	0·001 (0)	0·002	0·003 (8)
0·20	0·054	0·057	0·06 (2)	0·009 (1)	0·014	0·01 (2)	0·004 (5)	0·012	0·03 (5)	0·004 (0)	0·026	0·14
0·25	0·081	0·078	0·10	0·015 (5)	0·02 (5)	0·05	0·011 (8)	0·07 (4)	0·18	0·022	0·15	0·45
0·30	0·112	0·13	0·12	0·025	0·06	0·09	0·032	0·1 (9)	0·4 (8)	0·106		
0·35	0·14 (5)	0·21	0·19	0·028	0·08	0·14		0·3 (6)				
0·40	0·17 (8)	0·27	0·2 (8)	0·033	0·11	0·2 (3)		0·4 (4)				
0·45	0·21 (2)	0·35	0·3 (3)	0·039	0·15	0·2 (7)		0·5 (1)				
0·50	0·25 (8)	0·4	0·3 (6)									
0·60	0·34		0·4 (1)									
0·70	0·43		0·4 (4)									
0·80	0·52		0·4 (7)									

Table 8

Values of $\partial\bar{\beta}/\partial\mu$ for $a/\lambda_0 = 0{\cdot}4$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\mu$								
	$\varepsilon = 5$			$\varepsilon = 10$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$	$\alpha = -\frac{1}{2}$	$\alpha = 0$	$\alpha = +\frac{1}{2}$
0	0	0	0	0	0	0	0	0	0
0·05	0·008	0·003	0·000	0·003	0·003	0·003	0·003	0·000	0·000
0·10	0·015	0·008	0·008	0·015	0·010	0·008	0·018	0·010	0·010
0·15	0·035	0·025	0·020	0·045	0·025	0·030	0·055	0·050	0·050
0·20	0·070	0·058	0·050	0·11	0·11	0·17	0·19	0·54	1·30
0·25	0·13	0·13	0·11 (5)	0·29	0·83	1·33	0·64	2·54	1·80
0·30	0·21	0·29	0·36	0·69	1·78	1·49			
0·35	0·33	0·60	0·70						
0·40	0·51	0·88	0·86						
0·45	0·66	1·04	0·89						

Table 9

Values of $\partial\bar{\beta}/\partial\alpha$ for $a/\lambda_0 = 0{\cdot}4$, $\mu = 1$

b/a	Values of $\partial\bar{\beta}/\partial\alpha$											
	$\varepsilon = 1$			$\varepsilon = 5$			$\varepsilon = 10$			$\varepsilon = 15$		
	$\alpha = -\frac{1}{4}$	$\alpha = 0$	$\alpha = +\frac{1}{4}$	$\alpha = -\frac{1}{4}$	$\alpha = 0$	$\alpha = +\frac{1}{4}$	$\alpha = -\frac{1}{4}$	$\alpha = 0$	$\alpha = +\frac{1}{4}$	$\alpha = -\frac{1}{4}$	$\alpha = 0$	$\alpha = +\frac{1}{4}$
0	0	0	0	0	0	0	0	0	0	0	0	0
0·05				0·004	0·002	0·001	0·004	0·002	0·002	0·004	0·002	0·001
0·10	0·008	0·007	0·004	0·014	0·010	0·006	0·017	0·010	0·006	0·016	0·012	0·008
0·15				0·036	0·022	0·019	0·104	0·134	0·126	0·056	0·050	0·042
0·20	0·05 (2)	0·028	0·01 (0)	0·072	0·055	0·048	0·22 (1)	0·68 (3)	0·88	0·36 (4)	0·42 (7)	0·68 (6)
0·25				0·129	0·130	0·125	0·80	1·56	1·16	1·03	1·94	1·74
0·30	0·09 (0)	0·059	0·04 (6)	0·21 (0)	0·28 (4)	0·32 (3)						
0·35				0·32	0·55	0·59						
0·40	0·12	0·098	0·07	0·52	0·83	0·73						
0·45				0·71	0·96	0·80						
0·50	0·16	0·140	0·09									
0·60	0·24	0·176	0·10									
0·70	0·32	0·22 (1)	0·11									
0·80	0·37	0·21	0·09									
0·90	0·46	0·20	0·05									
1·00	0·55											

Table 10

Values of $\partial\bar{\beta}/\partial\varepsilon$ for $\varepsilon = 10$, $\mu = 1$, $\alpha = 0$

$a/\lambda_0 = 0{\cdot}2$		$a/\lambda_0 = 0{\cdot}25$		$a/\lambda_0 = 0{\cdot}275$		$a/\lambda_0 = 0{\cdot}29303$		$a/\lambda_0 = 0{\cdot}3$		$a/\lambda_0 = 0{\cdot}4$	
b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$
0·546	∞	0·32	∞	0·20	∞	0	0	0	0	0	0
0·55	0·46	0·33	0·07 (5)	0·22	0·03	0·25	0·007	0·10	0·001 (6)	0·05	0·000 (2)
0·56	0·4 (2)	0·34	0·06 (5)	0·26	0·02	0·30	0·032	0·20	0·006 (0)	0·10	0·000 (9)
0·57	0·3 (9)	0·36	0·06 (0)	0·28	0·02	0·325	0·08 (0)	0·30	0·053	0·15	0·003 (4)
0·58	0·3 (6)	0·38	0·07 (5)	0·30	0·03	0·35	0·144	0·35	0·12	0·20	0·012
0·59	0·3 (3)	0·39	0·10 (0)	0·32	0·04 (5)	0·36	0·20 (0)	0·40	0·19 (3)	0·25	0·07 (4)
0·60	0·29	0·40	0·15 (0)	0·35	0·08 (5)	0·38	0·219	0·45	0·22 (3)	0·30	0·1 (9)
0·65	0·25	0·42	0·211	0·365	0·12 (5)	0·40	0·228	0·50	0·24 (8)	0·35	0·3 (6)
0·70	0·23 (3)	0·44	0·251	0·38	0·16 (5)	0·45	0·222	0·60	0·19	0·40	0·4 (4)
0·75	0·22 (5)	0·46	0·257	0·40	0·225	0·50	0·206	0·70	0·17	0·45	0·5 (1)
0·80	0·22 (3)	0·48	0·24 (0)	0·42	0·239	0·55	0·197	0·80	0·17 (5)		
0·85	0·21 (5)	0·49	0·22 (0)	0·44	0·237						
0·90	0·19 (3)	0·50	0·19 (1)	0·48	0·223						
0·95	0·18 (3)	0·52	0·19 (0)	0·50	0·218						
1·00	0·17 (6)	0·54	0·20 (6)	0·55	0·205						
		0·55	0·222	0·60	0·195						

Table 11

Values of $\partial\bar{\beta}/\partial\varepsilon$ for $\varepsilon = 15$, $\mu = 1$, $\alpha = 0$

$a/\lambda_0 = 0{\cdot}2$		$a/\lambda_0 = 0{\cdot}25$		$a/\lambda_0 = 0{\cdot}275$		$a/\lambda_0 = 0{\cdot}29303$		$a/\lambda_0 = 0{\cdot}3$		$a/\lambda_0 = 0{\cdot}4$	
b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$	b/a	$\partial\bar{\beta}/\partial\varepsilon$
0·514	$-\ \infty$	0·307	∞	0·187	∞	0	0 & ∞	0	0	0	0
0·50	− 0·07 (0)	0·310	0·09 (5)	0·28	0·03 (0)	0·25	0·03 (5)	0·10	0·000 (4)	0·05	0·000 (1)
0·49	− 0·11 (0)	0·315	0·06 (5)	0·29	0·06 (5)	0·26	0·05 (0)	0·20	0·004 (3)	0·10	0·000 (2)
0·48	− 0·15 (0)	0·320	0·06 (5)	0·30	0·10 (5)	0·27	0·07 (0)	0·25	0·05 (8)	0·15	0·002
0·47	− 0·44 (0)	0·325	0·07 (3)	0·31	0·16 (0)	0·28	0·09 (5)	0·30	0·21	0·20	0·026
0·462	$\pm\ \infty$	0·330	0·09 (0)	0·32	0·19 (5)	0·29	0·13 (0)	0·35	0·32	0·25	0·15
0·47	+ 0·51 (5)	0·335	0·13 (5)	0·325	0·21 (5)	0·295	0·163	0·40	0·23 (3)		
0·48	0·36 (5)	0·340	0·20 (5)	0·33	0·24 (5)	0·305	0·219	0·45	0·15 (3)		
0·49	0·31 (5)	0·345	0·25 (5)	0·34	0·249	0·31	0·231	0·50	0·08 (3)		
0·50	0·28 (0)	0·350	0·27 (5)	0·35	0·236	0·315	0·239				
0·51	0·26 (0)	0·355	0·28 (0)			0·33	0·230				
0·55	0·210	0·365	0·29 (0)			0·34	0·22 (8)				
0·60	0·187	0·375	0·28 (4)			0·35	0·22 (6)				
0·65	0·173	0·380	0·27 (5)								
0·70	0·166	0·390	0·25 (5)								
		0·400	0·22 (5)								

Manuscript received by the Institution on 4th May 1962. (Paper No. 805.)

U.D.C. 681.32

An Instructional Aid for Digital Computer Logic

By

G. J. LINGWOOD†

AND

R. H. GRIGG‡

Presented at a meeting of the Computer Group in London on 16th November 1960.

Summary: A design is presented of a self-contained instructional aid with which the function of basic digital computing elements may be demonstrated, and in which the elements may be combined to demonstrate the more complex circuits which are part of all digital computers. The speed of operation, and the ease with which elements may be interchanged, combine to make the device particularly suitable for instructional purposes, and the apparatus may also be used for computing experiments.

Diodes are used to perform most of the logical functions, with transistors to switch indicating lamps and regenerate digits. Serial operation is used and the machine can be switched to operate at approximately 1000 bits per second, one bit in two seconds, or one bit at a time, i.e. at step-by-step.

1. Introduction

There is an increasing need to extend the understanding of digital computing and data transmission both in industry and the armed forces. The basic principles are not easily demonstrated or made the subject of a practical laboratory experiment if a complete computer is used, not only because of the speed at which such equipment operates, but also because of the lack of visual indication of the flow of data within the machine.

It is quite evident, therefore, that a requirement exists for an instructional aid which not only operates at suitable speeds but also incorporates a display of the flow of information. In addition the device must possess a high degree of flexibility so that the more complex computing circuits can be formed by suitable interconnection of the basic elements in such a way that their behaviour and function can readily be understood and demonstrated.

It was decided that this requirement could best be met at Shrivenham by making it the subject of a design exercise. It should be explained that among the courses held at The Royal Military College of Science there is a two-year course of training given to Technical Staff Officers. During the last four months of instruction, the students are split into small groups, each of these being given a design problem to solve in a specified time. These are known as "Design Exercises" and the solutions are presented to an invited audience towards the end of the course. The digital training aid was therefore given as a design exercise to a syndicate of four officers under the general supervision of a member of the Military instructional staff.

As a result of the work carried out by the syndicate, an engineered prototype was constructed with the aid of the College workshop within the allotted time, and tribute must be paid to the energy and enthusiasm of all concerned.

2. Assessment of the Requirement

When teaching the basic principles of any subject, it is essential that any demonstration apparatus used be as simple as possible, easily explained and hence readily understood. It is of paramount importance that the student should thoroughly grasp the fundamentals, since these are the foundations upon which future knowledge is built.

The subject of digital computing is no exception, and instruction must therefore start with the most basic computing elements, these being the logical gates OR, AND and NOT. For effective demonstration, each element should preferably be mounted on a separate panel, the front being inscribed with the appropriate circuit symbol rather than a circuit diagram, and its operation illustrated by some suitable indication, say a lamp. Interconnection between gates should be possible so that more advanced logical circuits may be developed and when the action of these combinations has been understood, single panels should be available to replace them. For example, an AND and a NOT/OR gate suitably connected may be replaced by a half adder, while a full adder would replace a number of units. Operation of the basic logical elements can easily be demonstrated by the use of steady potentials since no time or

† Formerly Royal Military College of Science, Shrivenham, Wiltshire; now with U.K. Atomic Energy Authority, Engineering Group Headquarters, Risley, Lancashire.

‡ Royal Military College of Science, Engineering Physics Branch, Shrivenham, Wiltshire.

sequence is involved, but synchronizing waveforms are necessary for the more advanced processes. These processes should be capable of operation at very slow speeds and, when necessary, of being stopped and held at any desired stage. Storage elements will be needed to hold numbers upon which arithmetic is to be performed and also to hold the result. A decision must also be made regarding the system of computing to be used, whether serial or parallel.

Thus a system is envisaged whereby the instructor or a student can proceed from the basic elements and build up more and more complex arrangements, the aim being finally to evolve circuits capable of arithmetical operations. In addition it is desirable to have panels available to demonstrate the conversion of information that occurs not only in computing but also in the transmission of data.

This conception immediately poses a number of problems since a considerable number of panels will be required, but only a few will be in use at a time. A convenient storage system is therefore desirable for the unused units, and a rack suitably wired for power is necessary for display. Insertion or removal of a panel should not be difficult, and this demands that each should be of standard size. For convenience in use in the classroom or laboratory, the complete apparatus should be self-contained and readily transportable.

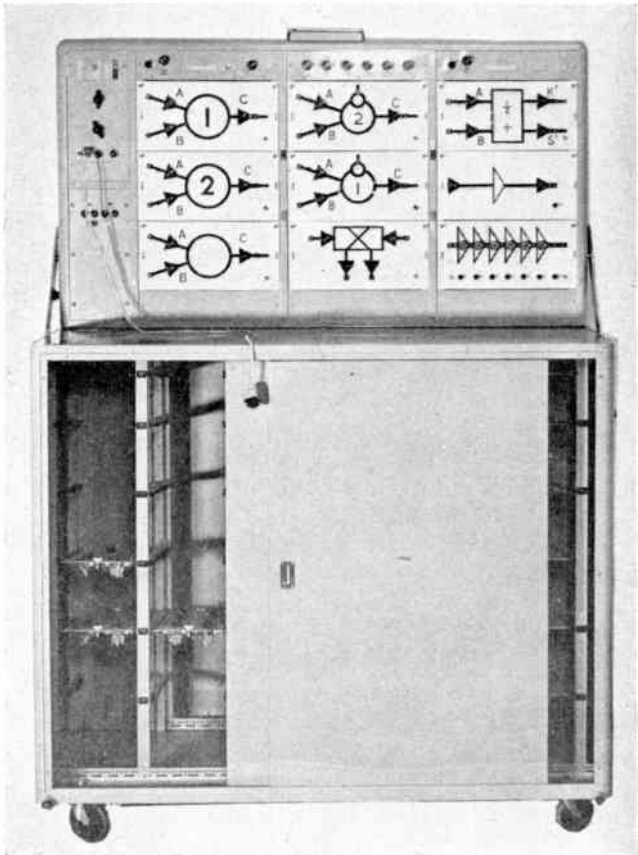


Fig. 1. Complete equipment with rack erected.

3. General Design

Although the design could equally well have been based on the use of thermionic valves or electromagnetic relays, it was decided that semiconductors would be used throughout. Serial operation was adopted since this gives simpler circuits, which are easier to construct, and fewer elements are needed.

Investigation showed that in order to provide for demonstration or laboratory work, a maximum of nine standard panels would be required at any one time and a basic size of 14 in. × 7 in. was considered to be the optimum for each to meet the requirements of good visibility and ease of connection. A considerable amount of thought was also given to the colour scheme required for the front panels and after tests, black circuit symbols on a yellow background were finally chosen as the best.

Consideration was then given to decide the minimum number and type of elements required, the following being selected for construction.

	Quantity
(*a*) Elements with d.c. inputs	
AND gate	2
Inclusive OR gate	2
Exclusive OR gate	2
NOT/AND gate	1
NOT/OR gate	1
Half adder	2
(*b*) Elements requiring a pulse drive	
Full adder/subtractor	1
Trigger pair	1
Shifting register	2
Shifting register/binary counter	1
Single digit delay	1
Stop pulse gate	1
Clock pulse gate	1
(*c*) Elements for data transmission	
Staticizer	1
Dynamicizer	1
Analogue-to-digital converter	1
Digital-to-analogue converter	1
(*d*) Auxiliary circuits required to make the apparatus electrically self-contained.	1
Clock pulse generator	1
P-pulse and sequence generator	
Stabilized d.c. power supply	1

The overall dimensions of the storage cabinet were determined not only by the requirement that it should pass easily through a standard doorway, (3 ft wide by

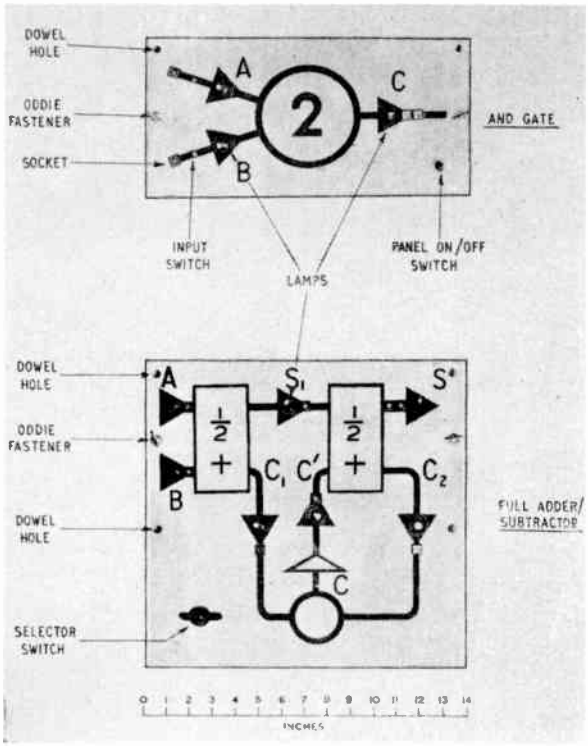


Fig. 2. Typical demonstration panels.

7 ft high at Shrivenham) but also by the fact that it should form the support for the display rack. When not in use, the latter can be folded down in front of the cabinet. The overall height of the complete equipment with the rack erected is 6 ft 5 in., while the width when folded is 2 ft 8 in. The cabinet is at a convenient height to allow all panels mounted in the rack to be easily seen by seated students (Fig. 1).

4. Mechanical Design

To combine strength with lightness, the main framework of the apparatus is made of Widney-Dorlec standard components and the enclosing panels are of light alloy.

The storage cabinet is mounted on castors to permit easy movement of the whole unit and storage space is provided for thirty demonstration panels, spare parts and tools. Removable jacks are available to prevent movement when the equipment is in use, and to conserve space, sliding doors are fitted. Its overall dimensions are 4 ft 5 in. wide, 3 ft to 6½ in. high and 2 ft deep.

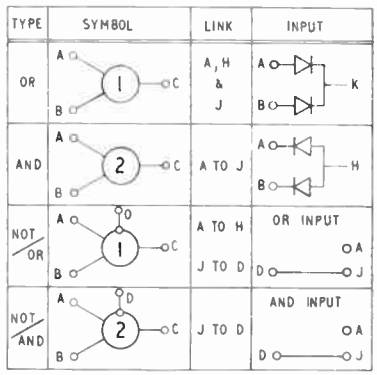
The rack contains all the auxiliary circuits required to make the apparatus self-contained except for an a.c. mains supply. The control and indicating lights of these circuits are displayed on the control panel, and also along the top of the frame, and two 5 A a.c. power outlets are provided for the connection of test instruments, e.g. an oscilloscope. Each demonstration panel may be mounted at any of the nine positions available on the rack and automatic connection of d.c. power and clock pulse drive is then made. The units are secured to the framework by means of Oddie fasteners, and are therefore easily detached after use. Paralleling sockets are provided on the centre support struts of the rack to facilitate wiring of the more complex circuits and to avoid the need for long interconnecting leads, which may obscure the view of the panels.

Figure 2 shows examples of the demonstration panels which are constructed of aluminium alloy reinforced by right-angle stiffeners to prevent flexing. The electronic circuits are mounted on a sub-chassis attached to the back, while the front is inscribed with the circuit symbol. Lights, switches and connecting sockets are merged into the diagram as much as possible, to ensure that the student is not confused by detail. Certain elements, the full adder for example, were found to need a larger panel than the standard so these were made 14 in. square.

5. Electrical Design

To keep abreast with modern techniques, semiconductors were selected for the project and it was intended that printed circuitry should be used wherever possible. Most logical functions are performed by diodes, which were readily available, and transistors are used to switch lamps and regenerate digits; transistors could have been used throughout at slightly greater cost. There is no requirement for high-speed working, hence the switching speed of an audio frequency transistor is more than adequate, and it can switch a 6 V, 0·35 W lamp which is satisfactory as an indicator.

In a binary system, only two states are possible, either "0" or "1", but it is essential for clear recognition that the condition representing one or the other is properly defined. The convention adopted for this machine is that a "1" is indicated by lighting a lamp. The indicator is connected in series with the collector of a transistor, which is used as a switch. In the cut-off condition, the collector current is negligible, the lamp is off and a "0" is indicated. When the transistor is turned on, the collector current increases rapidly to a maximum limited by the load resistance, the lamp lights and indicates a "1". A series resistor is included to reduce the effect of lamp resistance variation. Hence the output voltage for a "1" is very low, about zero volts, and is maintained as long as the transistor is bottomed. In the case of a "0", the transistor is cut off and the nominal value is −8 V, but the actual voltage will depend on the number of circuits connected to the output. All circuits are therefore designed to accept a range of "0" conditions.



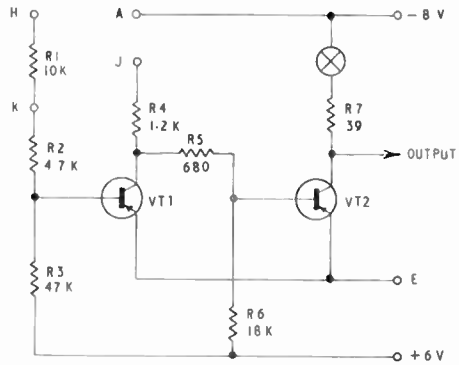


Fig. 3. Basic gate circuit.

Two-stage direct-coupled amplifiers are needed to maintain the correct output polarity.

As the action of the basic logical elements are similar, a single circuit was designed so that by re-arrangement of simple links and diode connections an OR, AND, NOT/OR or NOT/AND gate may be obtained. This is shown in Fig. 3 and the arrangement required for a particular gate can be established by reference to the table. Since, in many cases, at least two of these are required to form one element, the sub-chassis consists of a standard size paxolin panel drilled and fitted with turret tags to accommodate two gates. Although the operation of the various gates can be demonstrated easily by the application of suitable d.c. potentials, the input circuits are fitted with three-position toggle switches, thus allowing for the application of an external pulse signal. The operation of each type of gate will now be examined in detail.

5.1. *D.C. Input Elements*

5.1.1. Inclusive OR gate

The circuit shown in Fig. 4 is achieved by linking points A, H, and J, the inputs being applied at K from the cathodes of a pair of diodes. If both inputs are at "0" (−8 V), the diodes are back biased and due to base current through the network R1, R2, VT1 bottoms driving the collector and hence the base of VT2 positive. Thus VT2 is cut off, the lamp is off and a "0" is indicated. When either or both inputs are "1" (about zero volts), either or both diodes conduct driving the base VT1 positive. Hence VT1 cuts off, VT2 bottoms, the increased collector current lighting the lamp to indicate a "1". Thus this gate may be used to detect the presence of a signal on either input in this case, or in general on any one of a number of inputs, which can be provided by including an additional diode in the circuit for each source.

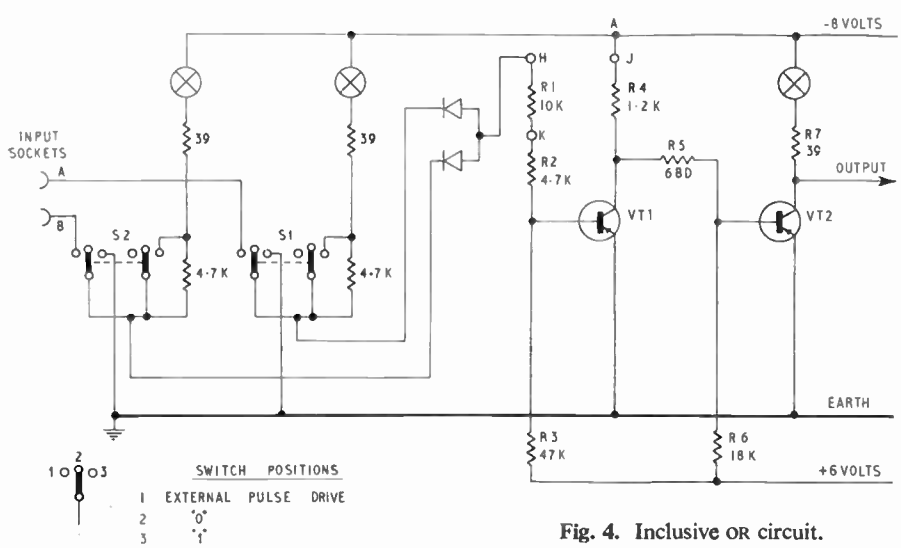


Fig. 4. Inclusive OR circuit.

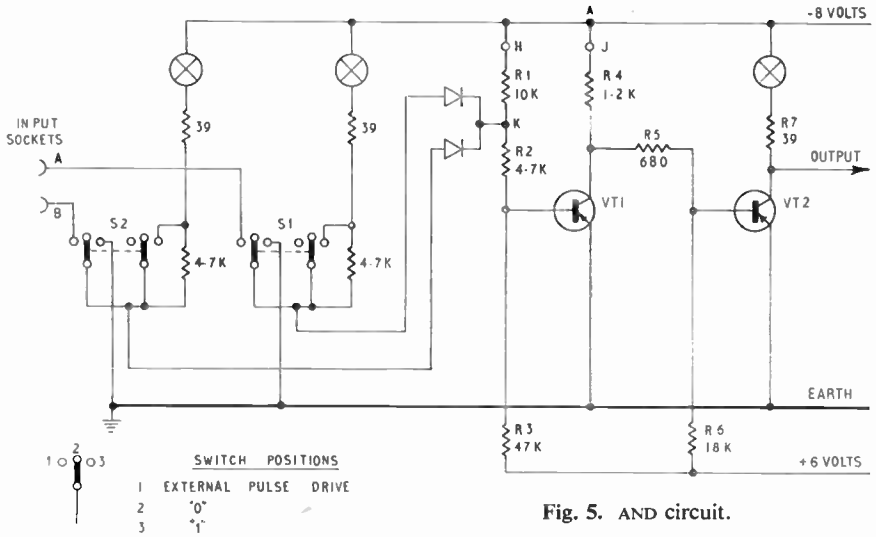


Fig. 5. AND circuit.

5.1.2. AND gate

Points A and J are connected to form this circuit and the signals are applied at H from the anodes of a pair of diodes as shown in Fig. 5. If either input is "0" (−8 V), one of the diodes is conducting and base current is supplied to VT1 through the resistors R1 and R2. Hence VT1 is on, VT2 off and the output is "0". When both inputs are "1" (about zero volts), the potential network R1, R2 and R3 cuts off the base current to VT1, whose collector current falls allowing VT2 to bottom thus giving an output of "1" and lighting the indicator. This gate can be used therefore to detect coincidence between the two input signals.

5.1.3. NOT/OR and NOT/AND gates

In certain circumstances, it is required to prevent the normal output being available from a gate by the introduction of a further input known as an "inhibit" signal. This facility when applied to the elements already described gives a NOT/OR or a NOT/AND gate. Referring to Fig. 3, the link between points A and J is removed and the latter is connected to an input socket on the front panel to facilitate the application of an inhibit signal. When the input is a "0" (−8 V), quite obviously the gate functions normally since this simulates the reconnection of A and J, but when a "1" is applied, (about zero volts) VT1 is inoperative and VT2 is held at cut-off hence no output can be obtained.

The logical elements described form the basic building blocks of the system, and it is now a matter of logical design to establish the particular configurations which are needed. It is essential that the machine should be able to add two binary numbers and it is proposed to examine in some detail precisely how this is achieved.

Operation in the serial mode infers that the digits appear in time sequence, hence the process of addition is carried out just as one would add up two numbers on paper, starting with the least significant digits and dealing with pairs of digits at successive time intervals. If a pair of digits is represented by A and B respectively and their sum by S, all the possible values for A, B and S can be written down in tabular form as shown in Fig. 6, which is known as a "Truth Table". An examination of this reveals that an inclusive OR gate can fulfil the first three conditions postulated, but not the last. This obviously requires an AND gate to detect coincidence and, when this occurs, the output from the OR gate must be zero. If the output of the AND gate is used to inhibit the NOT/OR as shown in the logical diagram, Fig. 6, the desired result is achieved. This combination is known as an "exclusive OR" gate.

The summation of two digits in any scale can produce a "carry", which must be taken into account for the process to be accurate. Further examination of the Truth Table shows that the carry is available when coincidence occurs and can be taken from the

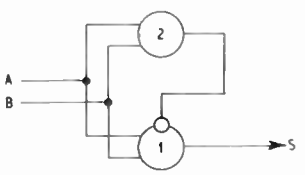
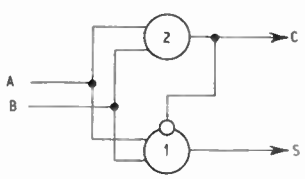


TABLE 1

A	B	S
0	0	0
1	0	1
0	1	1
1	1	0

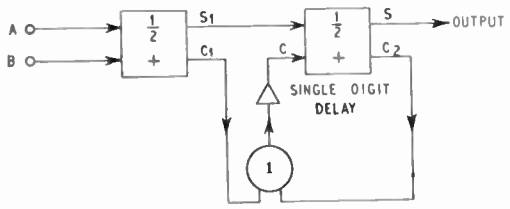
Fig. 6. Exclusive OR logical diagram.



TRUTH TABLE

A	B	S	C
0	0	0	0
1	0	1	0
0	1	1	0
1	1	0	1

Fig. 7. Half adder logical diagram.



A	B	C	C_1	S_1	C_2	S
0	0	0	0	0	0	0
1	0	0	0	1	0	1
0	1	0	0	1	0	1
0	0	1	0	0	0	1
1	1	0	1	0	0	0
1	0	1	0	1	1	0
0	1	1	0	1	1	0
1	1	1	1	0	0	1

Fig. 8. Basic full adder.

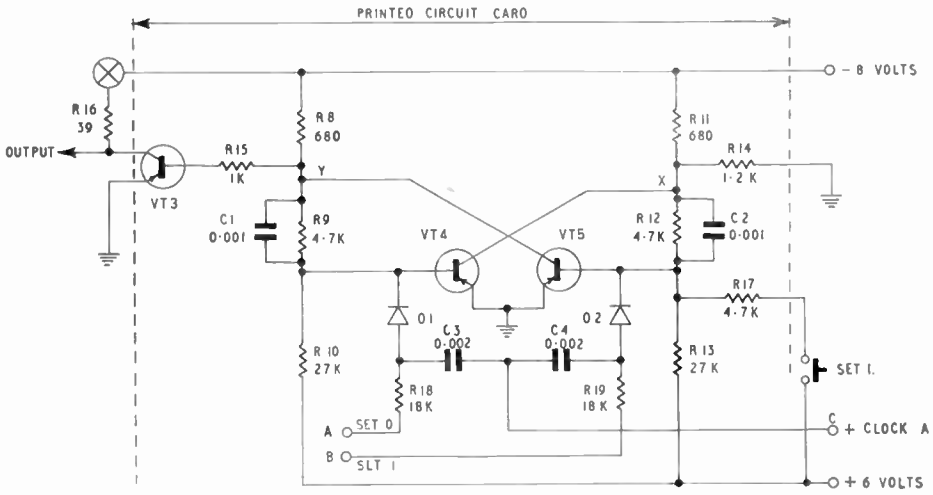


Fig. 9. Bistable circuit.

output of the AND gate as shown in the logical diagram Fig. 7. The combination is then called a "half adder" because it can deal with the summation of two digits, but not two numbers.

5.2. *Pulse Driven Elements*

5.2.1. Basic full adder

The serial addition of two binary numbers, as opposed to two binary digits, introduces a complication since any carry produced must be delayed in time by one digit period, so that it may be taken into account with the next significant pair of digits to arrive. Thus essentially a three-input device is required, the delayed carry constituting the third input. The logical diagram, Fig. 8, shows a possible solution to the problem, and it can be seen that the arrangement consists of two half adders, an inclusive OR gate and a single digit delay, hence four standard panels are required for the circuit.

The operation of the half adder has been described and needs no further comment, but it is of interest to examine the production of third input. The carry derived from either half adder is recirculated through the OR gate, and stored by the single digit delay for one digit period, thus appearing at the input to the second half adder in time sequence with the sum of the next significant pair of digits. The Truth Table given in Fig. 8 illustrates the action of the circuit, and also shows that it is impossible for both half adders to produce a carry signal simultaneously.

5.2.2. Storage

A convenient system of storage is essential for any digital computer, and the designer must consider not only the type of store, but also its capacity. For this instructional aid, it was decided that an Eccles-Jordan bistable circuit should form the basic storage element and that six digits would be sufficient to provide adequate demonstrations of the techniques involved in digital computing.

Referring to Fig. 9 it can be seen that the indicator and output is separated from the bistable circuit and a resistor (R14) is included to balance the load on the collectors of the transistors VT4 and VT5. If the former is conducting, the point X is almost at earth potential, the resistors R12 R13 maintaining a positive voltage on the base of VT5, which is therefore cut-off. The point Y is then at about -4 V, the resistors R9 R10 holding the base of VT4 negative and hence the state is stable. In this condition, R8 and R15 supply base current to VT3, which conducts lighting the lamp and indicating a "1". When VT3 and VT4 are turned off, VT5 conducts, the condition again being stable and representing a "0", since the indicator lamp is now off. The capacitors C1 and C2 are provided to improve the transient switching conditions.

Two stable states are therefore possible, and a change can be initiated by triggering the conducting transistor towards cut-off by the application of a positive-going pulse to its base. Since either VT4 or VT5 may be conducting, the trigger pulse must be properly routed, and this is achieved by the use of a well known steering circuit.[1] In Fig. 9, the components R18 R19, C3, C4 with the two diodes D1 and D2 form this circuit, which consists of two AND gates with one input terminal common. The junction of C3 and C4 is the common input point C and is supplied with positive going pulses from the clock, while the other inputs must be complementary quantities, i.e. if A is 0 then B is 1 and vice versa.

Consider the case when VT4 conducts, VT5 is cut-off and a 1 is indicated. The d.c. potentials on the bases of the bistable pair of transistors only vary about earth by a small amount, the cathode of D1 being slightly negative, whereas that of D2 is somewhat positive. Thus the bias condition of the diodes will depend largely on the complementary input signals applied. If A is 0 and B is 1, coincidence will only occur in the steering circuit AND gate which includes D2, and hence a positive trigger pulse is applied to the base of VT5, but since this transistor is already cut off, no change takes place. However, if A is 1 and B is 0, the diode D1 will conduct allowing the trigger pulse to be transferred to the base of VT4, which is driven towards cut-off, thus VT5 is turned on and the bistable circuit changes state to indicate a 0. It can readily be seen that the time interval for a 1 to be passed through the bistable circuit must of necessity be equal to one digit period, since two consecutive positive clock pulses must be applied. The rate at which alternate 1 and 0 conditions can be transferred is limited not only by the switching speed of the transistors but also the time constant of the components R18 and C3.

A standard printed circuit card was available and was used to accommodate the complete circuit including the lamp driving transistor and the steering circuit. In addition, a resistor R17 is connected from the base of VT5 to the positive supply through a manually-operated pushbutton switch mounted on the front panel. This allows a 1 to be set in by operation of the switch.

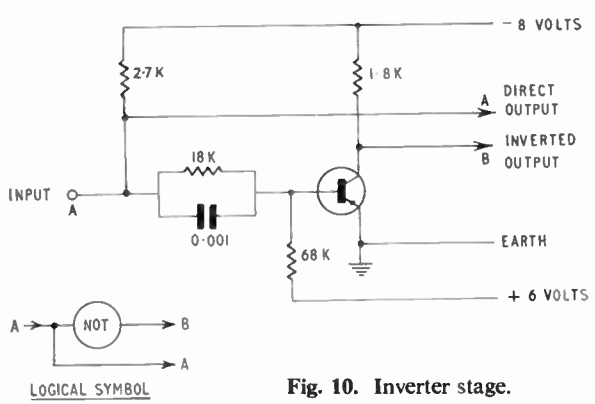


Fig. 10. Inverter stage.

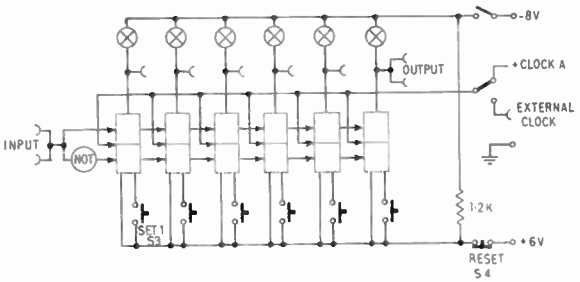


Fig. 11. Block diagram 6 stage shift register.

The action of the bistable circuit can easily be demonstrated by providing indicating lamps on both sides, and including two manually-operated push-button switches to initiate a change of state in either direction. The switches are wired from the positive supply through suitable resistors to the bases of VT4 and VT5. The components are fitted to a standard printed circuit card, which fits into an 18-point socket mounted on the chassis. The card may be used for a variety of purposes.

Since serial operation is employed, only single point connectors are necessary to link the various elements in any configuration, but the steering circuit described requires dual complementary inputs. In order to maintain compatibility, a device is needed which accepts a single input and provides direct and inverted outputs. These requirements are met by the simple circuit given in Fig. 10.

5.2.3. Single-digit delay

The unit contains a bistable element on a standard printed circuit card (see Fig. 9), preceded by an input circuit as shown in Fig. 10. The time interval for a 1 to be passed through the bistable circuit is that between two consecutive clock pulses, i.e. one digit period and hence the device may be used to obtain unit delay. A push-button switch is provided on the front panel for clearing purposes.

5.2.4. Six stage shift register

Figure 11 shows the block diagram for a six-stage shift register, three of which are used as logical stores for the computer. Six standard printed circuit cards are used, the collectors of one stage being coupled to the bases of the succeeding stage by means of the steering circuits. A positive-going clock pulse is applied simultaneously to all the interstage steering circuits, the state of each stage then being transferred to the succeeding stage, so that the stored number shifts to the right by one digit. Clearly, therefore, six successive trigger pulses are required to shift a number out of the register. A number can be placed in

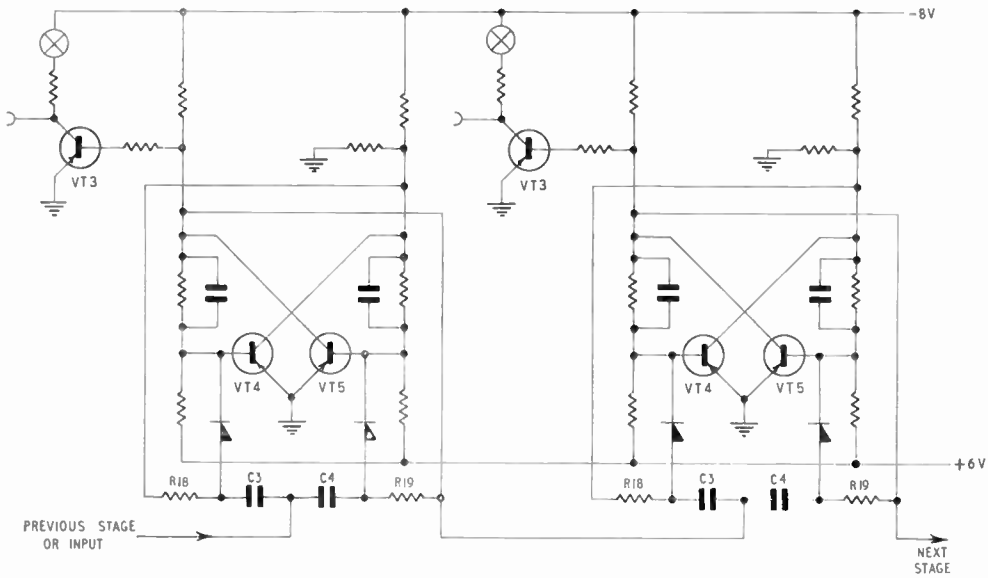


Fig. 12. Two stages of binary counter.

the store by manual operation of the appropriate push-button switches mounted on the front panel, and the store can be cleared by operation of the re-set button provided. A changeover switch allows the application of an external trigger pulse in place of the positive going clock pulses normally supplied. The output of each stage, which is derived from the collector of the lamp driving transistor, is wired to a single socket on the front panel.

5.2.5. Binary counter

Figure 12 shows the circuit diagram for two stages of a binary counter, and although standard printed circuit cards are used, the triggering arrangements are modified. The dual complementary inputs to the steering circuits are now derived from the collectors of the bistable pair of transistors, VT4 and VT5, and the input stimulus applied to the common junction of the capacitors C3 and C4. Two positive input pulses are required for each output pulse, so that each stage operates as a "divide-by-two" circuit.

Since the binary counter is basically similar to the six-stage shift register, it was found possible to combine both functions in one standard panel, the changeover being achieved by means of a multipole switch. This shift register/binary counter is therefore normally used to display the result of arithmetical processes.

5.2.6. Stop pulse gate

This element is formed by cascading two NOT/OR gates as shown in Fig. 13. The first gate is provided with additional diode inputs and may be operated as an OR gate with the inhibit switch in the position shown, the output being used to inhibit the second gate. Hence any signal applied to the latter will pass through this gate only when all the inputs to the system are at "0", and this output pulse is used to stop the machine.

When the inhibit facility is used on the first gate, the operation is similar, but production of an output stop pulse will now depend upon the time at which the inhibit pulse is applied to the input element. When this occurs, the inhibit signal is again removed from the second stage and an output stop pulse produced.

The standard paxolin sub-chassis is used in conjunction with the basic gate circuit, suitably modified to accept additional inputs, and the inhibit switch is located on the front panel.

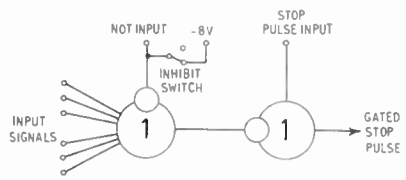


Fig. 13. Stop pulse gate.

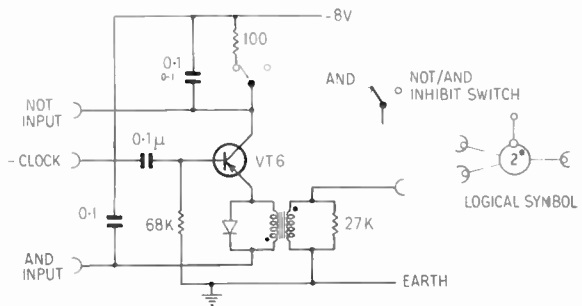


Fig. 14. Clock pulse gate.

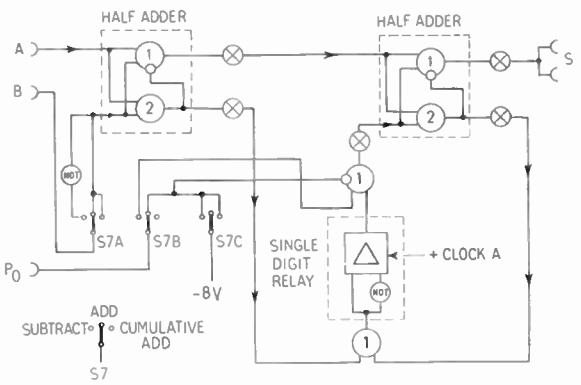


Fig. 15. Full adder/subtractor logical diagram.

5.2.7. Clock pulse gate

Referring to Fig. 14, the inhibit facility is again optional and the element is fundamentally an AND gate, but the logic in this case is performed by a transistor.

With the inhibit switch in the position shown, the device operates as an AND gate. Negative-going clock pulses are fed to the base of the transistor, which remains cut off until a "1" is applied to the AND input. This simulates an earth connection to the emitter circuit, and the transistor conducts during coincidence to produce an output pulse. The diode is provided to limit the overswing on the primary of the transformer, and a positive going output is selected by appropriate grounding of the secondary winding.

With the switch set to NOT/AND, normal gating takes place until an inhibit signal is applied. Since the element differs somewhat from the diode logic AND gate, it was decided that this should be indicated on the logical symbol displayed on the front panel by the inclusion of an asterisk as shown.

5.2.8. Full adder/subtractor

The full logical diagram is given in Fig. 15 and with the switch in the position shown, the system is set to add. The arrangement is similar to the basic full adder already described, except that automatic cancellation of any carry left from a previous calculation is provided. This is achieved by a pulse (Po) applied as an inhibit signal at the beginning of a sequence, to the NOT/OR gate coupling the single digit delay to the delayed carry input of the second half adder. It should be noted that the Po pulse occurs at a time corresponding to the least significant digit, and hence its numerical value is one.

With the system switch set at "cumulative add", the inhibit pulse (Po) is not injected and cancellation does not take place. This facility was provided for adding numbers not restricted to six digits.

Subtraction is carried out by adding the complement of the number being subtracted. In binary notation, the complement is easily formed by passing the number through an inverter stage and adding "1". Since the pulse Po is already available in the adder, and is numerically equal to one, it can conveniently be used to add in the extra one. When the switch is set to "subtract", the number B is inverted, the inverter stage being as shown in Fig. 15, and the pulse Po is connected through the OR gate to the input of the second half adder.

5.3. *Conversion Units*

5.3.1. Staticizer

The staticizer is used to demonstrate conversion from serial to parallel representation. A double size front panel, 14 in. square, was necessary to display the logical diagram, Fig. 16, and this shows that each stage consists of an AND gate coupled to a bistable element. Although standard printed circuit cards are used for the latter, the steering circuits are modified as shown in Fig. 17. This allows the application of separate positive going pulses to set "0" or "1", and at the beginning of a sequence, Po gates the positive clock B pulse to reset all the bistable elements to "0". The serial pulse train is applied to all the AND gates, and the second input stimulus to each is a P pulse. Any digit in the serial number, which is a "1" will pass through a gate when coincidence occurs with the appropriate P pulse, and hence the associated bistable element is set to "1" by the positive going clock A pulse. At the end of a sequence therefore, the parallel representation of the number is available.

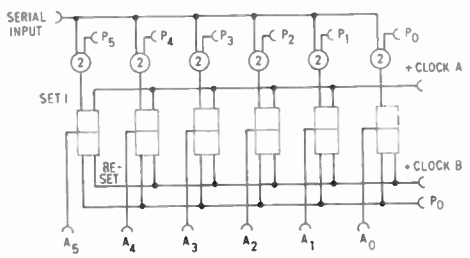


Fig. 16. Staticizer.

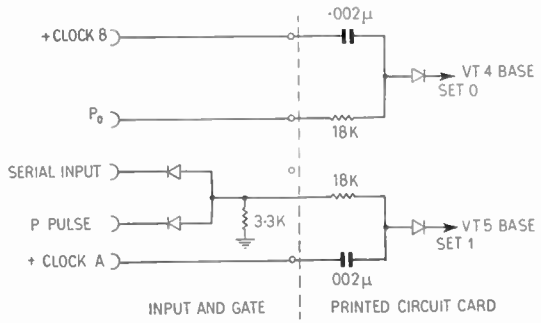


Fig. 17. Modified steering circuit for staticizer.

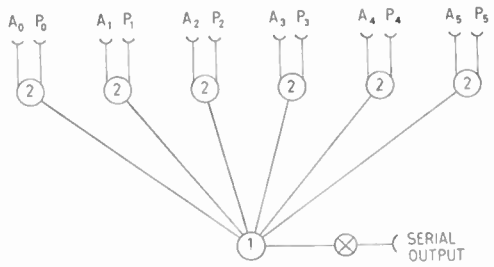


Fig. 18. Dynamicizer—logical diagram.

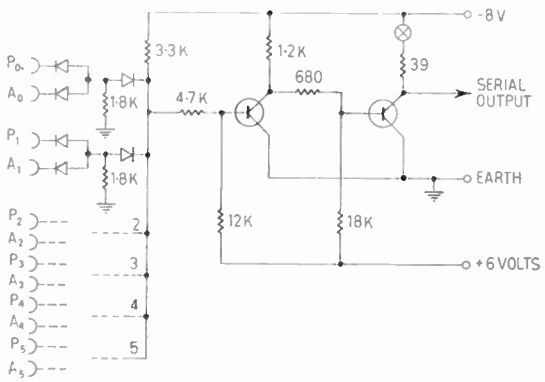


Fig. 19. Dynamicizer—circuit.

5.3.2. Dynamicizer

This unit performs parallel to serial conversion, and the logical diagram, Fig. 18 shows the input parallel number and P pulses (which appear in time sequence) applied to six AND gates, each feeding a common OR gate. If any digit of the input parallel number is a "1", coincidence occurs with the appropriate P pulse and the digits will pass through the OR gate to appear at the output in serial form with the least significant digit first. If the output is coupled to a shift register, the serial representation of the number can be displayed at the end of the sequence.

If the dynamicizer had been constructed using the basic gate circuit, fourteen transistors would have been required, and in order to economize, the basic arrangement was modified, the circuit then being as shown in Fig. 19. By combining the diode circuits, a larger input current was needed, but this was considered acceptable since only two transistors were necessary.

5.3.3. Analogue-to-digital converter

At the time of construction of the apparatus, a G.E.C. Induction Digitizer Type LB5–3 was readily available at Shrivenham, and it was decided that this should be used. The digitizer is essentially a transformer, in which the coupling between the primary winding and five secondary windings is such that the output voltages in the latter are either in or out of phase with the primary voltage, depending on the position of the rotor. The secondary voltages change phase rapidly with movement of the rotor, and remain fairly constant in amplitude between changes. The flux paths between the motor and stator are such that 32 shaft positions are defined in the progressive binary code by the secondary winding voltages which are in phase with the primary. The device can be used with the primary winding energized, when the binary output is available in parallel form from the five secondary windings. Alternatively, a serial output can be taken from the primary if the secondary windings are driven in sequence, and this also gives a considerable economy in the associated circuits.

Referring to the circuit diagram, Fig. 20, negative clock B pulses are passed through a transistor amplifier,

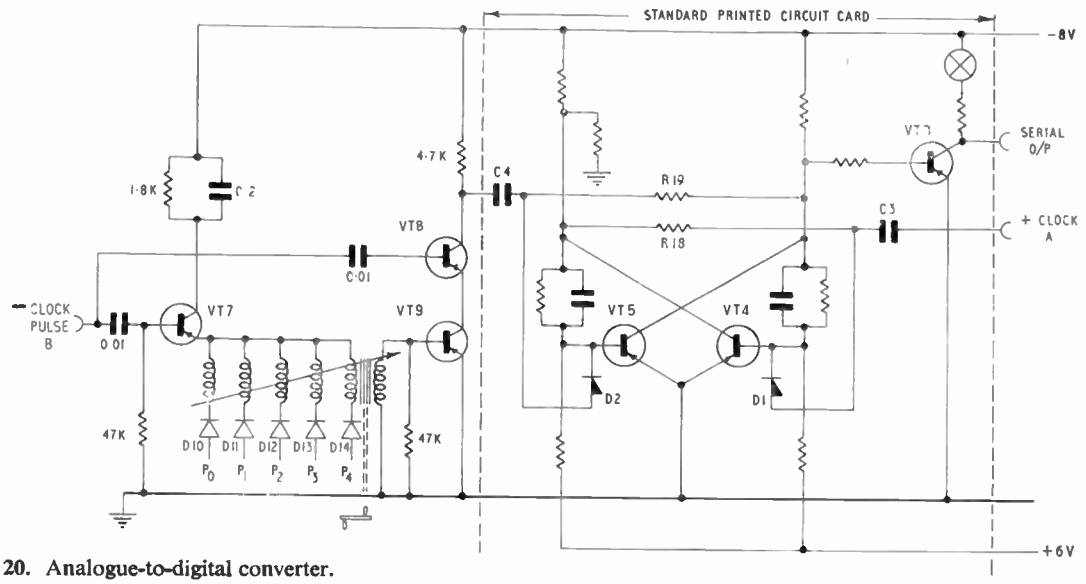


Fig. 20. Analogue-to-digital converter.

the output being connected to all the digitizer secondary windings. The amplified clock current pulse is routed through these windings in a time sequence by using P pulses to forward bias the junction diodes D10 to D14. A negative-going output pulse is obtained from the primary winding when a particular digit is present for a given shaft position, and absence of a digit gives rise to a positive-going pulse, which is followed by a negative-going overshoot. The latter is prevented from giving a false indication by the use of time and phase discrimination to select the correct pulse. Two transistors VT8 and VT9 form a series AND gate, the former being turned on when the negative clock B pulse is applied to its base, and the latter conducting when the digitizer primary winding supplies a negative going voltage. Hence when coincidence occurs, both transistors conduct, and the output pulse sets a "1" into the bistable stage. On the next clock A pulse, the stage is reset to "0". In this way, the binary number defining a particular shaft position will be displayed by the indicator lamp of the bistable stage, digit by digit if the machine is operated at "step by step". Alternatively if the output is connected to a shift register, the whole binary number will be indicated at the end of the sequence. It should be noted that the position of the shaft is represented by a maximum of five digits, this limitation being imposed by the digitizer, and hence only five P pulses are necessary for sequential switching of the diode gates. Any number in the range may be selected by means of a suitably engraved scale mounted on the front panel, and a pointer type knob fitted to the digitizer shaft.

5.3.4. Digital-to-analogue converter

The most convenient form of analogue output is a voltage, which may be displayed on a meter or on a cathode-ray oscilloscope. The circuit diagram, (Fig. 21) shows that a six-digit parallel drive is required, and briefly the principle of operation is that each input "1" produces a current proportional to its binary significance. The analogue voltage can then be formed by summing these currents, and passing the total through a resistor.

Referring to the diagram, diode gates are used in the input circuits, in order to prevent any significant drive current being applied when a "0" is present. This is achieved by holding the base of VT11 accurately to the reference voltage of the Zener diode, by the action of the feedback stabilization circuit comprising VT10, VT11, and VT12. Since the reference is $-3{\cdot}5$ V and a "0" is represented by -8 V, the appropriate diode is cut-off when any input is "0", and the reverse current that flows is insignificant compared to I_0.

The diode gates are enabled when a "1" is applied, and each input current is then defined by a constant voltage and a resistor which is chosen to provide the correct magnitude of current. However, a "1" is represented by a bottomed transistor voltage of approximately $-0{\cdot}2$ V, the exact value depending on the particular transistor and the pre-set resistors are included in the more significant inputs to allow for this possible variation. In this context, it should be mentioned that the six-digit parallel input required can be derived from several alternative sources, e.g. any of the shift registers or the staticizer, and hence for accurate results, the converter input circuits should be adjusted in conjunction with a particular driving source.

The emitter current of VT10 consists of the sum of the drive currents and a small component due to base current of VT11. In the collector circuit of VT10, the output impedance is very high compared with the load resistor, and hence the latter will not affect the collector current, which is approximately 0·98 of the total emitter current flowing. The quantized analogue voltage can therefore be taken from the collector of VT10, the zero error due to the base current of VT11 being eliminated by adjustment of the preset resistor.

5.4. *Auxiliary Circuits*

These are housed in the display rack, and consist of the control panel and the power supply unit. The former allows selection of the various modes of

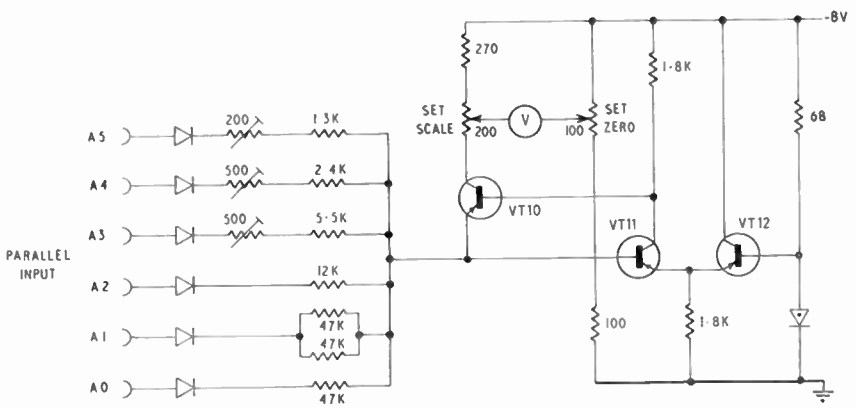


Fig. 21. Digital-to-analogue converter.

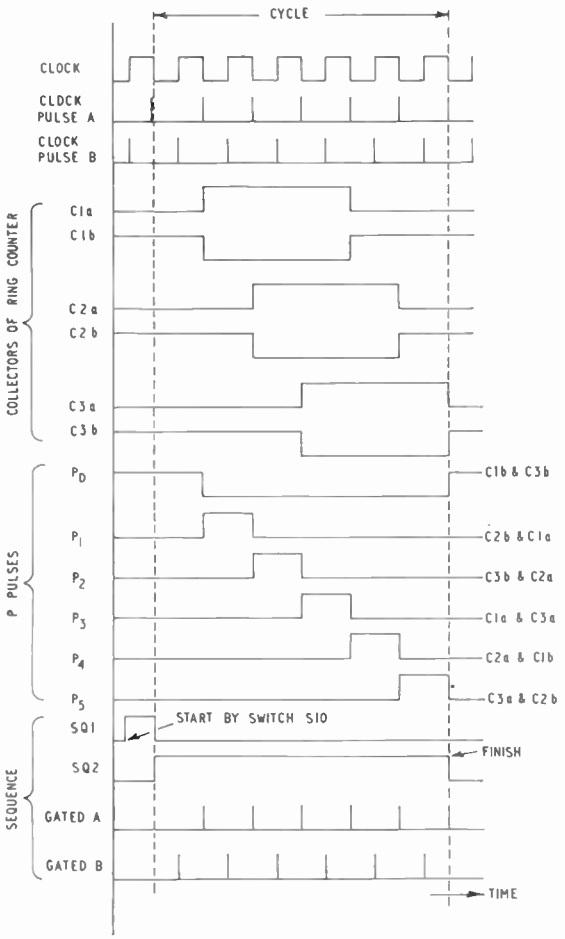


Fig. 22. Basic cycle waveforms.

operation, and also produces the synchronizing waveforms necessary for not only the processes of arithmetic, but also the conversion of data from one form to another, and it is proposed to discuss these in general terms first.

The clock multivibrator produces a square wave output, from which two phases of clock pulse are derived alternately. Clock A is a narrow pulse, which is used to initiate the movement of digits within the machine, while clock B is a similar pulse, appearing midway between the clock A pulses, and determines the synchronism of those events which must occur when the digits are stationary. In addition, the interval between clock A pulses is filled with square waveforms, which may be used to examine the significance of a binary digit. These waveforms are known as P pulses and are numbered from P_0 to P_5, the former being coincident in time with the least significant digit, whilst the latter is synchronized with the digit representing 2^5. Figure 22 illustrates the basic cycle, which consists of six A and B clock pulses and all the P pulses, the sequence commencing with P_0 immediately after a clock A pulse.

Referring to the block diagram, (Fig. 23) the clock multivibrator runs continuously providing two anti-phase outputs to drive the clock A and B generators, which are connected to two AND gates, output pulses then being produced only when coincidence occurs with the sequence gate pulse, Sq2. The sequence circuit consists of two cascaded bistable elements, Sq1 and Sq2, the clock A pulses being applied through one pole of the push button "start" switch S10, which is connected to the control panel by a short length of flexible cable. When the "start" button is depressed,

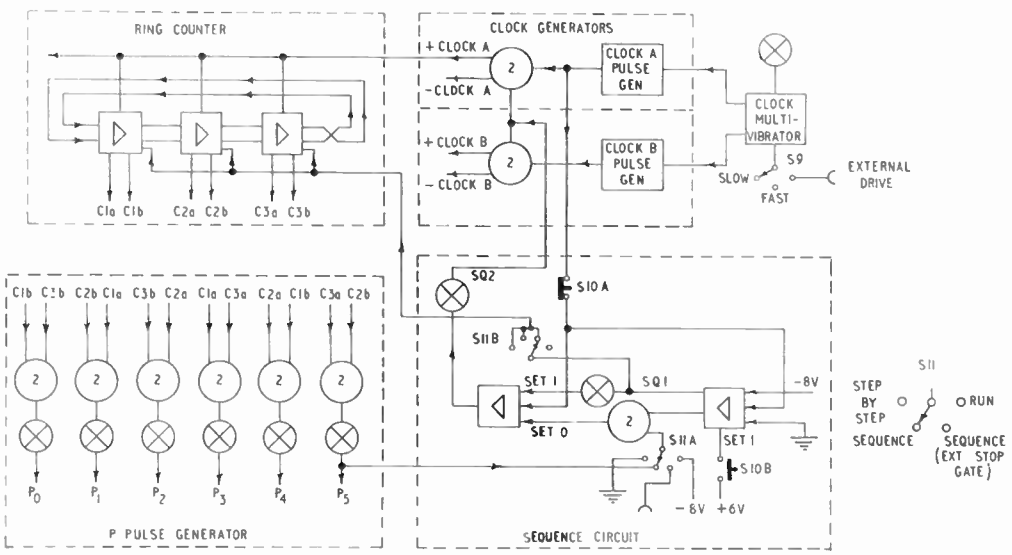


Fig. 23. Block diagram of clock circuits.

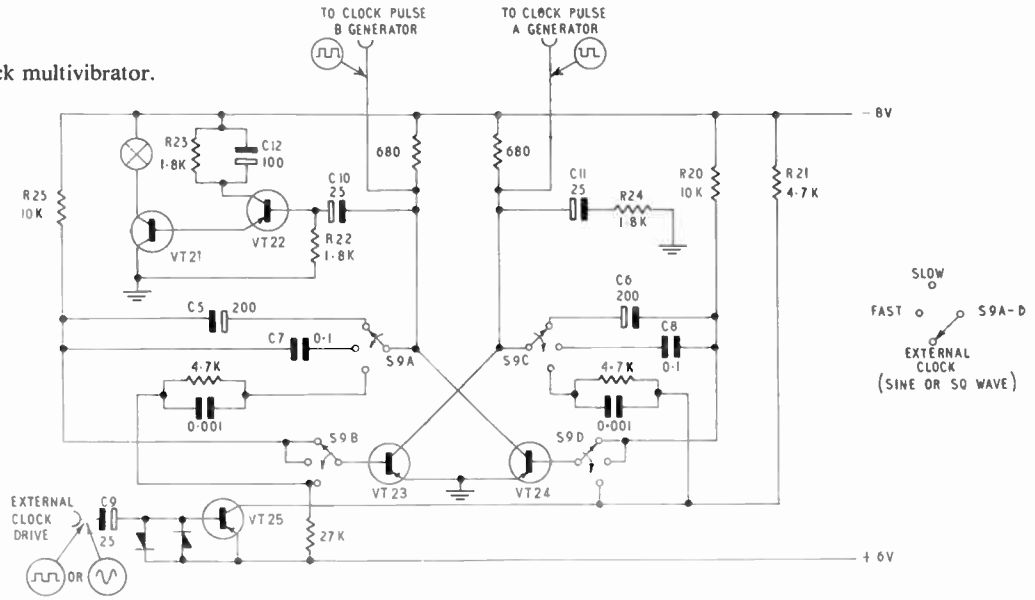


Fig. 24. Clock multivibrator.

S10A interrupts the clock drive while S10B set a "1" into Sq1, and when the push button is released, the clock drive is restored and the "1" passed into Sq2 on the next clock A pulse. This enables the two clock AND gates to allow clock pulses to pass to the ring counter and other circuits. Although Sq1 reverts to the "0" condition, the diode AND gate prevents this passing to Sq2, which remains set to "1" until the end of the cycle, when P_5 opens the AND gate to permit the next ungated clock A pulse to set Sq2 to "0" and so stop the sequence.[2] This is the mode of operation with S11 in the position shown, i.e. set at "sequence", and hence the machine completes one word and stops. The processes of multiplication and division require the apparatus to run for several words, and this is possible with S11 set to "Sequence External Stop Gate" and the stop pulse is then applied from an external source. On "Run" operation, "0" (−8 V) is fed to the diode AND gate and the machine will run continuously, once started, until S11 is returned to "Sequence" or "Step by Step". With the switch in the latter position, an earth ("1") is fed directly to the AND gate and hence when the "Start" push button is operated, Sq2 is set to "1" and then "0" on consecutive clock A pulses and consequently only one B and one A pulse are gated through the output AND gates. Thus the sequence is completed at "Step by step" only after the push button has been operated six times.

The ring counter consists of three shift register stages with the inverted output applied to the input and two possible patterns can circulate—either 010101 or 000111000111, the latter being the required cycle of six. Although when switched on, the stages are more likely to arrive at a state to give the required pattern, Sq1 is used to set the "0" condition into all three stages simultaneously at the beginning of a sequence. This is not possible, however, at step-by-step, because the cycle would not progress beyond P_1, due to the re-setting action of Sq1, and hence the connection is routed via a pole of S11 to ensure correct operation of the system.

The gated clock A pulses synchronize the movement of digits around the ring counter, and the collector voltage waveforms are used in pairs to drive six AND gates, which in turn produce the P pulses. This is clearly illustrated by Fig. 22.

5.4.1. Clock multivibrator

In the circuit diagram, (Fig. 24) the control switch S9 selects "Fast" or "Slow" speed operation, and also permits the application of an external drive signal if required. At "Fast" (about 1000 c/s) and "Slow" (about 0·5 c/s), the transistors VT23 and VT24 are connected as an astable multivibrator, the frequency being determined by the time constant of the components R25, C5 (or C7) and R20, C6 (or C8). At "Slow", it is desirable to have an indicator which flashes at each clock A pulse. The negative-going collector voltage of VT24 is differentiated by C10, R22, amplified in the emitter follower VT22 and direct coupled to the lamp driving transistor VT21. At low frequencies the charge for the current pulse to VT22 is provided by the capacitor C12, and at high frequencies the resistor R23 limits the total current taken by VT22. The components R24 and C11 are included to balance the circuit.

When the switch S9 is set to "External Clock" the transistors VT23 and VT24 are connected as a monostable multivibrator providing a square wave output from the input triggering waveform, which

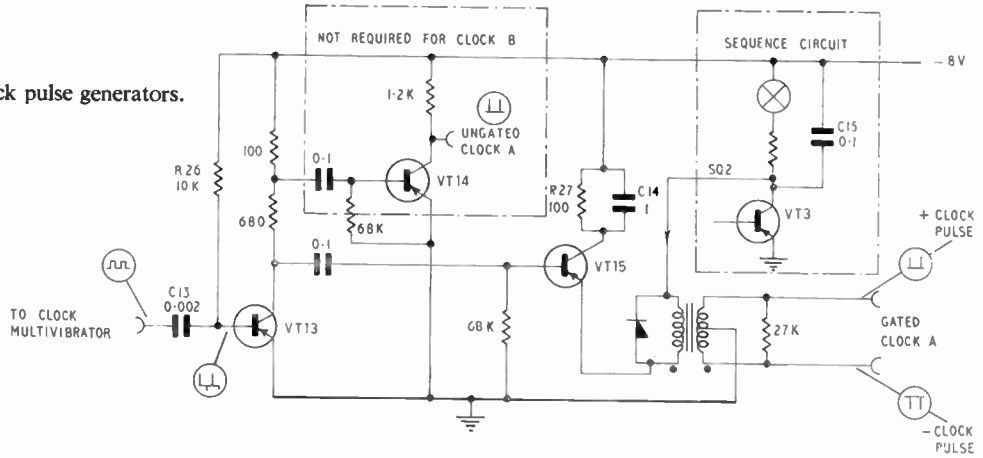


Fig. 25. Clock pulse generators.

can be either a sine or square wave. With no external drive, VT25 draws very little current through R21, allowing a negative potential to appear at the base of VT24, which conducts to drive VT23 to cut-off. When a sine or square wave is applied through the input capacitor C9, the negative half cycles cause VT25 to conduct more heavily, and the increase of current through the resistor R21 drives the base of VT24 to a less negative potential. Hence VT24 turns off and VT23 conducts, the normal cross-coupling producing a quick transition. The input diodes assist in squaring and also ensure that the mark/space ratio is near unity. The circuit works satisfactorily from a few cycles per second to approximately 15 kc/s on an input sine wave of about 1 V, and the turn-on rise times of VT23 and VT24 are always fast enough for differentiation in the clock pulse generators.

5.4.2. Clock pulse generators

Figure 25 shows the circuit diagram for the clock A pulse generator, the arrangement for clock B being similar except for the production of ungated pulses, which are not required.

Transistor VT13 is normally bottomed due to base current through the resistor R26 and the incoming clock square wave is differentiated by the capacitor C13 and the transistor input resistance. A positive and negative spike is produced, but only the former is amplified by VT13, and hence a negative spike of almost the full supply voltage appears at the collector. VT14 provides the positive ungated clock A pulses for the sequence circuit, whose output controls the operation of the emitter follower amplifier VT15. When Sq2 is "0" (−8 V), VT15 cannot conduct due to reverse bias on its base emitter junction. If Sq2 is "1", the transformer is earthed through a bottomed transistor in the sequence circuit and VT15 conducts. The diode is used to limit the primary overswing, and the centre-tapped secondary provides positive and negative gated clock pulses of about 7 V amplitude at very low impedance. The resistor R27 limits the mean current to VT15 and C14 provides the high peak cur-

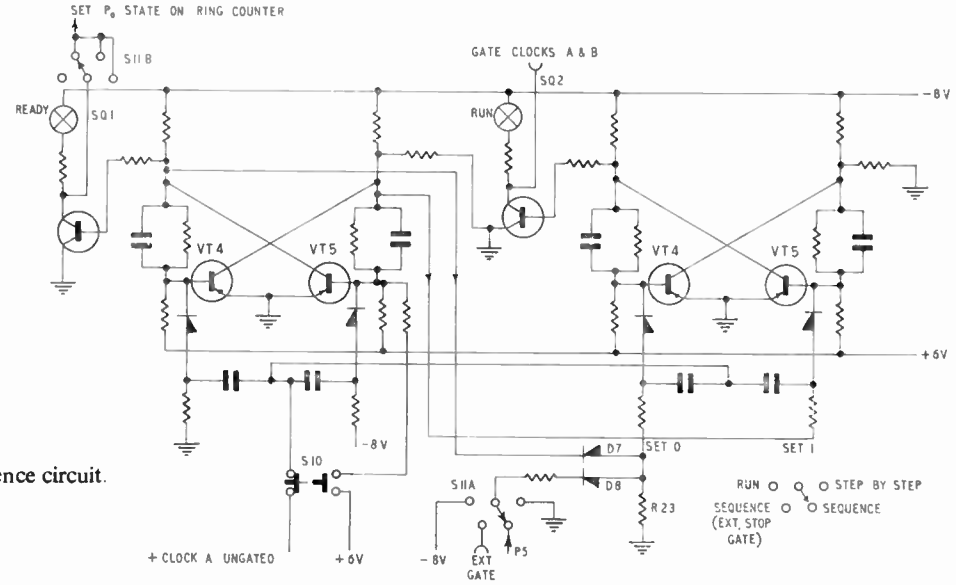


Fig. 26. Sequence circuit.

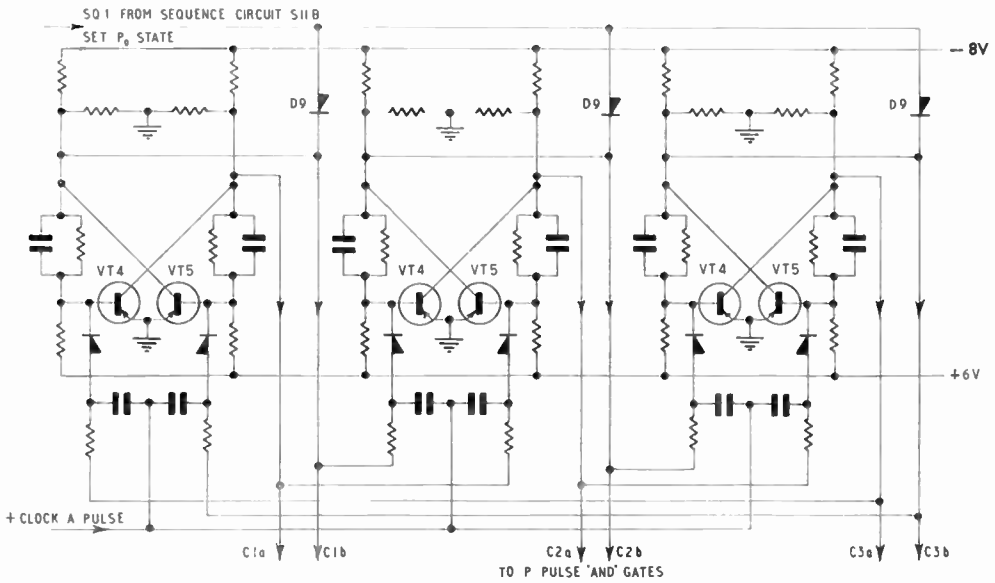


Fig. 27. Ring counter.

rent required during the pulses. It was found that on step-by-step operation, two clock A pulses were passed instead of one, because Sq2 was turned on too quickly by the ungated clock pulse. This difficulty was overcome by the capacitor C15, which increases the rise time of the gate pulse Sq2.

5.4.3. Sequence circuit

The principle of operation has already been described, and the circuit diagram is given in Fig. 26. The AND gate coupling the set "0" condition into Sq2 is composed of the diodes D7, D8 and the resistor R28. The gate is enabled only when the inputs to both diodes are at or near earth potential, and the position of the switch S11 decides when this will occur. The indicating lamps of the two bistable elements are mounted on the front panel and are labelled "Ready" and "Run".

5.4.4. Ring counter

Three standard printed circuit cards are used, but for economy the lamp driving transistors are omitted. The circuit diagram (Fig. 27), shows an additional diode D9 connected to the collector of each transistor VT5 and Sq1 applied to all the anodes of the diodes. If any VT5 is not conducting at the beginning of a sequence, its collector potential is negative with

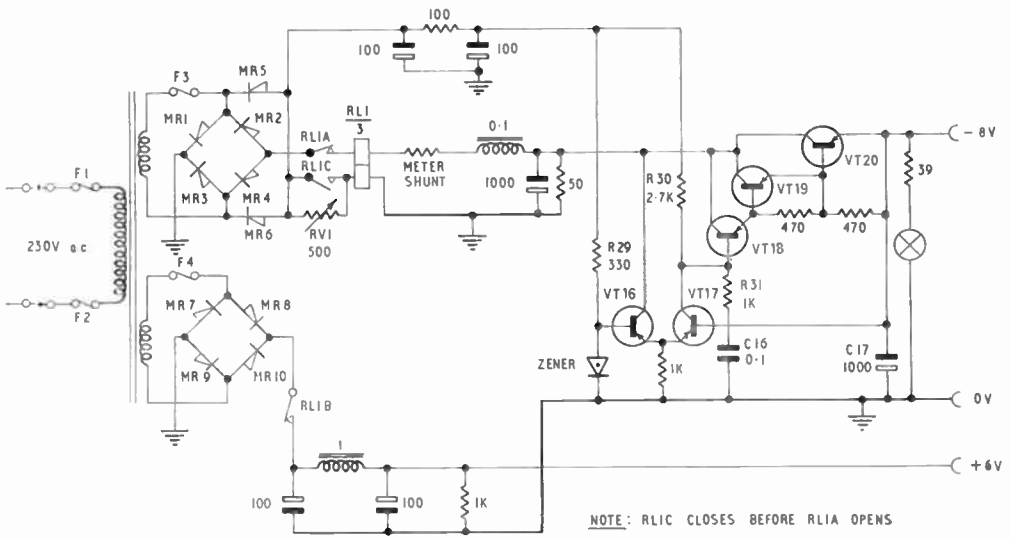


Fig. 28. Power supply unit.

respect to Sq1, hence D9 conducts pulling the collector towards earth. This positive-going change of potential is applied to the base of VT4, which cuts off and by normal switching action drives VT5 on. In this way all stages of the counter are set to "0" at the beginning of a sequence. Before the first clock pulse is applied to the counter, Sq1 returns to "0" (−8 V), hence the diodes D9 are normally cut off and have a negligible effect on the operation of the ring counter.

5.4.5. Power supply unit

The computing elements require two d.c. power supplies at −8 V and +6 V. The latter is comparatively lightly loaded at 50 mA, but the maximum load on the −8 V supply is about 4 A and may vary considerably from one clock pulse to another if many lamps light together. It was decided therefore that a stabilized negative supply with a low output impedance would be required to eliminate any undesirable fluctuations or cross couplings. No special arrangements are necessary for the positive supply.

Figure 28 shows the circuit diagram of the complete power unit. The bridge rectifier MR1–4 feeds a choke input filter to give a −11 V roughly-smoothed supply, and a low current −16 V supply for a part of the stabilizer circuit is derived from another bridge rectifier MR 1, 2, 5 and 6 with a capacitance input filter. The +6 V supply is a conventional capacitance input circuit.

The negative stabilized output voltage is compared with the Zener diode reference voltage by the long tailed pair VT16 and VT17, and the output from the latter is used to control a compound emitter follower of very high current gain, VT18, 19 and 20, and hence to control the output voltage.[4] A much improved performance is obtained by returning R29 and R30 to the −16 V line rather than the −11 V supply. The components R31 and C16 reduce the loop gain at the higher frequencies where the transistor phase shifts would cause oscillation and C17 maintains a low output impedance at frequencies where the stabilizer is ineffective.

A self-locking relay is used to safeguard the series transistors in the event of an overload or short circuit. The steady current in the "hold" winding is adjusted by RV1 so that the relay operates with 4 A flowing in the main winding. Both supplies are broken by the relay and can only be reconnected by opening the mains switch.

6. Experimental Procedure

The student is instructed to examine the operation of all the basic logical elements and to prepare a truth table for each by direct observation of the results. A NOT/OR and an AND gate can then be arranged to form a half adder (see Fig. 7) and the

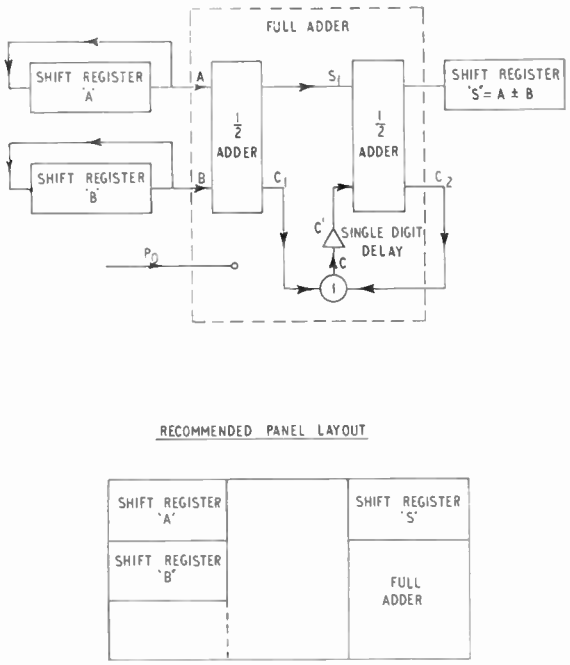


Fig. 29. Addition and subtraction.

results compared with those obtained from a standard half adder panel. Simple d.c. potentials are used for this preliminary work and are applied by operation of the input switches provided on each panel.

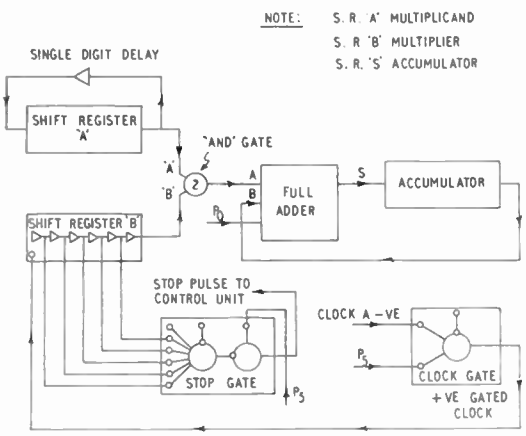
In order to carry out all the processes of arithmetic more complex circuits are needed. For addition and subtraction the full adder panel is used as in Fig. 29 and the student is invited to work out suitable examples on paper showing all "carries". The problem is then carried out on the computer at "Step-by-Step" and the operation checked at each stage against the paper calculation. The example is repeated at "Sequence/Fast" to show that the machine can perform almost instantaneously. The layout recommended in Fig. 29 has been specifically chosen so that further panels can readily be added for multiplication or division.

Figure 30 shows the arrangement required for multiplication, which is carried out by cumulative addition with shift, and the procedure requires that:

(i) The multiplier should double in significance after each word period and this is achieved by including a single digit delay in its recirculatory loop.

(ii) The multiplicand should be examined digit by digit beginning with the least significant and hence requires a shift pulse at the end of each word period. The pulse is obtained by gating clock A negative pulses with P5.

(iii) The calculation should terminate when the multiplier is empty.



RECOMMENDED PANEL LAYOUT

SHIFT REGISTER 'A'	SINGLE DIGIT DELAY	SHIFT REGISTER 'S'
SHIFT REGISTER 'B'	"AND" GATE	FULL ADDER
STOP GATE	CLOCK GATE	

Fig. 30. Multiplication.

The state of the multiplier store is monitored by the stop pulse gate and a stop pulse is generated at the end of the final word period. The machine is run at "External Stop Gate", and will run for several words depending on the particular example chosen. The student is again advised to select and work out an example on paper before operating the machine.

Division is performed by repeated subtraction of the divisor from the dividend until the remainder is zero or negative. In the first case the quotient is the number of subtractions, but when the remainder is negative, the quotient is one less than the number of subtractions. The most significant digit of the dividend is used as a sign digit and hence the dividend is restricted to five digits.

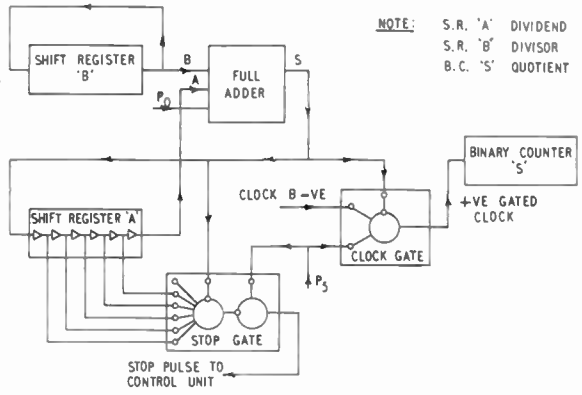
Referring to Fig. 31 during each word period the divisor is subtracted from the dividend and the remainder fed back into the dividend register, and P5 gates a clock pulse to the binary counter which records the number of subtractions. The divisor is also recirculated. The process is repeated until the dividend is reduced to zero or becomes negative. When the remainder is zero the stop pulse unit gates P5 to produce a stop pulse as in the case for multiplication. If the dividend is not exactly divisible, the final subtraction gives "1" in the most significant digit of the dividend to indicate a negative remainder. The "1" is used to inhibit the clock pulse gate and hence the binary counter records one less than the total number of subtractions. The "1" is also used to inhibit the OR gate in the stop pulse unit and its output no longer inhibits the second OR gate, thus P5 is gated to produce the stop pulse. The true remainder is then the sum of the divisor and the dividend after the final sequence. The student is advised to choose two examples, one being exactly divisible and the other producing a remainder.

By means of the various conversion units available, data transmission techniques can also be studied, but it is not considered necessary to include specific layout diagrams for these operations.

7. Conclusions

The equipment has been in use at Shrivenham for approximately two years and during this time has proved to be of immense value, not only for experimental work but also for lecture demonstrations. Experience has shown that the slow speed operation is almost unnecessary, since the "step-by-step" facility is more versatile and allows the speed to be adjusted to individual requirements. If the provision of an automatic slow speed is required, it is suggested that several bits per second would be preferable to the one bit in two seconds provided.

It is sometimes necessary to demonstrate all the processes of arithmetic during one lecture period, and it was found that in practice short delays were inevitable to allow rearrangement of the circuits. In order to avoid this, the wiring can be routed through a multipole switch mounted on a small sub-chassis, and



RECOMMENDED PANEL LAYOUT

S.R. 'B'		BINARY COUNTER
S.R 'A'		FULL ADDER
STOP GATE	CLOCK GATE	

Fig. 31. Division.

separate block diagrams used to illustrate each of the individual circuits.

Another difficulty which perhaps is not so obvious is deciding on a suitable length for the patch cords, which ideally should be straight line connections between panels. By using electrolastic cable which extends to two and a half times its original length, it is possible to use only a few standard lengths and this also simplifies the storage problem. Suitable peg boards are mounted internally on each side of the storage cabinet so that the patch cords can be kept in reasonable order.

Several inquiries have been received regarding the cost of manufacturing the training aid, which it is very difficult to estimate, particularly with regard to the labour costs. It can be stated, however, that all the materials and components used were purchased for approximately £300. It is suggested that anyone intending to build a similar device might well be advised to start with a less ambitious project initially, as long as the basic framework is large enough to allow expansion at some later date.

8. Acknowledgment

The authors wish to record their thanks to the Commandant of the Royal Military College of Science for granting permission to publish this paper.

The work was carried out as a design exercise under the general supervision of Lt.-Col. I. W. Peck, R.A. (Associate Member) by a syndicate of four officers, Captains T. Graham and B. E. Thorteinson, Royal Canadian Artillery, Captains J. L. Akass and J. L. Heard, Royal Signals.

9. References

1. C. G. Bradshaw and E. F. Stallard, "Junction Transistors and Their Application to a Shift Register", S.R.D.E. Technical Memorandum No. RES 177, March 1955.
2. C. G. Bradshaw, "A versatile 100 kc/s Transistor Counter Chronometer for Frequency Measurement", S.R.D.E. Report No. 1106, October 1956.
3. E. F. Stallard, "Variable Frequency Division Using a Transistor Shifting Register", S.R.D.E. Technical Memorandum No. RES 199, July 1957.
4. W. L. Stephenson, "A mains-operated d.c. stabilized transistor power supply for laboratory use", *Mullard Tech. Communs*, 3, No. 29, pp. 282–284, March 1958.
5. S. H. Hollingdale, "High Speed Computing: Methods and Applications" (English Universities Press, London 1959).

Manuscript first received by the Institution on 5th March 1960, in revised form on 7th March 1962, and in final form on 7th July 1962. (Paper No. 806/C51.)

U.D.C. 621.375.9

Parametric Amplifiers: Static and Dynamic Inductance and Capacitance and their Significance in the Non-Linear and Time-Varying Approaches

By

F. J. HYDE, D.Sc.†

AND

Professor D. G. TUCKER, D.Sc. (*Member*)‡

This paper is sponsored by the Institution's Education Group.

Summary: The analysis of circuits such as parametric amplifiers can be approached in two main ways, by regarding the variable element as having either non-linear characteristics largely controlled by a dominant pumping current or voltage, or linear characteristics subject to a specified time-variation. In the former case, "dynamic" inductance or capacitance may be used with advantage in analysis and measurement in place of the "static" parameters. In the latter case the distinction is invalid. The relation between the two cases is discussed in this paper.

1. Introduction

The analysis of circuits such as rectifier modulators and parametric amplifiers can be approached in two main ways:[1,2] in specific configurations it is often convenient to regard the rectifier, inductor or capacitor as a non-linear element largely controlled by a dominant carrier or pumping current (or voltage), while in the development from general network theory it is convenient to regard the variable elements as being linear but having a specified time-variation. In the case of rectifiers and time-varying resistances there is generally no difficulty in reconciling these approaches, and there is ample literature. For the non-linear or time-varying inductance and capacitance it seems that the matter still requires some clarification.

One of the authors has recently discussed[3] the use of "dynamic" parameters of inductance and capacitance in place of "static" parameters in the analysis of parametric amplifiers by the non-linear approach. Whereas static inductance and capacitance are defined as

$$L_s = \phi/i$$

where ϕ = flux-linkage

and i = current

and

$$C_s = q/v$$

where q = charge

and v = voltage

the dynamic parameters are defined as:

$$L_d = \mathrm{d}\phi/\mathrm{d}i$$

and

$$C_d = \mathrm{d}q/\mathrm{d}v$$

The use of dynamic parameters has some advantages. In making practical measurements of the parameters, L_d is more easily measured than L_s, and C_d more easily than C_s. It is often therefore advantageous and more realistic to use L_d and C_d in the analysis.

The other author has been concerned with teaching the theory of circuits with time-varying parameters,[4] and discussion has arisen between the authors on the significance of the static and dynamic conceptions in this approach. The conclusions reached are made clear in the following treatment of the parametric amplifier with pumped inductance. The case of pumped capacitance is the exact dual of this, and needs no separate treatment.

The symbols used in this paper are, as far as possible, the same as in the paper[4] on time-varying parameters.

2. Non-linear Approach

In the approach through non-linearity we assume currents in the inductance thus:

$I = \hat{I} \cos \omega_p t$ due to the pumping source at angular frequency ω_p

$i_0 = \hat{\imath}_0 \cos(\omega_q t + \theta_0)$ at the angular frequency (ω_q) of the input signal,

$i_{+1} = \hat{\imath}_{+1} \cos[(\omega_q + \omega_p)t + \theta_{+1})]$ at the angular frequency $(\omega_q + \omega_p)$ of the output signal.

It is assumed that the circuit is tuned so that no other currents exist.

Now we have to ascribe a non-linear function, i.e. a current dependence, to the inductance. It is complicated to analyse the problem in general terms, but analysis is simple if we assume a second-order power-series dependence of flux on current, namely,

$$\phi = L_0 i + L_1' i^2$$

giving static inductance,

$$L_s = \phi/i = L_0 + L_1' i \qquad \ldots\ldots(1)$$

and dynamic inductance,

$$L_d = \mathrm{d}\phi/\mathrm{d}i = L_0 + 2L_1' i = L_0(1+\alpha i) \text{ say} \qquad \ldots\ldots(2)$$

where $i = I + i_0 + i_{+1}$.

† Department of Electronic Engineering, University College of North Wales, Bangor.

‡ Electrical Engineering Department, University of Birmingham.

To obtain this kind of non-linear law clearly requires the presence of a bias current in a real inductor.

The voltage v across the inductance may be expressed as either

$$\frac{d\phi}{dt} = \frac{d}{dt}(L_s i) = \frac{d}{dt}(L_0 i + L_1 i^2) \quad \text{.....(3)}$$

or

$$\frac{d\phi}{dt} = \frac{d\phi}{di} \cdot \frac{di}{dt} = L_d \frac{di}{dt} \quad \text{.....(4)}$$

according to whether static or dynamic inductance is used.

Now eqn. (4) gives simply

$$v = L_0[1+\alpha(I+i_0+i_{+1})]\frac{d}{dt}(I+i_0+i_{+1}) \quad \text{.....(5)}$$

and eqn. (3) is readily shown to give exactly the same result. On separating out the terms of the various frequencies, and assuming that no currents other than I, i_0 and i_{+1} have finite amplitude, eqn. (5) leads readily to the equations for frequencies ω_q and $\omega_q+\omega_p$. These may be regarded as the basic equations of the circuit of Fig. 1 on putting

$v = E\cos\omega_q t - Z_0 \hat{\imath}_0 \cos(\omega_q t+\theta_0)$ at frequency ω_q

and

$v = Z_{+1}\hat{\imath}_{+1}\cos[(\omega_q+\omega_p)t+\theta_{+1}]$ at frequency $\omega_q+\omega_p$.

Replacing the cosine terms by current vectors $\boldsymbol{i}_0$ (carrying the phase angle θ_0) and $\boldsymbol{i}_{+1}$ (carrying the phase angle θ_{+1}), and putting $L_1 = \alpha L_0 \hat{I}$, we obtain the equations:

For frequency ω_q:

$$E - Z_0\boldsymbol{i}_0 = j\omega_q L_0 \boldsymbol{i}_0 + j\tfrac{1}{2}\omega_q L_1 \boldsymbol{i}_{+1} \quad \text{.....(6)}$$

For frequency $\omega_q+\omega_p$:

$$-Z_{+1}\boldsymbol{i}_{+1} = j\tfrac{1}{2}(\omega_q+\omega_p)L_1\boldsymbol{i}_0 + j(\omega_q+\omega_p)L_0\boldsymbol{i}_{+1} \quad \text{.....(7)}$$

It is implied above that $\hat{I}$ is given; if it is not, then a third equation, for frequency ω_p, must be taken so that a solution is obtained for I as well as i_0 and i_{+1}.

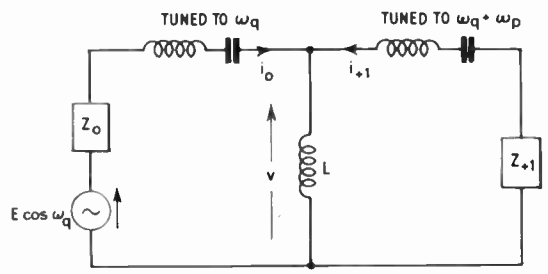


Fig. 1. Parametric amplifier.

In the non-linear approach, L is a function of the total current through it, and has also a third circuit, with a generator, driving current I at frequency ω_p through it.

In the time-varying approach, L is a specified periodic function of time with fundamental frequency ω_p.

It is apparent that, for the particular dependence of flux on current which has been assumed, the only difference in the analysis between the use of static inductance (L_s) and dynamic inductance (L_d) is that the variable part of the former has to be only half as great as that of the latter to give the same result.

Note that if we had taken an infinite power series to represent the non-linearity in eqns. (1) and (2), then to obtain eqn. (6) and (7) we should have had to specify that the pumping current was much larger than the other currents. Furthermore, we should have had to define L_1 as an infinite series involving all odd-order coefficients and powers of I, and L_0 as an infinite series involving all even-order coefficients and powers of I.

From equations (6) and (7) the performance of the up-converter or of the negative-resistance amplifier is readily worked out.[4] (In the latter case, the frequency $\omega_q-\omega_p$ has to be substituted for $\omega_q+\omega_p$ in all the working.)

3. Time-varying Approach

In the approach through time-varying parameters, we put L_s or L_d as a function of time, but not of current. Thus a Fourier expansion is possible:

$$\frac{\phi}{i}(t) = L_s(t) = \sum_{n=0}^{\infty} L_n \cos n\omega_p t \quad \text{.....(8)}$$

and

$$v = \frac{d}{dt}[L_s(t)\,.\,i] \quad \text{.....(9)}$$

where i now comprises only i_0 and i_{+1}.

Since the inductance is not a function of current,

$$\frac{d\phi}{di} = \frac{\phi}{i} \text{ or } L_d = L_s \quad \text{.....(10)}$$

and both $\frac{d\phi}{di}$ and $\frac{\phi}{i}$ are the same functions of time. *There is therefore no distinction between static and dynamic parameters in this approach.* Thus we cannot apply eqn. (4) to this case, and the expansion must be made only from eqns. (8) and (9) taken together. Expansion leads to exactly the same equations as (6) and (7), (where L_1 is now the value of L_n in (8) for $n = 1$), without the loss of generality implied by the limited non-linear representation of eqn. (1); on the other hand the assumption of linear response to the signal has been necessary. This is equivalent to saying that, when pumping is done by a current driven through the inductance rather than by mechanical variation, the pumping current must be much larger than the signal currents.

The fact that identical equations are obtained by the two approaches needs a little further explanation

in view of the fact that L_1 is defined differently in the two cases. Referring to the second paragraph from the end of Section 2, and observing that L_0 is obviously the same in both cases, it becomes clear that with a dominant pumping current, L_1 determined from the general non-linear equation must be identical with L_1 in the Fourier series of eqn. (8), since the two circuits are now operationally identical, have identical equations, and must have identical solutions.[5] The fact that L_1 is also the same for the simple non-linear law of eqns. (1) and (2), which does not require the assumption of a dominant pumping current, indicates that this is evidently a special case.

4. References

1. S. Bloom and K. K. N. Chang, "Theory of parametric amplification using non-linear reactances", *R.C.A. Rev.*, **18**, p. 578, December 1957.
2. H. Heffner and G. Wade, "Gain, bandwidth and noise characteristics of the variable-parameter amplifier", *J. Appl. Phys.*, **29**, p. 1321, September 1958.
3. F. J. Hyde, Discussion on paper by A. L. Cullen, "Theory of the travelling-wave parametric amplifier", *Proc. Instn Elect. Engrs*, **107**, Part B, p. 124, 1960.
4. D. G. Tucker, "Circuits with time-varying parameters", *J. Brit.I.R.E.*, **25**, No. 3, p. 263, March 1963.
5. F. J. Hyde, "Analysis of parametric amplifiers incorporating varactor diodes", *Proc. Instn Elect. Engrs*, in course of publication. (This paper gives a fuller discussion of the matter.)

Manuscript first received by the Institution on 20th May 1962 and in final form on 14th December 1962 (Contribution No. 63/Ed.10).

Brit.I.R.E. GRADUATESHIP EXAMINATION, NOVEMBER 1962—PASS LISTS

The following candidates who sat the November 1962 examination at centres outside Great Britain and Ireland succeeded in the sections indicated. The examination, which was conducted at 65 centres throughout the world, attracted entries from 618 candidates. Of these 187 sat the examination at centres in Great Britain and Ireland and 220 sat the examination at centres overseas. The names of successful candidates resident in Great Britain and Ireland were published in the March issue of the *Proceedings* of the Brit.I.R.E.

Section A	*Candidates appearing*	*Pass*	*Fail*	*Refer*
Great Britain	89	54	27	8
Overseas	124	38	73	13
Section B				
Great Britain	98	27	46	25
Overseas	96	16	71	9

OVERSEAS

The following candidates have now completed the Graduateship Examination and thus qualify for transfer or election to Graduate or a higher grade of membership.

AMIRTHALINGHAM, Savanamuthu (S), *Ceylon.*
AVERY, Elliott George (S), *Nairobi, Kenya.*
CALLENDER, Peter Charles (S), *Auckland, New Zealand.*
CHAN, Chee Khong (S), *Singapore 13.*
FRY, Anthony Jack (S), *Elizabeth, South Australia.*
GHALLEY, Jaspal Singh (S), *Nairobi, Kenya.*
GOYAL, Badrinarain (S), *Bombay 4, India.*
JONES, K. Philip (S), *Poona 6, India.*
JOSHI, Sivasanker Rajaram (S), *Madras, India.*
KHOO, Poon Tong (S), *Singapore 10.*
KOHEN, Dan, *Givatayim, Israel.*
LOBO, Leo Cyril (S), *Kampala, Uganda.*
MEHTANI, Mahendra Nath (S), *New Delhi, India.*
ONDRICH, Milan (S), *Scarborough, Ontario, Canada.*
TAM, Shun Kwong (S), *Hong Kong.*
WONG, Yau Liong (S), *Singapore 1.*

The following candidates have now satisfied the requirements of Section A of the Graduateship Examination.

AIKHOJE, Samson Omo (S), *Lagos, Nigeria.*
ASLAM, Mohamed (S), *Karachi, Pakistan.*
BERLINER, Carlo Carmi (S), *Tel-Aviv, Israel.*
BISWAS, Pares Nath (S), *West Bengal, India.*
BODEWIG, Hugo Wilhelm Karl (S), *Johannesburg South Africa.*
BOSWORTH, Richard William (Associate), *Salisbury, Southern Rhodesia.*
DORON, Itzhak (S), *Tel-Aviv, Israel.*
GAUR, Bibhuti Bhusham (S), *Allahabad, India.*
GOWDARA, Shdaicsharappa (S), *Sekondi, Ghana.*
IFTIKHAR (S), *Rawalpindi, Pakistan.*
JAYAWARDHANA, Stanley Cecil (S), *Panadura, Ceylon.*
JOHNSON, Ernest Arthur (S), *Toronto, Canada.*
KALYAN, Singh (S), *Poona 6, India.*
KAPOOR, Suresh Kumar (S), *Bangalore 15, India.*
KRUEGER, Horst (S), *New South Wales, Australia.*
KUSHWAHA, Ramesh Chandra (S), *Kanpur, India.*
LANDEN, Joseph (S), *Ramat-Gan, Israel.*
LEE, Hon Wah (S), *Hong Kong.*
MORGAN, David Hambry (S), *Geneva, Switzerland.*
MURTHY, A. V. R. (S), *Bangalore, India.*
NWADUKWE, Dennis Enweluzo (S), *Nigeria.*
ODURUKWE, Boniface Chukuemeka (S), *Lagos, Nigeria.*
OLANIYAN, Emman B. (S), *Lagos, Nigeria.*
PARANN, Moshe, *Tel Aviv, Israel.*
PIROOZMANDI, Rouhollah (S), *Abadan, Iran.*
PRABHU, Ramachandra (S), *Kerala, India.*
RAMACHANDRA, K. (S), *Bangalore, India.*
RAMACHANDRAN, Chinniah (S), *Colombo, Ceylon.*
RAMAKUMARM, Rajupeta (S), *Bangalore, India.*
RAMAKRISHNA, J. S. (S), *Bangalore, India.*
RAMANUNNI, Anupurath Pisharath (S), *Kerala, India.*
ROSSI, Francis (S), *Malta.*
SAINT-CYR, Frederic Gaetan (S), *Montreal 35, Canada.*
SAX, Nicolaas (S), *Suva, Fiji.*
SCAFE, Roy Barrington (S), *Jamaica.*
SCHER, Joseph (S), *Ramat Chen, Israel.*
SETHURAMAIAH, Hanasoge Naganappa (S), *Hyderabad, India.*
SHOLANKE, Patrick Antanda (S), *Lagos, Nigeria.*
SPRAGGS, Malcolm John (S), *Devonshire, Bermuda.*
TAN, Seng Siew (S), *Singapore 25.*
THAMPI, Sukumaran (S), *Bombay, India.*
VENUGOPALAN, K. (S), *Jamnagar, India.*

The question papers set in the November 1962 Graduateship Examination have been published in the January-February, March and April 1963 issues of the *Proceedings* of the Brit.I.R.E., together with answers to numerical questions and examiners' comments.

These questions, answers and comments are now available as reprints, price 3s. 6d. for the question papers of Section A, and 5s. for those of Section B.

(S) denotes a Registered Student.

U.D.C. 621.391.833.24: 621.397.743

The Performance Evaluation of Television Line and Microwave Links using Sine-Squared Test Techniques

By

K. H. POTTS,
(*Associate Member*) †

Summary: The principles involved in the use of sine-squared pulse test techniques and the application of this type of test to the assessment of the linear transmission performance of television line and microwave radio links are surveyed in detail. The characteristics of the sine-squared pulse are dealt with from both a mathematical and practical point of view, attention being drawn to the essential difference and flexibility of waveform test techniques as opposed to the steady state type of tests (gain/frequency —phase/frequency). The application of waveform testing to the linear performance evaluation of television line and radio links on a day-to-day and commissioning basis are dealt with and methods of expressing the inherent waveform distortion in terms of a *K*-rating with reference to a given waveform tolerance are explained. Finally, there are examples of waveform tests as applied to the linear performance assessment to studio centre—transmitter links and other items of equipment.

1. Introduction

The basic requirement for the distortionless transmission of television signals over any form of link or system is that the signal at the receive terminal be identical to that being transmitted within the specified tolerances of the system under test. In other words, the quality of the transmitted picture is directly related to the impairment and deviation from the specified limits of the received waveform after transmission over the system.

The majority of television links have well defined upper frequency limits or fall-off points, the links being generally designed and commissioned to approximate to the characteristics of an ideal low-pass filter with a video-to-video fall-off point of around 3·0 Mc/s. The object of this is to enable links to be operated in tandem with ideally no accumulative loss up to the upper frequency limit.

To ensure the distortionless transmission of signals over a television system it is necessary for the gain frequency characteristic of the system to be constant with frequency over the system bandwidth and the phase shift to be proportional to frequency. Of these two parameters, the former is easily measured by standard techniques whilst the measurement of phase shift over the system bandwidth is a more involved process. However, the assessment of the characteristics of a television system using these steady-state methods has a serious disadvantage in that a subjective appreciation of the system performance is not obtained. In other words, the distortion present in the system which may be the sum of several different types of distortion cannot be directly related in terms of picture impairment which, after all, is the end test of any system used for the transmission of picture signals.

† Wales, West and North Television Ltd., Cardiff; formerly with Anglia Television Ltd.

As an alternative to quoting the performance of a television link in terms of gain frequency and phase shift it can be expressed in terms of time with reference to a specific waveform of a recurrent nature. This is a more desirable method of estimating the performance of a television system since a dynamic measurement is obtained and furthermore, the results obtained can be more readily related to a given series of subjective effects.

Extending the waveform treatment, it is found that if a television picture is considered in relation to its waveform, then the picture resolution is found to be a function of the time of rise and fall of the components of the transmitted complete waveform, whilst picture ringing is found to be a function of the overshoots of the signal components. Furthermore, if the picture traverses a transmission link with inherent phase/frequency distortion, then the signal is so affected, as to result in a received picture which displays streaking after objects contained in the picture.

Having considered the signal waveform with respect to the subjective picture effect it is clear that a relationship exists which can be extended to a form of testing other than the normal steady-state sine-wave tests, which are not related to a subjective function of a television picture. This type of testing is termed waveform testing and as the name implies, the test waveform bears a positive relationship to given subjective picture effects.

The logical application of waveform testing is further substantiated when the Fourier Theorem is considered in the following light. This states that any periodic function can be expressed as a series of sinusoids whose frequencies are integral multiples of the fundamental frequency; in other words, harmonically related to the fundamental frequency of recurrence. This method of treatment can be extended

to non-recurrent waveforms in the Fourier Integral by letting the fundamental tend to zero. Thus a transmission system passing sinusoidal waves within its pass band will transmit without distortion any other waveform whose spectrum is within the pass band of the system. This means that it is a sine-wave test and may be replaced by a waveform test if all the restrictions are met. On considering the points made, it is clear that waveform testing must offer advantages when used for the testing of television transmission systems.

It is the object of this paper to show the advantages of waveform testing in relation to the steady-state type of sine-wave test and to explain the basic theory and method of application of waveform testing to the evaluation of the linear distortion present in a television line or microwave radio link.

2. The Transmission Properties of Television Links in Terms of Steady-state Test Methods

The two main transmission properties of any system used for the carrying of television signals will now be examined in relation to steady-state sine wave testing.

The steady-state test indicates the performance of the system at any one given frequency and as such is not a dynamic measurement, in short the test is not related to a given subjective picture effect. The two main transmission properties are gain/frequency response and phase/frequency response, which will be explained briefly in order that the full implications of waveform testing are appreciated later in the paper.

2.1. *Gain/Frequency Response*

The gain/frequency characteristic of a television link should be constant with frequency over the video bandwidth to the extent that if a sinusoidal signal of 1 volt d.a.p. is transmitted over the link then the received signal at the distant terminal is identical in amplitude over the whole of the video working band, within the specified tolerances of the system. If the standard television waveform for the 405-line system is examined it will be seen that the components of this will be affected by gain/frequency distortion such that the line synchronizing pulses have the time of rise increased for a falling gain/frequency characteristic, in other words the rise time is slow. This affects the received picture to reduce fine detail contained in the transmitted signal to general background level, due to the attenuation of the upper video frequency components and consequently the picture is of poor definition.

The gain/frequency characteristic of a typical television transmission system[1] is shown in Fig. 1 and is the video-to-video resultant of a complete transmission system whose cable attenuation is compensated and equalized at given intervals by amplifiers, waveform correctors and equalizers. This result is the observed characteristic at the receive terminal of the system, when the transmit terminal is sending a standard 1 volt sinusoidal signal. It will be noted that the video level is constant over the working bandwidth and tends to fall off at 3·0 Mc/s, in a fashion approximating to that of an ideal low-pass filter, this being a desirable feature for the operation of several television systems in tandem.

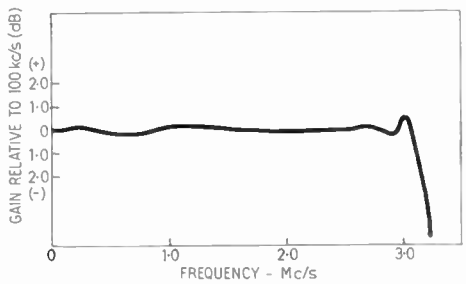


Fig. 1. The overall (video to video) gain/frequency characteristic of a typical line television link.

2.2. *Phase/Frequency and Group Delay Characteristic*

It is of considerable importance in a television transmission system that the individual times of transmission of all frequencies in the video spectrum are substantially equal. Otherwise the distortion due to a difference in transmission time would be very obvious on picture transmission.

When the time of propagation of a signal through a television system varies with frequency then there is phase frequency distortion present in the system, in other words the phase/frequency response is non-linear.

The ideal characteristic of phase shift/frequency is shown in Fig. 2. The time of transmission of a signal through a television system is given by

$$\tau = \frac{\beta}{\omega} \qquad \text{.....(1)}$$

where β = phase shift

τ = phase delay

ω = angular frequency.

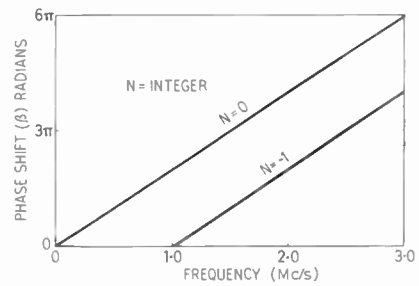


Fig. 2. Ideal characteristic of phase shift/frequency.

If a signal is transmitted through a system having inherent phase/frequency distortion then the result is a variation of the group velocities through the system which results in a spreading-out of the transmitted finite waveform and resultant distortion.

The transmission time of a single frequency over a system is termed the absolute phase delay which is derived as follows:

The phase shift through a circuit can be expressed as

$$\beta = 2\pi f\tau + 2N\pi \qquad \text{.....(2)}$$

Thus

$$\tau' = \frac{\beta - 2N\pi}{2\pi f} = \frac{\beta}{\omega} - \frac{N}{f} \qquad \text{.....(3)}$$

where f = frequency in kc/s

τ' = absolute phase delay

N = integer.

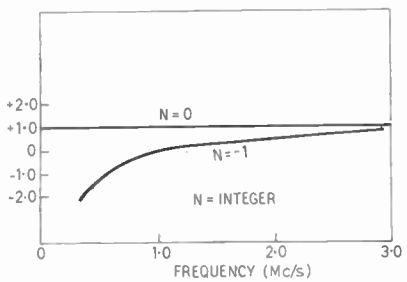


Fig. 3. Ideal characteristic of phase delay/frequency.

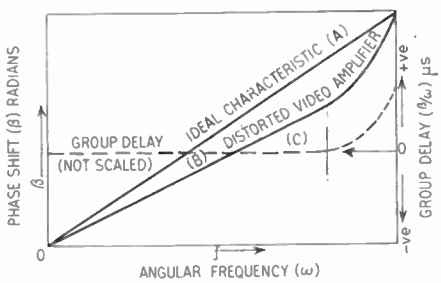


Fig. 4. The relationship of phase shift to group delay.

Phase delay is essentially a steady-state characteristic and in utilizing it there can be no differentiation between the successive cycles of a sinusoidal waveform. Thus when using this method it is usual to plot the phase delay characteristic nearest to the horizontal which is the ideal, as shown in Fig. 3 where the inherent distortion is indicated by the deviation from the mean horizontal line (i.e. $N \neq 0$).

Considering a video amplifier with inherent phase/frequency distortion as shown in curve (B) in Fig. 4, then it is difficult to assess the phase distortion present by phase delay methods when the amount of distortion is small; this is due to difficulties in obtaining accurate tangents to points on the phase/frequency curve. A more satisfactory method is to consider the slope of

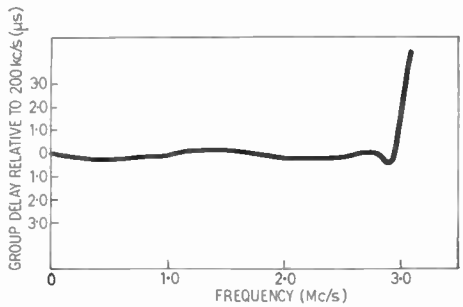


Fig. 5. The overall (video to video) group delay/frequency characteristic of a typical line television link.

the phase/frequency characteristic which gives the group delay $d\beta/d\omega$; this is illustrated in Fig. 4 whose curve (C) shows the approximate derivative of the phase/frequency curve (B).

Thus for television transmission systems operating on a carrier basis it is desirable to assess the system performance over the working range by taking the slope of the phase/frequency characteristic and obtaining the group delay $d\beta/d\omega$. It might be mentioned that the measurement of group delay can be obtained directly using special measuring equipment. A typical group delay characteristic,[1,2] for a coaxial cable carrier television link is shown in Fig. 5; this indicates the amount of delay which is normally expected at the video receive terminals of such a system. Table 1 shows the order of delay which is expected from a correctly aligned system,[1,2] over the working video range of the 405-line monochrome television system.

Table 1

Frequency (kc/s)	Group Delay Limits relative to 0 μs at 200 kc/s
200–2000	± 0·15 μs
2000–2500	± 0·25 μs
2500–2800	± 0·5 μs

3. Principles of Waveform Testing

Having considered the application of steady-state test methods to television transmission systems, it is readily apparent that there is no relationship between the results of this type of test and a given subjective picture effect, in other words, the tolerances of the system could not be expressed in terms of the test signal relative to a subjective picture effect. This is understandable since the sine wave is not representative of the television picture signal and in fact, a step function or impulse function is a closer approximation to a television picture signal.

As already explained, it is apparent that it would be of great help to be able to specify the performance of

a television transmission system in terms of time with reference to a specific waveform of a recurrent nature. In this way, the distortion existing in the system can be examined by passing a suitable waveform over the system whose tolerances are related to a given subjective picture effect. This is a logical step when it is considered that a steady-state test cannot be related to a subjective picture effect but a waveform test can always be related to a steady-state condition. As an interim step between complete waveform testing and steady-state testing, it is worth considering the pulse tests which were introduced as a rough method of assessing the linear performance of television systems from a transmission point of view.

The basis of the two tests was to pass a pulse over the system to assess (a) the h.f. response of the system and (b) to pass a pulse which indicated the l.f. response, these two pulses are now considered in brief.

3.1. *Previous Types of Test Pulses Used*

3.1.1. High frequency test pulse (Spike)

This pulse is termed a "spike" and is a step function type of pulse having the parameters as enumerated. It indicates the h.f. response of a system by measurement of the time of rise and fall of the pulse from its 10% to 90% amplitude limits. This indicates roughly the bandwidth of the system or upper operating frequency from:

$$f_c = \frac{1}{\text{T.O.R.}}$$

where $\begin{cases} f_c = \text{cut-off point of system} \\ \text{T.O.R.} = \text{time of rise.} \end{cases}$

The pulse also serves to indicate the amount of overshoot obtained after passing a transient over the system and the overshoot can actually be noted on a picture monitor. An illustration of a typical spike is shown in Fig. 6.

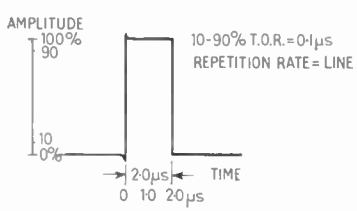


Fig. 6. H.f. test pulse (Spike).

Parameters of a typical h.f. test pulse.

Repetition frequency	= line. = 10·125 kc/s.
Pulse duration	= 1·8–2·0 μs.
Spike rise/decay time	= ⩾ 0·1 μs.
Spike rise/decay overshoot	= ⩾ 2%.

This pulse is of course, incorporated with the standard system synchronizing pulses in order that it can be transmitted as a complete waveform.

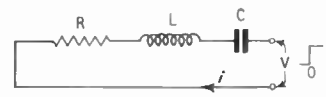


Fig. 7. Series mesh.

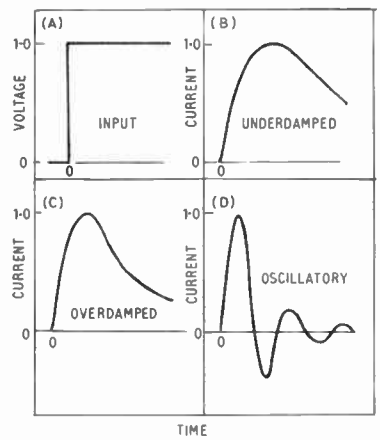


Fig. 8. The response of a series mesh to a step function.

3.1.2. Low frequency pulse (Window)

The purpose of the l.f. test pulse which is termed a "window" is to examine the l.f. portion of a television transmission system from a subjective point of view. Any distortion at low frequencies on the system shows up as a distortion of the test signal on a picture monitor (i.e. as a smeared signal).

Parameters of a typical l.f. test pulse.

Window rise/decay time	= ⩾ 0·25 μs.
Window rise/decay overshoots	= ⩾ 5%
Frame tilt	= ⩾ 1%
Window width	= 15–65% of picture width
Window vertical width	= 15–65% picture width.

As with the previous h.f. test pulse this pulse is incorporated with standard system synchronizing pulses in order that it can be transmitted as a complete waveform. It will be appreciated that these test pulses are not meant to be a precision test but an operational check on the system performance; nevertheless, they are a step towards waveform testing.

Before progressing to the requirements of a waveform for the performance assessment of television systems, the characteristics of the two basic types of pulse available for this work, namely, the step function and sinusoidal function will be considered.

3.2. *Step Function Analysis*

Consider a step function of unit amplitude applied to a series circuit of L, C and R as shown in Fig. 7.

Let $f(t) = 0$ when $-\infty < t < 0$

$f(t) = 1$ when $0 \leqslant t < \infty$

By Laplace transform:

For an LCR circuit i and V are related by:

$$\bar{i} = \frac{1}{L} \frac{p}{p^2 + pR/L + 1/LC} \cdot \bar{V} \qquad \text{.....(4)}$$

For unit step function $\bar{V} = V/p$

Thus $$\bar{i} = \frac{V}{L} \frac{1}{p^2 + R/L \cdot p + 1/LC}$$

or $$\bar{i} = \frac{V}{L} \frac{1}{(n+\mu)^2 + n^2} \qquad \text{.....(5)}$$

where $\mu = R/2L$ and $n^2 = 1/LC - R^2/4L^2$.

From eqn (5) three values of current are thus obtained depending on the relationship of $1/LC$ to R^2/L^2.

Solution (1)

For $n^2 > 0$ and +ve. $i = (V/Ln)\, e^{-\mu t} \sin nt$. Figure 8(*d*) shows the current waveform.

Solution (2)

For $n^2 = 0$, $i = (V/L)\, te^{-\mu t}$. Figure 8(*c*) shows the current waveform.

Solution (3)

For $n^2 > 0$ and −ve. $i = (V/kL)\, e^{-\mu t} \sinh kt$. Figure 8(*b*) shows the current waveform.

3.3. *Analysis of Sinusoidal Function*

Consider a signal of value $v = V \sin \omega t$ applied to an identical series combination of L, C and R as shown in Fig. 7.

Thus $$v = 1/C \int i\,dt + L\,di/dt + iR \qquad \text{.....(6)}$$

$$V \sin \omega t = i(j\omega L + 1/j\omega C + R)$$

$$i = \frac{V \sin \omega t}{j\omega L + 1/j\omega C + R}$$

Therefore $$i = \frac{V}{Z} \sin \omega t \qquad \text{.....(7)}$$

It is seen therefore that from eqn (7) there is only one solution of current when a simple sinusoidal signal is applied to the mesh. Furthermore, the same waveshape is present across any one component of the mesh since the derivative and integral of a sinusoidal function have the same waveshape as the original function, except for a time and amplitude displacement.

Thus, it has been established that the application of a unit step function to a series mesh results in three possible solutions of current depending on the degree of damping present in the mesh. Each of these three solutions gives a different waveshape across L and C to the applied waveshape and if the voltage across any one component is required then it must be calculated separately. For example, if the solution to the current is as shown in Section 3.2 solution (3)

$$i = \frac{V}{bL} e^{-at} \sinh bt$$

Then $$v_C = V\left[1 - e^{-at}\left(\cosh bt + \frac{a}{b} \sinh bt\right)\right]$$

and $$v_L = V\left[e^{-at}\left(\cosh bt - \frac{a}{b} \sinh bt\right)\right]$$

where $a = R/2L$ and $b = (R/4L - 1/LC)^{\frac{1}{2}}$.

From the treatment considered in this section it is clear that the application of a sinusoidal function gives a simple mathematical solution when compared to that of the step function. This will be appreciated further when the mathematical treatment of the sine-squared pulse is outlined.

4. The Specification of a Test Waveform for the Performance Evaluation of Television Systems with particular reference to the Pulse Component

The principle requirements for a test waveform which is to be used in the performance evaluation of television transmission systems can be summarized as follows:

(*a*) The waveform should present the information regarding the test in manner such that it is clearly understood by all grades of technical staff.

(*b*) The waveform should be reproducible by all authorities concerned.

(*c*) The waveform generator should not be complex or expensive to produce.

(*d*) The main parameter of the waveform which is the pulse component should be easily expressed as a mathematical function, thereby facilitating consistent and accurate analysis.

(*e*) In addition to the testing of the upper portion of the video spectrum it is desirable to have a composite signal such that information on more than one parameter of the system under test can be presented.

4.1. *Pulse Shape and Spectrum*

If the test pulse contains considerable energy outside the specified limits of the video band by virtue of its infinite frequency spectrum then clearly a considerable amount of distortion contained in the received pulse will be irrelevant for evaluating the performance of a television system having a specified upper frequency limit. In a pulse with a frequency limited spectrum no information of an irrelevant nature is contained within the pulse and standard test equipment as used on the link can be employed

in the setting up and checking of the waveform, since both have the same limiting frequency.

The problem of pulse shape which is interrelated with pulse spectrum has in this case two main solutions, as applicable to transient testing, namely, the use of a step function type of pulse or a sinusoidal function. The first suffers from the disadvantage of an extended frequency spectrum together with difficulties in precise measurement of its parameters at fast sweep speeds on the oscilloscope. Taking into consideration the previous treatment of the sinusoidal function coupled with the requirements of the test waveform it was found by the Post Office Engineering Department[3,5,6] that the most suitable type of pulse to use in this application was the sine-squared function. The choice of this pulse is readily apparent when its characteristics in the transient testing of television transmission systems are examined later.

4.2. *Pulse Duration*

It has been established that the time of rise of a step function bears an approximate relation to the bandwidth of the system under test. The sine-squared pulse must bear a more precise relationship to the characteristics of the system under test and it is for this purpose that the half amplitude duration of the pulse is used as a reference parameter. In order to illustrate the connection between the half amplitude duration of the sine-squared test pulse and the parameters of the system under test, the following fundamentals of the 405-line system are considered.

(1) In any television system the horizontal resolution of the system is its capacity to reproduce sharp changes in tonal value occurring along the horizontal scan lines, whilst the vertical resolution is its capacity to reproduce sharp changes in tonal value at 90 deg to the scan lines; in other words this represents the sharpness of the image. Clearly, vertical resolution can only be improved by increasing the number of lines in the system, whilst the horizontal resolution improves with an increase in the video bandwidth of the system.

(2) The limiting frequencies in the 405-line monochrome system and a relationship to the half amplitude duration of a sine-squared test pulse can be obtained from the fundamental consideration of the television camera pick-up tube.[4]

In any such tube the output at a given instant is directly proportional to the amount of light which has been received by that area of the target occupied by the scanning beam during the previous picture period. Due to the finite size of the scanning beam the output of such a tube is sinusoidal and therefore proportional to the cross-sectional area of the scanning beam and the tonal value of the scanned image.

The a.c. output of such a tube can be expressed as in eqn (8) which is derived from the basic equation

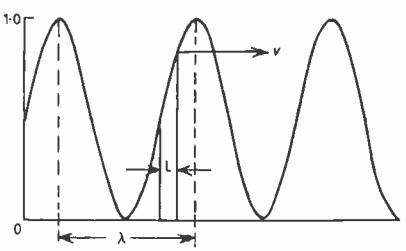


Fig. 9. The sinusoidal variation of tonal value in a television pick-up tube due to the finite size of scanning beam.

to the waveshape shown in Fig. 9 which illustrates the sinusoidal variation of tonal value due to the finite size of the scanning beam.

$$A = \frac{\lambda}{\pi} \sin \frac{l\pi}{\lambda} \qquad \text{.....(8)}$$

where λ = distance corresponding to one cycle of tonal value.

l = rectangular cross-sectional area of scanning beam.

Since $$\lambda = \frac{v}{f}$$

where v = velocity of the scanning beam.

$$A = \frac{lv}{f\pi l} \sin \frac{fl\pi}{v}$$

$$A = l \frac{\sin \theta}{\theta}$$

where $$\theta = \frac{\pi l f}{v}$$

Thus given a specific scan speed and beam area the frequency response of a camera pick-up tube is:

$$A = 20 \log \frac{\sin \theta}{\theta} \text{ dB} \qquad \text{.....(9)}$$

Now when $\sin \theta = 0$, zero output occurs from the tube. ($\theta = \pi$ rad)

This is when $\pi lf/v = \pi$

Thus $$f = v/l \qquad \text{....(10)}$$

f is the extinction frequency of the tube. For a given value of λ, i.e. for a given image size, there is an optimum value of scanning beam area. The tube output is a maximum when $\sin \theta = 1$, which occurs when $\theta = \pi/2$.

Therefore $$l = v/2_f \qquad \text{....(11)}$$

Let the picture contain L lines and the scanning beam be of cross sectional area of side l, where $l = b$ = width of scanning lines, and h = height of image on tube target.

Then $l = h/L$

The width of the image is ha, where a is the electrical aspect ratio of the system in which allowance is made for the effect of suppression time.

The length of one scan line = $ha = lLa$ and in one picture the beam scans L lines giving a total length of lL^2a.

In one second the total distance travelled by the beam = PaL^2l.

where P = the picture frequency of the system.

Thus the scan velocity is $v = PaL^2l$.

The extinction frequency f is thus

$$f = \frac{PaL^2l}{l} = PaL^2 \qquad \text{.....(12)}$$

For the 405-line system the extinction frequency can be deduced by substitution in (12)

$$f = PaL^2 = 25 \times 1 \times 516 \times 405^2 = 6{\cdot}22 \text{ Mc/s} \quad (13)$$

In the 405-line system $a = 1{\cdot}516$.

To transmit the full system resolution in the 405-line system the maximum transmitted frequency is given by $f/2$ since the highest frequency necessary to transmit detail the size of an element is given by $PaL^2/2$.

This therefore gives a figure of 3·11 Mc/s as the highest frequency transmitted in the 405-line system in order to attain detail of elemental size. The pulse duration of the test pulse must thus be closely related to the two limiting frequencies derived in the treatment of the pick-up tube which are fundamental to the system as a whole.

This is in fact the case as the two pulse durations (h.a.d.) chosen are $T = 0{\cdot}17$ μs and $2T = 0{\cdot}33$ μs which are seen to be equivalent to the figures already produced for the limiting frequencies.

5. Characteristics of the Sine-Squared Pulse

The salient characteristics of the sine-squared pulse when used as an integral part of a composite test waveform for the assessment of the linear distortion present in the upper region of the video band on the 405-line transmission links as used in this country will now be considered. The C.C.I.F. recommendations for the sine-squared pulse when used as one of the signals for the complete testing of a television system are:

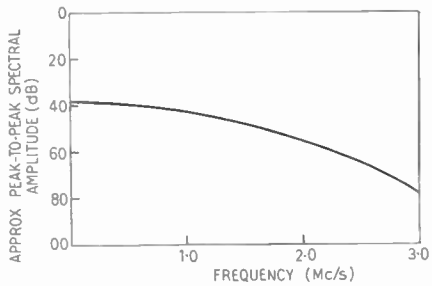


Fig. 10. The envelope of the spectra of a $2T$ pulse consisting of harmonics of the repetition frequency 10 kc/s.

The sine-squared pulse is to have a repetition frequency equal to that of the picture line frequency and the half-amplitude duration of the pulse is to be equal to one half period of the upper cut off frequency of the system.

The characteristics of the sine-squared pulse may thus be defined as follows:

(1) The pulse is finite in duration thereby facilitating evaluation of the test results since no irrelevant information is contained in the received pulse. This is clear when Fig. 10 is considered which shows the line spectra of the $2T$ pulse[6] plotted against frequency, the spectra are harmonics of the fundamental line frequency of 10 kc/s.

(2) The sine-squared pulse conforms to the requirement of a standard test pulse generated by a unit which can be manufactured at a reasonable cost and available to all authorities.

(3) The half-amplitude duration of the fast pulse (T) 0·17 μs corresponds very closely to the smallest element which can be resolved on the 405-line system. (See Section 4.2.)

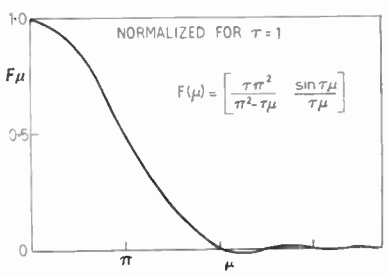


Fig. 11. The spectrum of the sine-squared pulse.

(4) The pulse amplitude is clearly defined between blanking level and peak white, thus the half-amplitude duration can be accurately measured since the half-amplitude level is simple to obtain.

(5) The components of the frequency spectrum relative to the fundamental have a greater amplitude towards the cut-off point than those of a sine-squared step function.

(6) Phase frequency distortion is clearly observed since this results in a direct asymmetry of the pulse about the y axis cutting the peak white point.

(7) The amplitude of the pulse frequency spectrum falls gradually as the frequency increases such that the pulse amplitude is 6 dB down at the nominal cut-off frequency f if the half-amplitude duration $T = f/2$ and reaches zero at f_c if the half-amplitude duration is $2T$. This is clearly observed in Fig. 11 which shows the amplitude/frequency spectrum of the pulse.

(8) For a 405-line television system, a cable link is generally designed and commissioned such that it approximates to the response obtained from a 3·0 Mc/s ideal low pass filter. Thus if a sine-squared pulse is applied to an ideal low pass filter having a cut-off at 3·0 Mc/s then the observed results would be similar to those obtained after transit of an identical pulse over a correctly aligned 3·0 Mc/s television transmission system. Figure 12 shows the response of a 3·0 Mc/s ideal low-pass filter[3] to a sine-squared pulse having unit amplitude and a half-amplitude duration of 0·17 μs.

(9) If a sine-squared pulse is applied to a 3·0 Mc/s ideal low-pass filter, the pulse being of unit amplitude and half-amplitude duration 0·17 μs, then an overshoot is obtained of 13%, this being four times in excess of that obtained when passing a step function into the same filter. Thus it is clear that the sine-squared pulse will be a more sensitive test for ringing and will thus indicate the ringing with a greater degree of accuracy than is possible with a step function. This result is clearly seen in Fig. 12 which illustrates the response of a 3·0 Mc/s ideal low-pass filter to a sine-squared pulse of unit amplitude and half-amplitude duration 0·17 μs.

(10) The mathematical treatment of the sine-squared pulse is as follows:

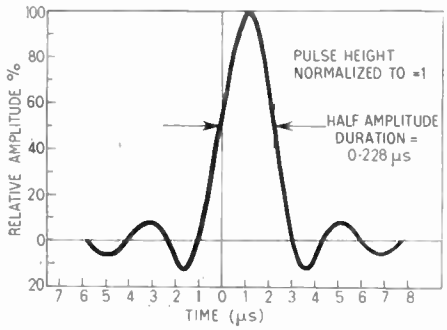


Fig. 12. The response of an ideal low pass (300 Mc/s) filter to a sine-squared pulse of unit amplitude and h.a.d. = 0·17 μs.

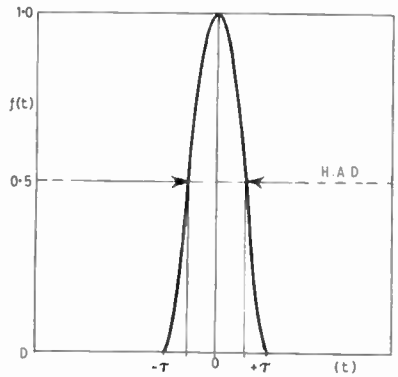


Fig. 13. The sine-squared pulse.

Consider the sine-squared pulse as shown in Fig. 13 symmetrically displaced about the zero y axis. This can be described as

$$f(t) = \sin^2\left(\frac{\pi}{2}\left(\frac{t-\tau}{\tau}\right)\right) \quad \text{for } -\tau < t < \tau \quad \ldots(14)$$

and since

$$f(t) = \cos^2\frac{\pi t}{2\tau}$$

then
$$f(t) = \frac{1}{2}\left(1 + \cos\frac{\pi t}{\tau}\right) \quad \ldots(15)$$

To derive the pulse spectrum:

$$F(\mu) = \int_{-\infty}^{+\infty} f(t)\, e^{-j\mu t}\, dt \quad \ldots(16)$$

$$F(\mu) = \frac{1}{2}\int_{-\tau}^{+\tau}\left[1+\cos\frac{\pi t}{\tau}\right] e^{-j\mu t}\, dt$$

$$= \frac{1}{2}\left[\frac{1}{-j\mu}e^{-j\mu t} + \frac{e^{-j\mu t}}{\left(\frac{\pi^2}{\tau^2}-\mu^2\right)}\left(\frac{\pi}{\tau}\sin\frac{\pi}{\tau}t - j\mu\cos\frac{\pi t}{\tau}\right)\right]_{-\tau}^{+\tau}$$

$$= \frac{1}{2}\left\{\left[\frac{e^{j\mu\tau}-e^{-j\mu\tau}}{j\mu}\right]\left[1+\frac{\mu^2}{\left(\frac{\pi^2}{\tau^2}-\mu^2\right)}\right]\right\}$$

Thus
$$F(\mu) = \frac{\sin\mu\tau}{\mu}\cdot\frac{\left(\frac{\pi^2}{\tau^2}\right)}{\left(\frac{\pi^2}{\tau^2}\right)-\mu^2} \quad \ldots(17)$$

or
$$F(\mu) = \frac{\pi^2}{(\pi^2-\mu^2\tau^2)}\cdot\frac{\sin\mu\tau}{\mu} \quad \ldots(18)$$

Thus the expression for the spectrum of the sine-squared pulse is as shown in (18) and the curve to this equation is illustrated in Fig. 11 which is normalized for $\tau = 1$.

Thus the two expressions for the sine-squared pulse are:

$$\text{Pulse shape} = f(t) = \sin^2\left[\frac{\pi}{2}\left(\frac{t-\tau}{\tau}\right)\right] \quad \ldots(19)$$

$$\text{Pulse spectrum} = F(\mu) = \frac{\pi^2}{(\pi^2-\mu^2\tau^2)}\cdot\frac{\sin\mu\tau}{\mu} \quad \ldots(20)$$

Whilst the half-amplitude duration is shown to be:

From (2)
$$f(t) = \frac{1}{2}\left[\cos\frac{\pi t}{\tau}+1\right] = \frac{1}{2}$$

Thus
$$\cos\frac{\pi t}{\tau} = 0 \quad \text{or} \quad \frac{\tau t}{\tau} = \frac{\pi}{2}$$

Therefore $t = \tau/2$ and the half-amplitude duration $= \tau$.

6. The *K*-Rating System

The *K*-rating system has been introduced by the Post Office Engineering Department for the testing of television links in co-operation with the various users. This seeks to place a limit on all the waveform characteristics required to describe the ability of a link as part of a television system to carry actual television signals. The limits are based, *inter alia*, on subjective tests, the need for economy and a wealth of experience on results obtained to date; the relationship between the limits and subjective picture distortion is readily perceived.

Distortion consisting of a single long-term echo has been adopted as the basis of reference in such a manner that the relative amplitude of this echo is numerically equal to the rating factor. Thus a complete set of performance limits is determined by assigning a particular numerical value to the factor *K*. As a consequence, if *K* is put into a percentage form then the corresponding standard of performance may be conveniently expressed as a percentage rating. It was mentioned previously that the distortion on which the system is based is that due to a single long term echo. A more precise definition of this is:

> The distortion is generally said to be equivalent to that caused by a single long-term echo of amplitude $K\%$ of the original pulse transmitted and distant by more than 4/3 μs from the received pulse.

The rating factor which is related to a subjective depreciation of picture quality as a result of distortion was arrived at by the following method. A party of trained observers compared the various forms of picture impairment with that of a typical ghost image. Different types of waveform distortion that gave the same degree of annoyance as that caused by a single long-term echo were then given the same rating and a set of limits established. It was found in the case of a 3·0 Mc/s system that the results could be equated, such that, for example, a 2% single term echo distant by more than 4/3 μs from the pulse is equivalent to an 8% echo distant by more than 1/3 μs from the originating pulse or an echo of 4% distant by 2/3 μs from the main pulse.

For various forms of distortion it was found that an amplitude of 3% was not objectionable but if the distortion increased to 5% then the distortion became objectionable. From the results stated it is thus possible to construct a graticule in which the received signal must lie within certain engraved limits in order for the signal to comply with the specified *K*-rating tolerance for the item under test.

Considering the complete test signal shown in Fig. 14, it has been found that if the difference between the pulse amplitude and the 40 μs bar amplitude is expressed as a percentage of the pulse amplitude, this difference $x\%$, is equivalent to an impairment caused by a distant echo of amplitude $x/4\%$. This measurement is termed the pulse-to-bar ratio.

The overall *K*-rating of a television system is taken as the worst figure of those obtained from a given series of tests. Links having the same rating should theoretically give the same standard of performance from a subjective point of view although the actual type of inherent impairment may be different.

Thus the values of *K* running from 1% for a good channel to 5% for a poor channel and down to ½% ($K = 0{\cdot}005$) for a simple video channel amplifier will be obtained. For example if the *K* limit in a particular case is specified as 0·03 or 3% and a certain type of distortion is rated as equal to 2*K* then that distortion must not exceed 0·06 or 6%. It has been found that generally with a large number of links working in tandem and having varying degrees of distortion and types, the overall *K*-rating factor tends to approximate to the root sum of the squares of the individual rating factors. A typical target performance suggested for four main links plus nine short links in tandem is $K = 4{\cdot}5\%$.

7. The Waveforms used for the Complete Testing of Television Links

If the types of distortion which generally occur in television line[7] transmission systems or radio links are considered it will be found that they can be classified generally into three approximate time regions associated with the television waveform as follows:

(1) 0–20 ms generally associated with field response.

(2) 0–0·1 ms generally associated with line response.

(3) Transients generally associated with h.f. response.

It is reasonable to say therefore that the linear performance of a 3·0 Mc/s television link can be evaluated if test signals are employed which will explore the distortion present in the three time regions enumerated.

Considering the waveform testing of television links using a sine squared pulse as a part of a composite test waveform then the component parts of the complete test waveform are related to the stipulated time regions as follows:

(a) Region (1) is not directly related to the pulse and bar test waveform but is tested by the application of a further test waveform consisting of a smooth 50 c/s square wave interrupted at 10 kc/s.

(b) Region (2) is related to the smoothed 10 kc/s bar of 40 μs duration.

(c) Region (3) is related to the sine-squared pulses of either 0·17 μs or 0·33 μs half-amplitude duration.

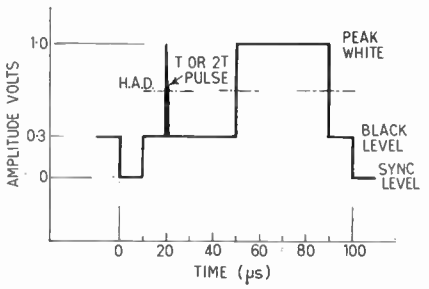


Fig. 14. Pulse and bar test signal (405 line).

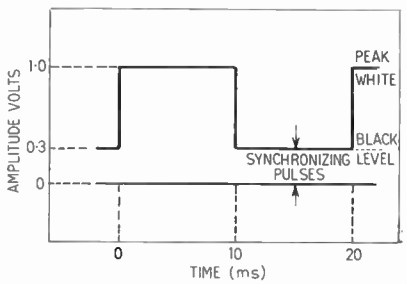


Fig. 15. 50 c/s square wave test signal.

In the standard type of pulse and bar generator signals (b) and (c) are combined to form the composite test waveform which thus enables several parameters of the circuit under test to be observed simultaneously, the required type of pulse being selected by means of a pulse selector switch which places in circuit the correct pulse shaping network. The complete test waveform is shown in Fig. 14, whilst Fig. 15 shows the 50 c/s square-wave test signal.

7.1. *Components of the Pulse and Bar Test Waveform*

(1) *Sync pulses:* Standard line synchronizing pulses are contained in the waveform having an amplitude of 0·3 V, pulse width of 10 μs and a time of rise and fall of better than 0·2 μs. These serve the dual purpose of being presented for evaluation in conjunction with the other components of the test waveform and triggering line amplifying and clamping equipment.

(2) 40 μs *bar:* The 40 μs bar as a repetition frequency of 10 kc/s since it is an integral component of the complete test waveform and a peak white amplitude of 0·7 V. Each bar transition approximates to an integrated sine-squared shape with a rise time of 0·33 μs, this value of rise time being chosen since the bandwidth of the standard system is 3·11 Mc/s. As a consequence no confusing bar overshoots are generated due to the pulse being restricted to 3·0 Mc/s. The bar transitions are determined by the same pulse shaping network as used for pulse shaping.

(3) *Sine-squared pulses:* The two types of pulses used for testing are termed T and $2T$ pulses, their characteristics being such that the former has a half-amplitude duration of 0·17 μs whilst the latter has a half-amplitude duration of 0·33 μs this being the duration of the 50% amplitude point. The pulse amplitude is 0·7 V and the shape is determined by the shaping networks as shown in Fig. 14. It might be noted that $T = \frac{1}{2}f_m$, where f_m is the maximum frequency of the system and for the 405-line system $T = 1/6$ μs.

(4) 50 *c/s square-wave:* The waveform is shown in Fig. 15 and it can be seen that for 10 ms there appears only sync pulses whilst for the other 10 ms the signal is peak white.

8. The General Method of Waveform Measurement to obtain a *K*-Rating

It will be appreciated that in order for the results of the waveform test to be related to a specified limit then a standard means of defining the agreed tolerance or limits of deviation must be employed. The method used in this type of testing is to observe the waveform received from the system under test in conjunction with a suitably engraved graticule, the mask having the waveform limits engraved on its surface in terms of K. One of the several graticules used in the K-rating system is illustrated in Fig. 16. The general method of measurement using this mask is firstly to set up the sweep of the oscilloscope to the correct value. Each unit interval of time on this graticule is equal to $\frac{1}{6}$ μs, thus the sweep speed of the oscilloscope is set to 0·367 μs/cm. Alternatively if the instrument does not have a calibrated time-base then the sweep speed is set up as follows. Applying a 3·0 Mc/s sinusoidal wave to the vertical amplifier of the oscilloscope the scan speed is set so that the positive peaks of the timing wave coincide with the even time markers on the graticule. The graticule in Fig. 16 is engraved for a rating of $K = 4\%$ with inner limits of 2% the various other limits on the mask being shown. Thus to obtain a rating on a pulse it would be fitted into the graticule as specified in the subsequent text and by suitable manipulation of the oscilloscope amplitude and shift controls the requisite K-rating can be derived, enabling a quick assessment of the performance of a television

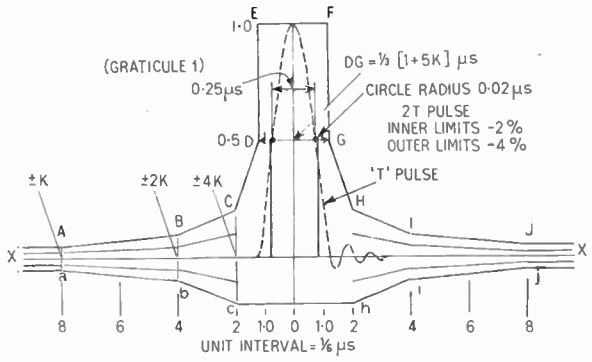


Fig. 16. Waveform response limits to $2T$ pulse test signal.

link to be obtained. By comparing the results obtained at the time of the test with the acceptance test figures a quick comparison of the difference existing is obtained.

9. The Routine Testing of Television Links using the *K*-Rating System

The routine type of test is used for the day-to-day checking of television systems, quick assessment of fault conditions or for the checking of mobile microwave radio link systems. The four signal components used for the routine test are the *T* pulse, 2*T* pulse, 40 μs bar and 50 c/s square wave, thus to meet a specified rating factor the response of the observed signals must be collectively within the limits associated with the factor. The parameters of each of the specified test components are evaluated as follows:

9.1. 2*T* *Pulse Response*

With the oscilloscope sweep speed set up as described in Section 8 the pulse component of the signal is fitted into the graticule shown in Fig. 16, the black level of the signal being coincident with the line XX′. The peak white point of the pulse touches the line EF, the half amplitude points being symetrically displayed about the line DG; the inner limits engraved then give a 2% rating whilst the outer limits give a 4% rating.

Since the link over which the test waveform is being received may tend to introduce lobes or overshoots before and after the pulse, in addition to echoes, it is thus necessary to take these into account when taking a *K*-rating of the pulse. This is achieved by making three effectively separate measurements:

(*a*) With a correctly positioned pulse the half-amplitude duration *K*-rating is given by pulse intercepts on the line DG.

(*b*) Inspect the skirts of the pulse and any lobes between the intervals − 4 and + 4 and take the *K*-rating of the worst distortion by interpolation.

(*c*) Noting the trace outside the limits − 4 and + 4 the *K*-rating of the worst long term echo is taken.

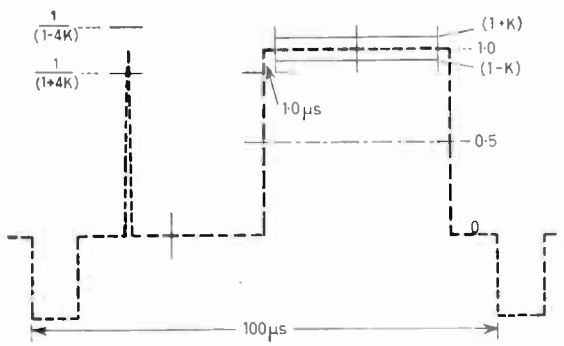


Fig. 17. Waveform response limits to 40 μs bar test signal.

Thus the *K*-rating of the 2*T* pulse will be given by the largest figure noted in the previous tests.

9.2. *Bar-to-2T-Pulse Ratio*

Using the graticule shown in Fig. 16, and having first checked that the pulse is of the correct amplitude, the complete pulse and bar test signal is positioned so that the pulse peak touches the line XX′. The mid-point of the bar should then be found to fall on a section of the mask outline so that the difference in level can be related to the graticule dimensions as follows:

The ordinates of the outer 4% line of the graticule correspond to the following percentage of the pulse height.

AX − Xa	4%	JX − Xj
BX − Xb	8%	IX − Xi
CX − Xc	16%	HX − Xh

The ordinates of the inner 2% line are half these values.

Thus the measurement made when divided by 4 gives the K-rating for this parameter using the ordinates as illustrated.

9.3. 40 μs *Bar Response*

The waveform limits of the bar are as illustrated in Fig. 17. The deviation of the top of the bar from its mid point expressed as a percentage of the bar amplitude is the *K*-rating of the bar, the first and last 1 μs of the bar being neglected since this region is covered by the pulse measurements.

The special graticule for this measurement is shown in Fig. 17 but it is more convenient to use the graticule shown in Fig. 16 for routine testing as this does not involve the changing of graticules. The measurement is made in the following manner:

The bar amplitude is adjusted with the oscilloscope Y gain such that it equals the engraved pulse height XX′–EF the width being set so that the leading edge falls at −6 on the time scale and the trailing edge at −14. The mid-point of the bar top is next set by the Y shift to sit on the axis XX′. The *K*-rating of the bar can thus be obtained directly in the same manner as described in the 2*T* response.

9.4. *T Pulse Response*

It is not possible to obtain a *K*-rating factor by direct observation of the *T* pulse but nevertheless the following features are measured which should fall within the stipulated limits:

9.4.1. Half-amplitude duration

The pulse is displayed and fitted into the graticule as shown in Fig. 16, its half-amplitude duration being measured with respect to the two engraved circles whose centres are 0·250 μs apart and of radius 0·02 μs.

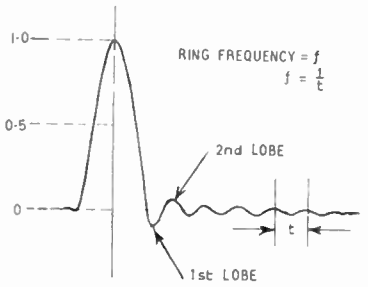


Fig. 18. The measurement of the ring frequency of the T pulse.

For a system with a sharp cut-off at 3·0 Mc/s, the K rating of the T pulse can be extracted from Table 2. For radio links or systems having a bandwidth in excess of 3·0 Mc/s the Table is not representative of the rating since for this type of link or system the overshoots and half-amplitude duration tend to be generally less than those in Table 2.[10]

Table 2

Features	Rating Factors 1%	2%	3%	4%	5%	6%
Half amplitude-duration max μs	0·245	0·250	0·255	0·260	0·265	0·270
Ring frequencies min. Mc/s	3	3	3	3	3	3
1st lobe trailing or leading, max. %	10	12	14	16	18	20
2nd lobe negative trailing or leading, max. %	6	8	9	10	11	12

9.4.2. Pulse lobes

With the pulse set up as shown in Fig. 16, the amplitude of the first two negative and positive lobes before and after the pulse are measured. These are expressed as a percentage of the pulse height.

9.4.3. Ringing

The ring frequency of a television system should be greater than 3·0 Mc/s. Ringing shows up as damped oscillation after the trailing edge of the T pulse as illustrated in Fig. 18, the frequency of ringing being measured three complete cycles past the main pulse lobe; the accepted limits for ringing are as enumerated in Table 3.[9] It might be noted that since the spectrum of the T pulse is wider than 3·0 Mc/s then clearly the response of equipment and systems will be dependent on the upper frequency limit. Thus it is likely that the received test results will deviate somewhat, in certain cases, from those laid down in Tables 2 and 3.

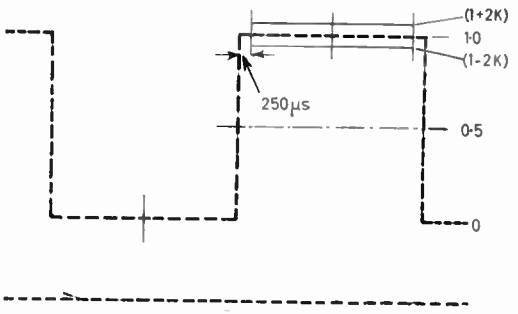


Fig. 19. Waveform response limits to 50 c/s square wave test signal.

Table 3

Maximum amplitude of ringing expressed as a percentage of T pulse amplitude

System rating factor K	1st lobe negative leading or trailing %	2nd lobe positive leading or trailing %
0·01	10	6
0·02	12	8
0·03	14	9
0·04	16	10
0·05	18	11
0·06	20	12

9.4.4. 50 c/s square wave test

The waveform response limits of the 50 c/s square wave test are as shown in Fig. 19, this being the special graticule engraved for this test. However, using graticule 1 as shown in Fig. 16 (for ease of use), the oscilloscope is set up to display one cycle of the test waveform, the gain of the instrument being adjusted so that the waveform lies between the limits XX′ and EF on the graticule. The sweep speed is next set to position the width of the bar half-amplitude points between the time limits −6 and +14 with the half-amplitude points coincident with the line XX′. The trace is then shifted so that the bar peak white level lies on the engraved line XX′. The maximum deviation over the half cycle under examination can then be given a K rating as in the 40 μs bar response.

10. The Acceptance Testing of Television Links using the K-Rating System

The test waveform is used for the commissioning and acceptance of new coaxial line television links. The principle component of this test is the T pulse whilst the 40 μs bar and 50 c/s square wave are dealt with as in the routine testing method.

The T pulse response of the circuit under test is photographed with a sinusoidal timing wave. The resultant print is then set up on a measuring microscope as shown in Fig. 20 so that the axes of the pulse response are in alignment with those of the microscope co-ordinates.

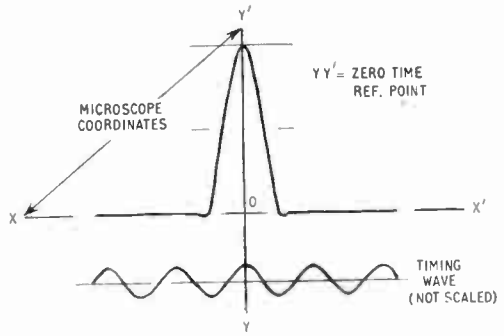


Fig. 20. The sampling of a T pulse to obtain a time series.

The ordinate YY′ represents the zero time reference point the timing wave being used to give a series of points spaced at $\frac{1}{12}$ μs intervals either side of the reference point. At each of these points which is indicated by a positive peak of the time marker, the Y amplitude of the pulse waveform is measured.

The technique described is equivalent to the sampling of the waveform in a time-division multiplex system and the time series obtained completely describes the waveform providing its components are limited to 6·0 Mc/s. This is complied with by virtue of the characteristics of the T pulse and expressing the waveform response in terms of a time series results in an expression particularly suitable for arithmetic manipulation. The principle of the method used for measurement is to obtain the filtered impulse response of the system under test, this being the response of the system to a pulse whose shape is given by

$$\sin\left(\frac{\pi t}{T}\right)\Big/\left(\frac{\pi t}{T}\right)$$

Its waveform[6] is shown in Fig. 21 and its spectrum is uniform in amplitude up to a frequency $T/2$ and zero at higher frequencies.

However, as a 3·0 Mc/s low-pass filter is involved in this measurement for bandwidth restriction then clearly difficulty will be encountered since an ideal 3·0 Mc/s low-pass filter is not practically obtainable and as a consequence the system will not be ideally restricted to 3·0 Mc/s.

This difficulty is overcome by taking the product of the time series of the link and of an ideal 3·0 Mc/s low-pass filter the resultant series having ordinates spaced at $\frac{1}{6}$ μs intervals and representing the tandem response of link and filter. Both the test equipment and link under test are subjected to the measurement and filtration process. Thus if the quotient of the time series of the link and test equipment is taken then the final series obtained which is the filtered impulse response of the system is not modified in any way by defects in the characteristics of the test equipment. The series representing the filtered impulse response of the system contains all the information about the distortion introduced by the system up to the maximum frequency of interest and the two results obtained will indicate either distortionless or distorted transmission.

(*a*) *Distortionless transmission*

If the link were distortionless then the filtered impulse response would be sin $(\pi t/T)/(\pi t/T)$ and the time series representing this would contain only one term, that of distortionless transmission. This is due to the fact that at each sampling interval in the final time series at $\frac{1}{6}$ μs spacing, the ringing associated with the ideal low pass filter is passing through zero as shown in Fig. 21. Thus the only non-zero sample is that of the time reference point of the pulse (Fig. 21) which coincides with the peak of the pulse.

(*b*) *Distorted transmission*

When there is inherent distortion in the system under test then clearly there will be further terms present in the final time series and each represents the amplitude of pulse sin $(\pi t/T)/(\pi t/T)$ whose peak occurs at a time appropriate to the term in question. The series thus contains a main term representing undistorted transmission plus a number of terms representing distorted transmission spaced at $\frac{1}{6}$ μs intervals either side of the main term, the amplitude of each distortion term being the amplitude of an echo of the main term occurring at the appropriate time. The echoes do not necessarily need to appear in symmetrical pairs about the main term since they do not represent attenuation or phase distortion considered separately.

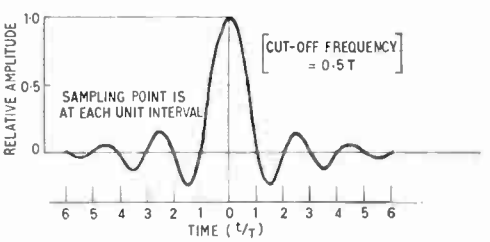


Fig. 21. The impulse response of an ideal low-pass filter.

10.1. *The Acceptance Rating Factor*

To obtain an acceptance rating factor a further time series is formed from the product of the filtered impulse response of the link and the series $(\frac{1}{2} : 1 : \frac{1}{2})$ which is a simple approximation to a $2T$ sine-squared pulse. Thus the final time series represents approximately the $2T$ pulse response of the system under test calculated from the T pulse response. The restrictions[10] placed on the system test results and their equivalence to the test graticules are as follows:

Let the filtered impulse response of a system be represented by the following normalized time series:

$$B(rT) = \ldots B_{-r} \ldots B_{-1} . B_0 . B_{+1} \ldots B_{+r} \quad \ldots(21)$$

Assuming that this has been normalized making $B_0 = 1$.

Let the serial product of $B(rT)$ and the series $(\frac{1}{2} . 1 . \frac{1}{2})$ be given by

$$C(rT) = \ldots C_{-r} \ldots C_{-1} . C_0 . C_{+1} \ldots C_{+r} \ldots(22)$$

where $C_r = \frac{1}{2}B_{r-1} + B_r + \frac{1}{2}B_{r+1}$

The main term of this series represents the distortionless characteristic of the link whilst the echo terms represent the distorting characteristic. To meet a specified rating factor $K\%$, the amplitudes of the echo terms must meet the following four restrictions, each representing a limit on a different property of the waveform response and yielding a different factor:

Restriction (1)

$$\tfrac{1}{8}(C_r/C_0 . -\tfrac{1}{2}) \leqslant K \quad r = -1$$

and

$$\tfrac{1}{8}(rC_r/C_0) \leqslant K \quad \begin{cases} -8 \leqslant r \leqslant -2 \\ +2 \leqslant r \leqslant +8 \end{cases}$$

$$(C_r/C_0) \leqslant K \quad \begin{cases} r < -8 \\ +8 < r \end{cases}$$

This is approximately equivalent to the limits indicated on the graticule shown in Fig. 16 for the $2T$ response.

Restriction (2)

$$\tfrac{1}{4}\left[1/C_0\left(\sum_{-8}^{+8} B_r\right) - 1\right] \leqslant K$$

This is approximately equivalent to the limits placed on the $2T$ bar/pulse amplitude ratio in the routine test method.

Restriction (3)

$$\tfrac{1}{6}\left(\sum_{-8}^{+8} B_r - 1\right) \leqslant K$$

This is equivalent to the limits placed on the bar/pulse amplitude ratio of the response to a hypothetical pulse and bar signal in which the pulse is an ideal filtered impulse.

Restriction (4)

$$\tfrac{1}{20}\left(\sum_{-8}^{+8} (B_r) - 1\right) \leqslant K$$

This is an upper limit placed on the average amplitude ignoring signs of the 16 central echo terms of the filtered impulse response time series.

Restrictions (1) and (2) have close equivalents in the routine test method whilst restriction (3) takes into account distortion near the upper frequency limit of the system under test. In the routine test method this type of distortion will affect the half-amplitude duration of the T pulse response and/or the initial amplitude of the ringing. Restriction (4) takes into account excessive ringing, a low system cut-off point and/or a long train of echoes whose magnitudes are not individually great enough to reach one of the other limits. Thus to obtain a single overall rating factor for the link the largest value of K obtained for the restrictions, the bar response or the 50 c/s square wave, is taken.

11. The Results to be Expected from the Waveform Testing of Television Systems

In television links and equipment there is general pattern of faults which normally occur, these being attributable to the two main characteristics of gain/frequency and phase/frequency which are of course interrelated. By interpretation of the results obtained from waveform testing a general pattern exists whereby the deformation of the pulse-and-bar test signal can be related to a particular subjective effect. A general relationship between the various components of the pulse and bar test signal and the subjective effect on a television picture is as follows.

11.1. *Tests over Linear Systems*

40 μs *bar*

Useful test range 10 kc/s to approaching 0·5 Mc/s.

(1) Distortion of the peak white portion of the bar is generally indicative of long term distortion in the lower region of the video band.

(2) Bar slope is indicative of l.f. loss and distortion.

(3) Rounding of the bar transitions is generally indicative of poor response in the h.f. region of the spectrum.

2T pulse

Useful test range 0·5 Mc/s to approximately just greater than 2·0 Mc/s.

(1) H.a.d. indicates the general response over the mid video band a large h.a.d. indicating poor rise time for pulses.

(2) Asymmetry of the pulse about its Y axis will indicate phase distortion over the appropriate region of the video spectrum.

(3) Pulse lobes and echoes indicate undershoots and overshoots.

T pulse

Useful test range greater than 2·0 Mc/s.

(1) H.a.d. indicates general response to greater than 3·0 Mc/s.

(2) Ringing shows up as a damped oscillation following the pulse.

(3) Shows up phase distortion as per the $2T$ pulse.

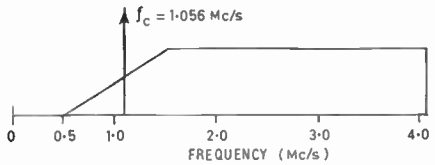


Fig. 22. C.C.I.T.T. frequency band for 405 line television land cables.

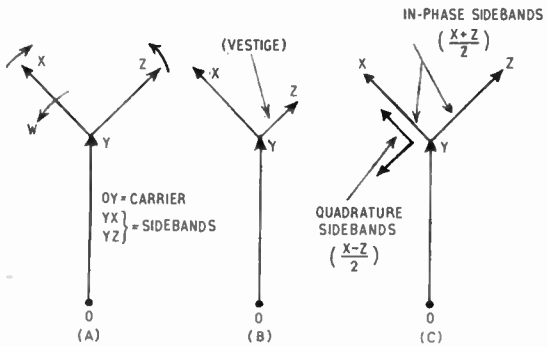


Fig. 23. Vector diagrams of a modulated wave showing
(*a*) Conventional diagram (double sideband).
(*b*) Shortened sideband.
(*c*) Resultant of vestigial sideband transmission.

Pulse-to-bar ratio

In the case of the 2*T* pulse this is an indication of the frequency differential response between 10 kc/s and approximately 2·0 Mc/s.

11.2. *Tests over Systems Subject to Non-linearity (Vestigial Sideband Cable Links)*

On long distance cable links television signals are transmitted as a vestigial-sideband signal according to C.C.I.T.T. specification. This recommends that television signals shall be transmitted as the upper sideband of a 1·056 Mc/s carrier frequency with a vestigial lower sideband extending down to approximately 0·5 Mc/s, and is illustrated in Fig. 22.

In the vestigial sideband transmission system, the video signal is modulated by a suitable carrier frequency and the resultant signal passed through a sideband-shaping filter, which passes one complete sideband and a vestige of the other sideband; this signal is then modulated by the final carrier for transmission to line. The reduction in amplitude of the one sideband, however, results in a form of distortion inherent in vestigial sideband systems, namely, quadrature distortion. Clearly, this distortion is only present when the system is not working with correct values of modulation index, carrier frequency level and sideband vestige.

The origination of quadrature distortion is illustrated in vector form in Figs. 23(*a*) and (*b*), the former showing the vector representation of the carrier and sidebands in a double sideband system. The carrier vector OY rotates about O with angular velocity ω and the modulating frequency is represented by the two equal vectors YX and YZ, rotating about the point Y with equal and opposite angular velocity. At any one instant these vectors are of equal magnitude and form equal angles with vector OY, thus their components at 90 deg to OY are always equal and opposite and therefore cancel out.

However, when one sideband is reduced as in vestigial sideband transmission, this results in an additional set of sidebands being formed, whose resultant is in quadrature with the original carrier, as shown, in Fig. 23(*b*). Another term for quadrature distortion is skew-symmetrical distortion, an example of this being a sinusoidal variation of amplitude frequency response about the carrier frequency. In the impulse response a pair of echoes appear modulated on a carrier in quadrature with the original, the leading echo being positive whilst the trailing echo is negative; the in-phase modulated waveform is undistorted. Since the result of quadrature distortion is to distort the carrier envelope voltage with some phase modulation of the carrier, then clearly, envelope detection will certainly follow any inherent quadrature distortion. The degree of the distortion is dependent on the modulation index and width of the vestige.

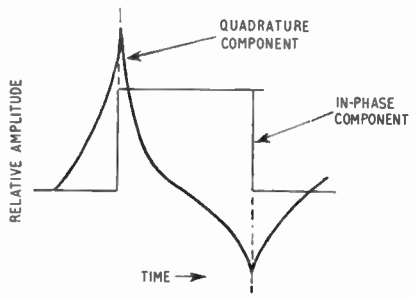


Fig. 24. The effect of quadrature distortion on a video square wave.

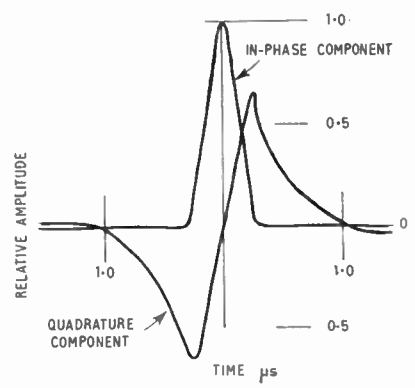


Fig. 25. The in-phase and quadrature modulating components for a 2*T* pulse in a 405 line transmission system with linear shaping to give a vestigial sideband.

Generally with this type of system, if the modulation depth is kept below 40% the amount of quadrature distortion affecting the transient response is negligible.

The effect of quadrature distortion on a rectangular video waveform after passing through a system with inherent quadrature distortion is shown in Fig. 24, which illustrates the in-phase and quadrature components.

A further example to illustrate the in-phase and quadrature modulating components of a signal is shown in Fig. 25. Here the components of a $2T$ pulse are illustrated for a 3·0 Mc/s 405-line system [11,12]; it is assumed that the vestigial response is linear, rising from zero at 0·5 Mc/s below carrier to unity 0·5 Mc/s above the carrier.

Now, a vestigial sideband system having a high modulation index is strictly speaking a non-linear system, and as such, the K-rating system will not hold a definite relationship so far as the pulse half-amplitude duration is concerned, since the pulse duration increases as the vestige and modulation index increases. However, if the test is preceded by a non-linearity check of the system using a sawtooth test waveform and the degree of non-linearity present is very slight, then a K-rating can still be obtained by relaxing the limits on the rating of the pulse half amplitude duration.

Generally, the effect of transmission over vestigial sideband systems is related to the manner of modulation used in the system and it is found that with positive modulation the pulse half amplitude duration is increased, giving a triangulated effect. The 40 μs bar is affected and positive going overshoots are observed on both bar edges above the peak white limit. The magnitude of the effects mentioned is dependent on the degree of modulation and width of vestige and affects a T pulse to a greater extent.

In a system having negative modulation the pulse half-amplitude duration tends to decrease together with a triangulated effect and negative going overshoots are observed on the 40 μs bar edges below black level.

To sum up; the use of pulse and bar tests over vestigial sideband links is valid below system modulation depths of about not greater than 40% but above this figure of modulation the pulse half amplitude duration rating has to be relaxed in light of the amount of non-linearity present in the system. Clearly, if a large amount of non-linearity is observed on a system after transmission of a sawtooth test waveform over it, then the K-rating assessment will not be valid.

12. The Principle of Pulse Generation

If a short but finite pulse is applied to the input terminals of a shaping network then the shape of the output pulse is determined only by the driving pulse. Thus in order to obtain a pulse whose shape and duration are independent of the driving pulse and only dependent on the characteristics of the network, the driving pulse should approximate to an ideal impulse.

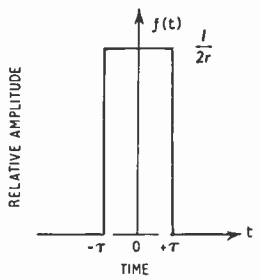


Fig. 26. Unit impulse.

An ideal impulse is one having a very short duration and large amplitude such that the product of amplitude and duration is unity. This can be clearly seen if the following treatment of an impulse function is examined.

Consider an impulse as shown in Fig. 26.

Dirac function is $\lim_{\tau \to 0} f(t)$

$$f(t) = 0. \quad \text{When } t < -\tau$$

$$f(t) = \frac{1}{2\tau} \quad -\tau < t < \tau$$

$$f(t) = 0 \quad t > \tau$$

Thus
$$F(\mu) = \lim_{\tau \to 0} \int_{-\infty}^{\infty} f(t)\, e^{-j\mu t}\, dt$$

$$= \lim_{\tau \to 0} \int_{-\tau}^{\tau} \frac{1}{2\tau}\, e^{-j\mu t}\, dt$$

$$= \lim_{\tau \to 0} \frac{1}{2}\left[-\frac{1}{j\mu\tau} (e^{-j\mu\tau} - e^{j\mu\tau}) \right]$$

Therefore
$$F(\mu) = \lim_{\tau \to 0} \left\{ \frac{\sin \mu\tau}{\mu\tau} \right\} = 1$$

It can be seen therefore that when τ approaches zero, which is the condition of an ideal unit impulse, the integral of the expression $F(\mu)$ becomes unity.

The impulse response of a network may be defined as the response of the network to an infinitesimally short pulse occurring at time zero and having an amplitude time integral of unity. The sine-squared pulse used for the testing of television systems is formed by feeding a standard drive pulse into a network whose impulse response is a good approximation to the sine-squared shape. To obtain a sine-squared shaped pulse of half-amplitude duration 0·17 μs a

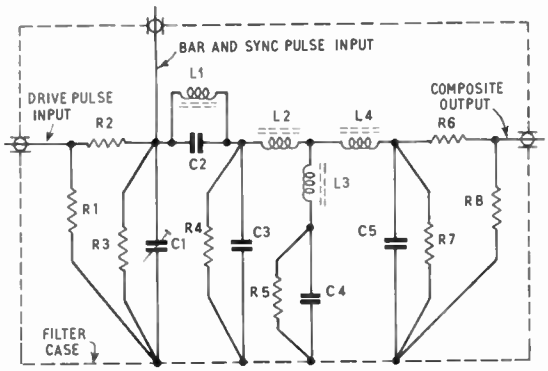


Fig. 27. 0·17 μs sine-squared shaping network.

drive pulse in the order of 0·03 μs is applied to the network[10] shown in Fig. 27.

By virtue of its small pulse width the drive pulse approximates to an ideal impulse when applied to the terminals of the network thereby producing a sine-squared pulse whose characteristics are only dependent on the component constants of the network.

13. The Test Waveform Generator

The sine-squared pulse which is an integral part of the composite test waveform is generated by the pulse and bar generator which is shown in block form in Fig. 28. The composite waveform generated is shown in Fig. 14 and from this it can be seen that the transitions of the waveform are spaced in multiples of 10 μs and the basic timing frequency is 100 kc/s.

V1 generates the 100 kc/s pulses and acts as a class C oscillator at whose anode appears a high peaked waveform. This is used as the trigger pulse to fire the 100 kc/s blocking oscillator stage V2, which is normally biased off.

V2 controls from its cathode, via diode gates, the generators of the three component parts of the composite test waveform which are line sync pulses, 40 μs bar and the sine-squared pulses. To open the diode gates at the required phases of the 10 kc/s waveform a cathode output of V2 is fed via a diode gate to the 10 kc/s divider blocking oscillator V3a whose output is then fed through its buffer stage V3b to trigger the 10 kc/s wide pulse blocking oscillator V4.

The output of V4 feeds the 100 μs delay line which is tapped at slightly less than 10, 20, 30, 40, 70 and 90 μs; these six tappings are the outputs to the six diode gates which are opened and closed in the following sequence.

Assuming the 100 kc/s wide pulse blocking oscillator V4 to have just been triggered, a positive pulse of nearly 10 μs width is fed into the delay line corresponding zero time delay. After a nominal 10 μs delay, diode gate No. 2 is opened, thereby allowing the next pulse from V2 to trigger the negative-going excursion of the bi-stable bar generator V6 and forming the end of the 40 μs bar. At 20 μs No. 3 gate is opened which allows V2 to trigger the negative-going excursion of the bi-stable sync pulse generator V5 thus forming the commencement of the 10 μs sync pulses; 10 μs after this the positive excursion of V5 is triggered thereby terminating the sync pulse.

At 40 μs the impulse trigger stage V10 is gated through No. 5 gate which in turn fires the blocking oscillator V11 the output of which is fed to one of the two pulse shaping networks which produce the sine-squared T or $2T$ pulses.

At 70 μs No. 1 gate is opened and a negative-going pulse is applied to V6 the bar generator which is triggered on its positive-going excursion thus completing the 40 μs bar. Finally at 90 μs a negative-going pulse is applied to No. 6 gate which is opened and applied to the 100 kc/s blocking oscillator V2. The ninth pulse from this stage is thereby suppressed to

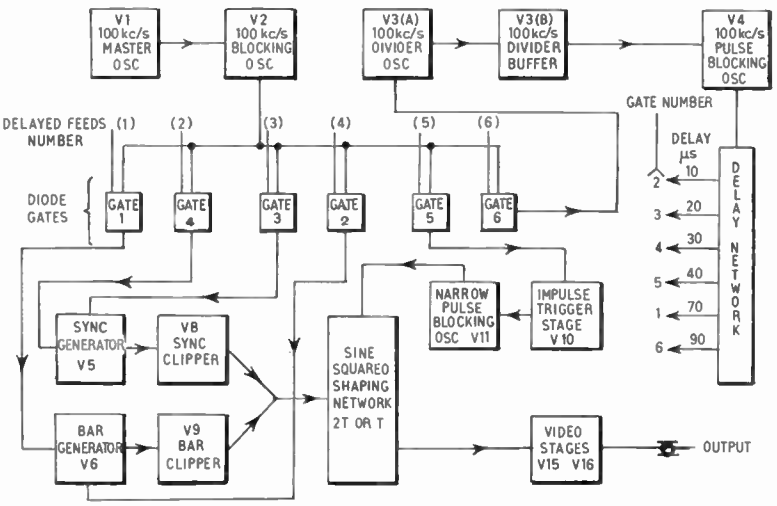


Fig. 28. Pulse-and-bar generator block schematic.

ensure that the self-oscillatory stage V3 is triggered by the tenth pulse 100 μs after the previous pulse from V3. V3 is thus synchronized to one tenth of the frequency of the master oscillator V1.

Thus the cycle of events 100 μs in duration is repeated at a repetition rate of 10 kc/s. The delay line which is quoted as 100 μs long is in actual fact 50 μs long, the 90 μs pulse being the reflection at the 10 μs tapping point of the positive-going pulse which passed that point 80 μs previously.

The sync and bar square waves which are originated by V5 and V6 respectively are clipped by V8 and V9 which are long tailed pairs, V8 is the sync clipper and V9 the bar clipper; V8 has its bias accurately maintained by a shunt stabilizer V7. The two square waves are next passed into the pulse-shaping network, where the transitions of these are shaped. The square waves are then combined with the output of V11 which is the sine-squared pulse to give the composite test waveform output of Fig. 14. The output of the shaping network is then passed to a three-stage video amplifier V15a, V15b, V16 whose h.t. line is shunted by the video amplifier shunt ripple remover, V14. The output of the last stage then appears at the output socket.

14. The Utilization of Pulse and Bar Testing in a Television Station

The following examples serve to illustrate the variety of applications to which waveform testing can be applied in a modern television centre.

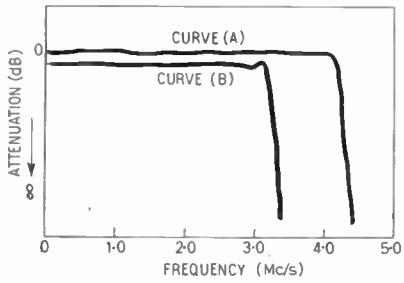


Fig. 29. Attenuation/frequency response of a practical linear phase low pass filter.

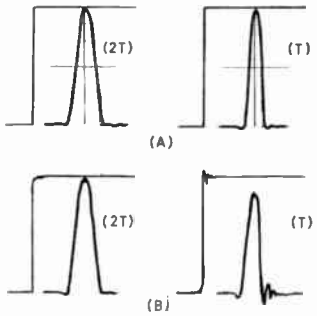


Fig. 30. The waveform response of a linear phase low pass filter
(*a*) Correct filter alignment.
(*b*) Filter mis-alignment to give sharper cut-off.

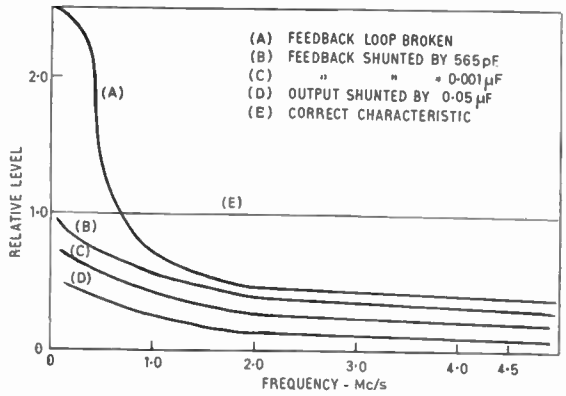


Fig. 31. Gain/frequency response of test amplifier.

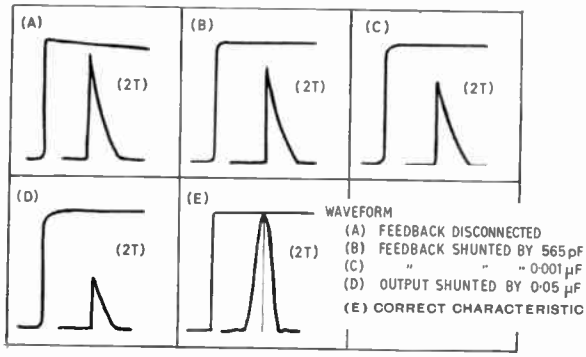


Fig. 32. The 2*T* pulse and bar response of the test video amplifier.

14.1. *Filter Testing*

The attenuation frequency characteristic of a linear-phase low-pass filter is shown in Fig. 29 having a design fall-off point of approximately 4·25 Mc/s; curve (A) is the characteristic of a correctly aligned filter whilst curve (B) is the characteristic of a mis-tuned filter. The waveform response obtained from the same two filters is shown in Fig. 30 from which it can be seen that there is ringing on the *T* pulse from the mis-tuned filter which is clearly a function of the filter cut-off point. In addition a slight amount of phase distortion is present on both *T* and 2*T* pulse due to the mis-alignment.

14.2. *The Testing of Pulse and Vision Distribution Amplifiers*

In a television station numbers of multi-stage video frequency amplifiers are used for the multiple distribution of pulse and vision signals. These amplifiers typically have an output impedance of 75 ohms with a gain of the order of 6 dB and there are up to four or more outlets per amplifier. A quick assessment of the response of these amplifiers can be obtained by using a 2*T* pulse. The attenuation frequency characteristic of an amplifier under test with introduced abnormalities is shown in Fig. 31.

The associated waveform response of the amplifier relative to the stated abnormalities using a $2T$ pulse for testing is illustrated in Fig. 32. With the feedback loop disconnected, the l.f. gain rises, resulting in a poor pulse to bar ratio and phase distortion as indicated by asymmetry of the pulse. With shunt capacitance added to the feedback network as in (B) and (C) the pulse-to-bar ratio becomes worse. As the capacitance decreases the h.f. signal is shunted and phase shift introduced as indicated by asymmetry of the pulse. With the amplifier output shunted by a capacitor the observed results are as shown in waveform (D); the pulse component is attenuated and the bar leading edge distorted. The normal working characteristics are shown in waveform (E).

14.3. *The Measurement of Studio-Transmitter Link Characteristics*

The video circuit between a particular television studio and its associated transmitter 30 radial miles distant, consists of approximately 0·7 miles of standard $\frac{3}{8}$ in. diameter coaxial tube to the repeater station, followed by a further 6 miles of identical tube feeding the transmit terminal of the radio link, which then completes the signal path to the transmitter station. The signal is carried at video frequency over the coaxial tubes and then modulates the transmitter of the microwave radio link via suitable frequency translation equipment. The layout of the signal path and routing is shown in Fig. 33.

In the tests carried out between the studios and transmitter it has been found that the $2T$ pulse-and-bar test is a very good pre-transmission check which enables the control engineer at the transmitter site to assess the condition of the signal path very quickly. This check is normally preceded by a sawtooth test waveform which indicates if any non-linearity is present in the system.

More precise checks using the pulse-and-bar method are carried out weekly and K ratings taken of all the signal parameters. Typical figures and waveform test results are shown in Fig. 34 which illustrates the waveform received at the transmitter when a waveform test is being carried out from the studio centre over a correctly aligned signal path.

15. Conclusions

The paper has illustrated with mathematical examples the advantages of waveform testing as compared to steady state testing when applied to the assessment of the linear distortion present in a television link.

The waveform testing of television links by the utilization of a composite test waveform, one of whose components is a sine squared pulse, results in the characteristics of the system under test being presented

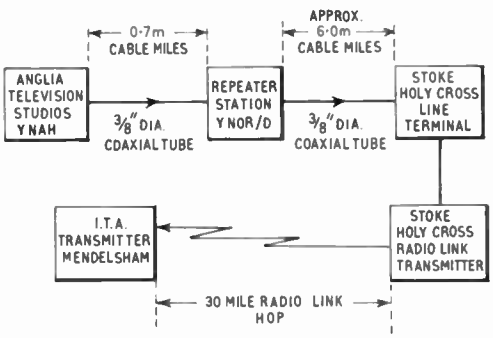


Fig. 33. Routing of video signal path from studies to associated transmitter.

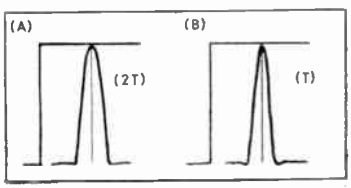


Fig. 34. Pulse and bar waveform resulting from studies to associated transmitter (Norwich studios to Mendelsham).

in such a manner that the result is related to the impairment of the received picture. In other words, a subjective result has been obtained.

The use of waveform testing applied to other features of television engineering has been illustrated with the practical results obtained. It must be stressed however, that these are purely experimental and are naturally dependent on the degree of inherent non-linear distortion present in the item under test and also on the cut-off point. The practical results do, nevertheless, give a rough guide to the waveform which might be expected from a given set of conditions.

It might be pointed out that the $2T$ response is the most useful test to the television engineer as this is operative even in the presence of a certain degree of non-linear distortion. On the other hand the acceptance test method is dependent on the calculation of the step response and the $2T$ response from the measurements on the T pulse and clearly will be affected the most by the existence of non-linearity. Calculations from the acceptance test method are only operative for a linear system.

In connection with the non-linear distortion it might be mentioned that the method employed at the present moment to assess the degree of non-linearity is by the transmission of a sawtooth test signal, whilst signal crushing is detected by the transmission of a variable lift signal. This method is however being superseded by the utilization of step waveform testing with a given number of steps in the sweep, this being a much superior test of non-linearity.

It has thus been shown that the amount of linear distortion present in a given system can be expressed in terms of K, this method of expression being such that it can be universally interpreted. Furthermore the measurement of distortion by the K rating method is in itself a simple task which does not call for complex and expensive test equipment.

It is hoped, therefore, that from this paper engineers interested in television transmission will obtain a lucid appreciation of waveform testing techniques and their application to problems of a varying nature in the television transmission field.

16. Acknowledgments

The author would like to thank the many engineers who gave advice on the preparation of this paper and to make acknowledgment to the Post Office Engineering Department for permission to make use of certain information and to also acknowledge all the other sources of information and assistance.

17. References

1. A. H. Roche and L. E. Weaver, "Television frequency translating equipment for the Birmingham-Holme Moss coaxial cable", *Proc. Instn Elect. Engrs*, **99**, Part 3A, page 455, 1952.
2. R. J. Halsey and H. Williams, "The Birmingham-Manchester Holme Moss television cable system", *Proc. Instn Elect. Engrs*, **99**, Part 3A, page 398, 1952.
3. I. F. Macdiarmid, "A test pulse for television links", *Proc. Instn Elect. Engrs*, **99**, Part 3A, page 456, 1952.
4. S. W. Amos and D. C. Birkinshaw, "Television Engineering", Vol. 1, page 273 (Iliffe, London, 1953).
5. I. F. Macdiarmid, "Waveform distortion in television links", *J. Brit.I.R.E.*, **20**, No. 3, page 201, March 1960.
6. I. F. Macdiarmid, "Waveform distortion in television links", *P.O. Elect. Engrs J.*, **52**, Part 2, page 108, Part 3, page 188, 1959.
7. A. R. A. Rendall, "Waveform testing methods for television links", *Electronic and Radio Engineers*, **34**, No. 12, page 451, December 1957.
8. N. W. Lewis, "Waveform computations by the time series method", *Proc. Instn Elect. Engrs*, **99**, Part 3, page 294, September 1952.
9. Pye Telecommunications Ltd., "Pulse and bar waveform generator technical handbook", PTC 1201.
10. "The waveform transmission performance of television links and equipment for 405 lines 3 Mc/s monochrome systems", P.O. Engineering Department, Specification RC 918.
11. H. Mumford, "Some aspects of television transmission over long distance cable links", *J. Brit.I.R.E.*, **19**, No. 8. page 509, August 1959.
12. B. W. Osborne, "Picture quality control equipment for wired television networks", *Proc. Soc. Relay Engrs*, **5**, No. 4, page 89, 1961.
13. G. J. Hunt and E. W. Elliott, "An introduction to the sine-squared pulse", *J. Television Soc.*, **7**, page 49, April–June 1953.
14. A. T. Starr, "Electronics", Appendix 5, page 361 (Pitman, London, 1959).

Manuscript first received by the Institution on 6th November 1961 and in final form on 20th March 1962 (Paper No. 807/T21).

U.D.C. 534.88: 621.396.677

The Directional Discrimination of Volume Arrays of Electro-acoustic Transducers

By

J. W. HORTON, Sc.D.†

Presented at the Symposium on "Sonar Systems" in Birmingham on 9th-11th July 1962.

Summary: The directivity factor of an array of electro-acoustic transducers distributed throughout a cubical volume has been found to vary inversely as the square of the element spacing, as does the factor for the elements occupying one face of the cube. Arrays occupying the volume of a sphere, its surface, and its plane projection show directivity factors which are not inversely proportional to the number of elements but which are nearly equal. At spacings for which these arrays have useful directivity patterns there is negligible advantage in using a volume array in an isotropic noise field.

1. The Evaluation of Directional Discrimination

The magnitude of the directional discrimination of an electro-acoustic transducer, or array of transducers, against acoustic interference reaching it along bearings other than its primary maximum response bearing is expressed by its directivity factor. This is the ratio of the power per unit frequency band of the electric energy generated in each of two systems receiving plane waves of acoustic energy having an intensity per unit frequency band which is the same for all bearings of both systems. One system is the directional transducer, or array, in question; the other is a hypothetical system having a response on any bearing equal to the primary maximum response of the actual system.

The directivity factor, η_D, is computed, in general, by integrating the relative response of the system over all bearings. This relative response, for any bearing, is measured as the ratio of the power of the electric energy generated in response to plane sinusoidal acoustic waves of given frequency and intensity received over that bearing to the power of the electric energy generated in response to plane sinusoidal acoustic waves of the same frequency and intensity received over the primary maximum response bearing.

$$\eta_D = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}(P_{\theta,\phi}/P_0)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$$

Here the angles θ and ϕ are the co-latitude angle and the azimuth angle giving the direction of the bearing at which the generated power is $P_{\theta,\phi}$ relative to the bearing of maximum response for which the generated power is P_0. (These angles are shown by the insert accompanying Fig. 2.)

† U.S. Navy Underwater Sound Laboratory, Fort Trumbull, New London, Connecticut.

In practice directional discrimination is usually expressed as a transmission loss, measured in decibels. This loss is known as the directivity index. Its numerical value is computed as the logarithm based on the tenth root of ten of the reciprocal of the directivity factor.

2. Nominal Relation Between Directivity Index and Size

It is frequently said that the directivity index of an array of non-directional sonar transducers is equal to the logarithm, based on the tenth root of ten, of the number of array elements. This is known to be true under some conditions. The question is repeatedly raised as to whether it would be true if these elements were distributed throughout a volume, as well as when they are located on a surface or along a line. If a large number of array elements could be packed into a volume and retain acceptable directional characteristics the overall dimensions needed for some specified performance would be greatly reduced. It appears desirable, therefore, to examine in some detail the nature of the directivity patterns and the magnitudes of the directivity indices of arrays having elements distributed throughout a volume.

Before looking at these volume arrays, however, it may be well to review some of the characteristics of more familiar forms. Of these the simplest is the linear array of uniformly spaced elements. Values of the directivity indices for linear arrays having $n = 2$ and $n = 5$ elements and for a continuous line of uniform sensitivity are shown graphically in Fig. 1. Values of the index are plotted against the length of the array measured in cycles. Letting the geometric length of the array be l, the values used for the abscissa are for $l/\lambda = (lf)/c$ cycles. They are, in other words proportional either to the geometric length or to the frequency, the constant of proportionality being

the measure of the velocity of propagation of the acoustic waves in question. On these curves points at which the element spacing is one half-wavelength are shown as indicated.

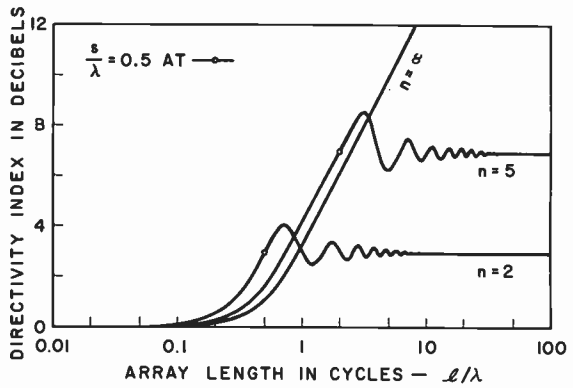


Fig. 1. The directivity indices of linear arrays. This shows the relation between number of elements, size, and frequency.

From these curves it is seen that the directivity index, considered as a function of frequency, does not reach a value equal to the logarithm of the number of elements until the frequency reaches the value for which the element spacing is one half-wavelength. Below this frequency the directivity factor varies inversely as the frequency and the directivity index characteristic has a slope of 3 dB/octave. As frequency is increased above the half-wavelength value the directivity index varies about the logarithm of the number of elements, passing through the value of this logarithm at each harmonic of the half-wavelength frequency. Deviations from this value diminish in amplitude as frequency increases. For spacings greater than five wavelengths these deviations are less than 0·2 dB.

For a linear array of elements having a spacing of five wavelengths, or greater, the directivity index is, indeed, approximately equal to the logarithm of the number of array elements. Below the half-wavelength frequency an increase in the number of elements which reduces the spacing but does not increase the array length will cause a reduction, rather than an increase, in the directivity index. The directivity index of a linear array having a half-wavelength spacing will remain equal to the logarithm of the number of array elements, as elements are added, only if the spacing remains constant. This, of course, requires that the array length increase.

An array having a spacing greater than one half-wavelength is valueless for many purposes. If the secondary maxima are to decrease progressively throughout one quadrant of the directivity pattern, as angle from the primary maximum response bearing increases, the spacing may not exceed this value.

3. Directivity Index of a Volume Array as a Function of Size

With this reminder of the manner in which the directivity index of a linear array varies with the number of elements and with frequency we may move on to look at surface and volume arrays. For this we first examine the case of an array of 8000 elements arranged within a cube. This may be considered as being formed of 20 parallel planar arrays, each of which is formed of 20 parallel linear arrays of 20 elements each. The spacing between elements in a line, between lines in a plane, and between planes are all equal.

The directivity patterns for linear and planar arrays have long been familiar to sonar engineers. The patterns of volume arrays are not so well known. At spacings of one half-wavelength or less a linear array has a directivity pattern which shows a relative response of unity for every bearing in a plane perpendicular to the line of the array, and for no other bearing. The pattern of a planar array shows a relative response of unity for each of the two bearings perpendicular to the plane of the array, and for no other bearing.

When we come to compute the directivity pattern of the volume array, making use of the product theorem, we find that the unit relative response bearing of a single element—a planar array—does not coincide with the unit relative response bearing of the hypothetical array formed by the acoustic centres of these elements. This means that there will be no bearing for which the responses of all 8000 nondirectional elements add in phase. To create such a bearing it is necessary to introduce electrical compensation. This is done by delaying the responses of the planar elements so that they combine in phase when receiving acoustic waves propagated along the line formed by their centres. When thus compensated the relative responses in both directions along this line will be unity when the spacing is one half-wavelength. To obtain a directivity pattern having secondary maxima which decrease progressively for 180 deg, as the angle from the primary maximum response bearing increases, it is necessary that the element spacing be one quarter-wavelength or less.

The relative response for any bearing of the compensated volume array may be computed by multiplying the relative response of one of the planar arrays by the relative response of the compensated linear array formed by the centres of these planar arrays. The general forms of the directivity patterns for each of these sub-arrays are well known. At spacings of one quarter-wavelength one is essentially the same as that of a linear transducer uniformly responsive over its entire length and the other as that of a square plate transducer uniformly responsive over its entire surface.

Because a spacing of exactly one quarter-wavelength was postulated for the computations, and because the number of planar arrays is a multiple of four, the response will be zero for a bearing 180 deg from the bearing for which the array of these subarrays has been compensated. Directivity patterns have been computed for planes passing through the bearing perpendicular to two opposite faces of the cube, for which all elements respond in phase. These patterns are shown in Fig. 2, in which relative response has been plotted as a function of co-latitude angle, θ, for various azimuth angles, ϕ.

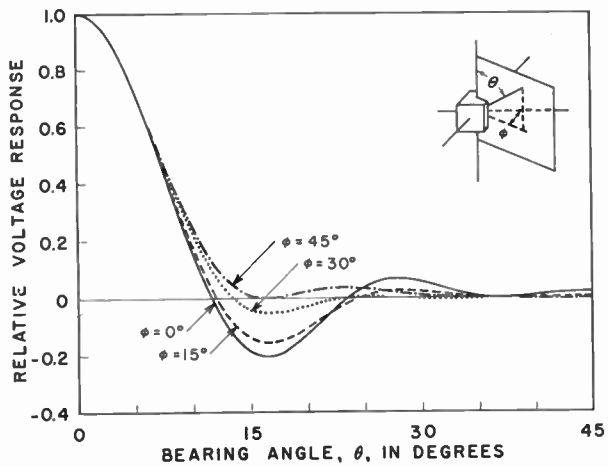


Fig. 2. Directivity patterns for a cubic array. This array has 20 × 20 × 20 elements uniformly spaced at quarter-wavelength intervals and compensated to respond in phase for energy propagated along the bearing $\theta = 0°$.

Since the relative response of the major lobe of the compensated linear array remains near unity for the range of bearings occupied by the major lobe of a planar array the major lobe of the compensated cubic array is nearly the same as that of a single planar array. A slight difference is found between the bearings of the first minor lobes. For the pattern plane parallel to two faces of the cube both the planar array and the volume array show this lobe as having a maximum response approximately 14 dB below the maximum response of the major lobe. As this plane is rotated the response of the first minor lobe decreases, becoming zero when the pattern plane, having been rotated through 45 deg, passes through an edge of the cube.

For any position of the pattern plane the secondary maxima decrease rapidly as angle from the maximum response bearing increases. They show, in general, much lower responses than are usually found with linear and planar arrays of comparable size. For bearing angles in the first 45 deg from the maximum response bearing the patterns of the cubical array resemble those of a square planar array for all positions of the pattern plane. For greater bearing angles the responses of the cubical array are virtually negligible.

The directivity indices for a single planar array, and for the volume array made up of the compensated planar arrays, have been computed by means of the general formula for directivity factor already mentioned. These computations are for spacings of one half-wavelength and less.

The results of these computations are shown in Fig. 3. Here directivity index in decibels has been plotted against element spacing in cycles. As for Fig. 1, values shown for the abscissa are proportional to either geometric length or to frequency.

One important result is the evidence that the directivity indices of both the planar array and the volume array increase at the rate of 6 dB/octave. It has long been known that the directivity factor of an array occupying a given area varies inversely as the square of the frequency. This, together with the knowledge that the directivity factor of a linear array varies

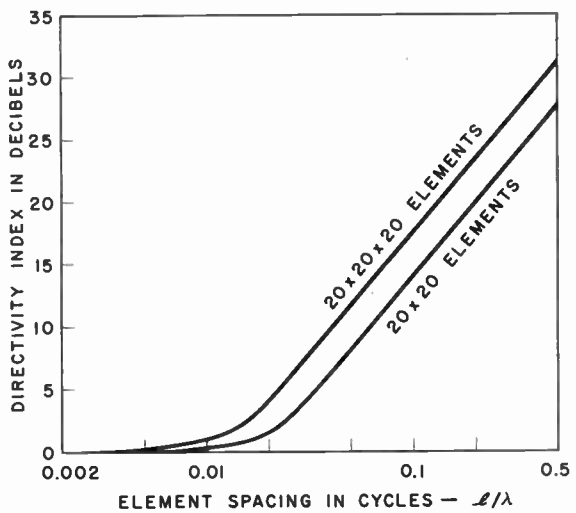


Fig. 3. The directivity indices of a planar and of a cubic array. The directivity factor of both of these arrays is inversely proportional to the second power of the linear dimension, for a fixed frequency, or of the frequency, for a fixed dimension.

inversely as the first power of the frequency, has led many to suppose that the directivity factor of a volume array would vary inversely as the third power of the frequency, or that its directivity index would increase at the rate of 9 dB/octave. This did, indeed, at one time, seem a reasonable supposition. It is not borne out by the behaviour of the particular volume array here under examination.

The second result has to do with the increase in the directivity index in going from a single planar array to an assembly of 20 such arrays, the frequency—or spacing—being the same for both. In the vicinity of

the quarter-wavelength point the increase computed is 3·5 dB. Of this difference 3 dB is to be attributed to the existence of two relative response bearings of unity for the planar array whereas the volume array has only one. It is clear that the increase in the directivity index in going from 400 elements to 8000 elements falls far short of the 13 dB given by the logarithm of the ratio of these numbers.

4. Directivity of a Volume Array as Compared with its Surface

There has been much speculation, in connection with volume arrays, as to whether the directional discrimination of an array of elements occupying some given volume is significantly greater than would be obtained from the elements on its surface. It is argued that no information reaches the interior of the volume which has not been intercepted by its surface. A brief examination of this question has been made for the case of a sphere.

The sphere, like the cube, will require the use of compensating delays to bring the responses of all elements into phase for some designated bearing. If this condition is to exist for a single bearing only the spacing must not greatly exceed one quarter-wavelength. Computations have been made for a sphere having a diameter of five wavelengths and for elements having a quarter-wavelength spacing. This calls for approximately 310 nondirectional elements on the plane surface bounded by a great circle. There would be 1240 elements on the surface of the sphere, and slightly more than 4200 in its volume. These numbers will vary somewhat depending on how "uniform spacing" within a spherical volume is defined.

We begin with the planar array bounded by a great circle. This will be considered as one of the sub-arrays which, when compensated, form the volume array. If 310 nondirectional elements are distributed over this surface the form of the directivity pattern will depend upon their arrangement. There is, moreover, no straightforward method for achieving a truly uniform distribution throughout the area within a circular boundary. Actually, a random distribution m ight be advantageous. Although the directivity pattern is affected significantly by the manner in which the elements are arranged the directivity index is believed to be affected very little, if at all. This question will be examined later.

Because of the size of the array and the closeness of the element spacing the directivity index computed for a circular piston having a diameter of five wavelengths and the index computed for 310 nondirectional elements distributed in some specified manner over a surface of the same area will be sufficiently alike for our present purpose. Although the directivity index of this circular piston may be computed by a well-established formula relative response as a function of bearing must also be computed in order later to compute the directivity pattern for the compensated spherical array. The pattern of a circular piston is too well known to need a detailed description. It is sufficient to note that this array, like the square planar array, has relative response bearings of unity in both directions along the perpendicular to its surface.

Elements may be distributed throughout the volume of the sphere by arranging a series of parallel planar arrays, of appropriate radii, spaced along a diameter of the sphere at quarter-wavelength intervals. There will be 21 such planar arrays, arranged as 10 pairs. It may be shown, when acoustic energy is received over any given bearing, that the responses of the two members of a pair, as connected by the required electrical delay, have quadrature components which are equal and opposite. The resulting response is in phase with the response of the centre element. This greatly simplifies the computation of the directivity pattern of the volume array. All that is needed is to sum the responses, for any bearing, of the pairs of planar arrays after adjusting for arrival angle and relative area. The results are shown graphically in Fig. 4.

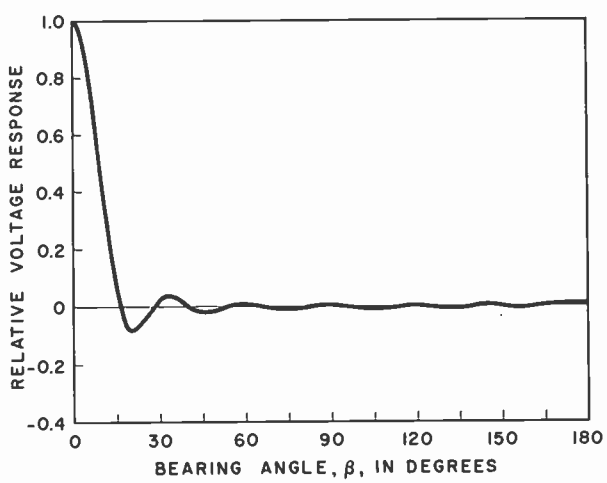


Fig. 4. Directivity pattern for a spherical array. The pattern for an array occupying the volume of a sphere having a diameter of five wavelengths, with a spacing of one quarter-wavelength.

The directivity pattern computed in this manner shows secondary maxima which decrease progressively for 180 deg as angle from the primary maximum increases. This is characteristic of a compensated array having a quarter-wavelength spacing. The secondary lobes pass through 12 maxima as the bearing angle passes through 180 deg, the response of the first being 22 dB below the primary maximum. Each secondary maximum shown by the directivity pattern

indicates the magnitude of the response for any bearing lying on a conical locus.

It remains to compute the directivity pattern corresponding to the elements on the surface of this sphere. For this we replace the 21 planar arrays already examined by 21 circular rings, positioned at the intersections of the planes of the planar arrays and the spherical surface. The directivity pattern of each of these continuous ring elements, like the continuous surfaces approximating the planar arrays, may be computed by an established formula. The pattern for one of these rings, like the pattern for one of the planar elements, has two primary maximum response bearings. These extend, in opposite directions, along the perpendicular to the plane of the ring. As with the spherical surface, these rings must be compensated to bring their responses into phase for acoustic energy propagated along the line through their centres.

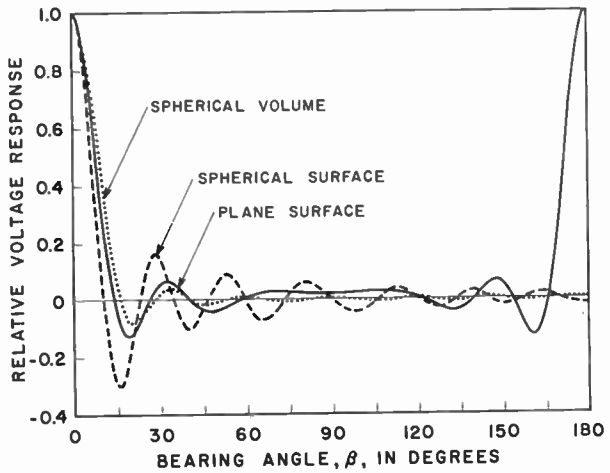


Fig. 5. The directivity patterns of spherical arrays. The patterns of arrays occupying a circular plane surface, a spherical surface, and a spherical volume, each having a diameter of five wavelengths. The element spacing assumed is one quarter-wavelength.

The directivity patterns for the plane bounded by a great circle, for the surface of the sphere, and for its volume, are shown in Fig. 5. From these patterns we may compute the directivity indices of these arrays. For comparison we may also compute the index for the single ring lying on the circumference of a great circle. The results are given in Table 1. In comparing the directivity indices it is to be remembered that the circular ring and circular plane both have two bearings for which the relative response is unity whereas the spherical surface and spherical volume have only one.

It is quite evident that in the frequency range where these arrays have useful directional characteristics the directivity indices are neither equal to nor proportional to the logarithms of the numbers of elements. It is interesting to compare the directivity patterns in the light of this knowledge. When we do so it appears at once that the patterns having the narrower major lobes also have the greater relative response for their first minor lobes. This is a well-known characteristic of shaded arrays. Given a plane surface array occupying a circular area an increase in the sensitivity of the elements near the centre will result in widening the major lobe and reducing the magnitude of the secondary maxima. Conversely, an increase in the sensitivity of elements near the circumference will result in narrowing the major lobe and increasing the magnitude of the secondary maxima. It is thus known that changes in the shading of a circular array are accompanied by changes in the directivity pattern of the same nature as those found in going from the circular array to the spherical surface and to the spherical volume.

Table 1

Characteristics of Circular and Spherical Arrays

	Equivalent no. of elements	Directivity index (dB)	Width of major lobe (deg)	Relative response of first secondary (dB)
Ring, lying on a great circle	63	15·5	17·6	— 7·9
Plane, bounded by a great circle	310	20·9	28·2	— 17·6
Spherical surface	1240	22·6	20·4	— 10·2
Spherical volume	4200	23·5	32·6	— 21·8

From the evidence obtained from the computations described it appears that the major effect of adding elementary transducers to those occupying the plane bounded by a great circle of the sphere postulated is to change its shading. It is clearly evident that the increase in the directional discrimination is not commensurate with the increase in the number of elements.

It must not be forgotten that these findings apply to frequencies for which the directivity patterns show not more than two bearings for which the relative response is unity. For higher frequencies there will be an increasing number of such bearings and the indications as to the bearing of some source of acoustic energy will be increasingly ambiguous.

5. Effect of a Specific Deviation from Ideal Symmetry

In the computational experiments which have been described the high-speed computer has been used to establish certain facts not readily evident by the inspection of some equation, or formula. We must use

great care in attempting to draw general conclusions from information obtained in this manner. It is not always easy, or even possible, to determine the extent to which the behaviour of a class may be inferred from the behaviour of a single sample. It is, in fact, often difficult to define the class to which some single sample belongs.

Many of the formulae and equations which we use daily, and which we speak of as expressing general relations, apply only within limited boundaries. For example, the familiar formula for the relative voltage response of a linear array of electro-acoustic transducers is

$$\frac{e_d}{e_o} = \frac{\sin n\psi}{n \sin \psi},$$

where

$$\psi = \frac{\pi s}{\lambda} \sin \beta.$$

Here

e_d = the r.m.s. voltage generated in the transducer in response to plane sinusoidal waves of given frequency propagated along a specified bearing,

e_o = the r.m.s. voltage generated in response to waves of the same frequency and amplitude propagated along the maximum response bearing,

n = the number of array elements,

s = the spacing between these elements,

λ = the wavelength of the propagated waves,

β = the angle between the bearing giving e_d and that giving e_o.

This equation applies only when the following conditions are satisfied:

1. All n elements are identical in response and impedance,
2. They all lie on the same straight line,
3. The separations, s, between them are exactly equal,
4. The acoustic waves reaching them along a given bearing, β, are ideally plane and monochromatic.

There are few actual arrays which fulfil all of these conditions. It is desirable, therefore, to inquire into the extent to which conclusions based on relations known to be valid under these restrictive conditions apply to arrays which do not meet them.

An equation for the relative response of an array of transducers is a statement of the result of summing a series of algebraic expressions having constant coefficients. Should any coefficient vary from element to element the expressions must be evaluated one by one and summed arithmetically. This usually calls for a high-speed computer.

We may obtain some idea of the consequences of deviations from uniform compliance with design specifications by computing the directivity pattern and the directivity index for a linear array of identical elements the spacings of which vary in a random manner. Except for spacing, all factors affecting the relative response will be assumed to conform exactly to prescribed specifications.

The array for which computations have been made has 21 nondirectional elements arranged along a straight line at a nominal spacing of four wavelengths. Each element has been displaced from the point which it would have occupied if all intervals had been equal by a distance and in a direction along the line which were determined by chance. The standard deviation of the actual spacing from the nominal value of four wavelengths was 0·5 wavelength.

Since the elements all lie on a straight line they all respond in phase to plane waves of acoustic energy propagated in a direction perpendicular to this line. The sum of the equal voltages generated in response to these waves is the reference voltage response. Since the actual distance of any element from the centre element is known the phase of the voltage generated in any element relative to the phase of the voltage generated in the centre element is known for plane waves propagated along any bearing. The ratio of the phasor sum of these voltages to the reference voltage is the relative voltage response for the bearing in question.

For a linear array of uniformly spaced elements having the length of 80 wavelengths postulated there would be 80 secondary maxima in each quadrant of the directivity pattern. When this pattern is drawn by plotting relative voltage response as a function of the sine of the bearing angle these maxima will occur midway between multiples of 0·0125. For a linear array of elements uniformly spaced at four-wavelength intervals four of these secondary maxima will have relative voltage response magnitudes of unity. These will occur at values of the sine of the bearing angle which are multiples of 0·25. The pattern appearing between the values 0 and 0·25 for the sine of the bearing angle is repeated in the three succeeding intervals of this magnitude.

Computations on the array having the non-uniform spacing postulated lead to a directivity pattern having 63 secondary maxima. Those occurring near values of the sine of the bearing angle which are multiples of 0·25 have responses which are 2, 6, 12, and 16 dB below the primary maximum response, or below the responses found at these bearings for the uniformly spaced array. These responses decrease in magnitude with increasing angle between the bearing in question and the primary maximum, at 90 deg to the line of the array. Although the positions of the elements on one

side of the array centre are not mirror images of those on the other the directivity patterns for the quadrants on either side of the primary maximum response bearing are mirror images of each other. A portion of the directivity pattern of this array is shown in Fig. 6. This portion applies to bearing angles for which the sines vary from 0·50 to 0·75. There are four such portions in each quadrant. The dotted line shows the pattern for a uniform spacing of exactly four wavelengths.

For the ideal uniform spacing there would have been eight secondary maxima, in one quadrant, for which the response would have been 16 dB below the primary maximum. For the non-uniform spacing postulated there are 46 maxima having greater responses than this.

Having computed relative voltage response as a function of bearing for the array of non-uniformly spaced elements its directivity factor may be found by integrating the square of this response over all bearings reaching the array through a spherical surface. Because arrays here in question all have an axis of acoustical symmetry which is also the primary maximum response bearing the general formula for directivity factor may be reduced to the simpler form

$$\eta_D = \int_0^{\pi/2} (e_d/e_o)^2 \cos\beta \, d\beta$$

$$= \int_0^1 (e_d/e_o)^2 \, d(\sin\beta)$$

It was because of the computational advantage of the second form of this integral that the directivity patterns have been plotted against the sine of the bearing angle.

When this integration is carried out the directivity index is found to be $(N_D) = 13{\cdot}2$ dB. When we apply the rule that the directivity factor is the reciprocal of the number of elements we see that $10 \log 21 = 13{\cdot}2$ dB also. When the pattern for the ideal uniformly spaced array was integrated in the same manner the directivity index was found to have this same value. For the case examined, therefore, we may say that for an actual array, in which the element positions depart from the ideal distribution generally assumed for

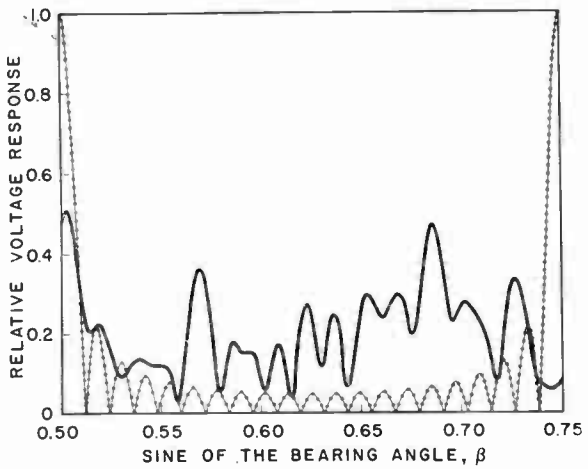


Fig. 6. Portion of the directivity pattern for a randomly spaced linear array. A portion of the pattern of an array of 20 elements having an average spacing of four wavelengths but varying from this value with a standard deviation of ± 0·5 wavelengths. The dotted line shows the pattern for a uniform spacing of exactly four wavelengths.

computations, the directivity pattern departs in many respects from the form usually considered characteristic of this type of array. The ability of the array to discriminate against unwanted interference, nevertheless, has remained unchanged.

The results of computations on an array having randomly spaced elements indicate that the bearing ambiguities usually found with arrays having elements uniformly spaced at separations greater than one half-wavelength, or than one quarter-wavelength in the case of compensated arrays, may be significantly reduced if the elements are randomly spaced. This is certainly true for the randomly spaced linear array.

The computations which have been reported here have answered a few of many questions raised as we attempt to increase the directional discrimination of sonar arrays. There are many more such questions to be answered. We are fortunate that the high-speed computer has arrived in time to assist in solving these problems.

Manuscript first received by the Institution on 18th June 1962 and in final form on 7th January 1963. (Paper No. 808/SS 22.)

POINTS FROM THE DISCUSSION

Mr. E. D. R. Shearman: Dr. Horton's paper has shown the lack of additional directional information provided by receiving elements distributed throughout a volume relative to that provided by elements on the bounding surface alone.

It occurs to me that the explanation of this may lie in Green's Theorem. This states that the distribution, within a particular bounded volume, of a scalar potential, such as the velocity potential defining a sound field, is known if the distribution of the potential and its normal

derivative over the bounding surface is known. This would suggest that samples of the field within the volume provide redundant information.

The author (*in reply*): I concur fully with Mr. Shearman's comment regarding Green's Theorem. This theorem is, of course, the justification for the argument to which I refer in my paper, that "no information reaches the interior of the volume which has not been intercepted by its surface."

Professor D. G. Tucker: The author shows that the directional pattern of a cubic distribution of omnidirectional transducers with phase compensation is given by the *product of the pattern* of a broadside planar array and a hypothetical endfire linear array. It is clear, therefore, that exactly the same pattern is obtained (for a receiving system) by *multiplying the signals* from a broadside planar array and an actual endfire linear array; the output signal is now, of course, a "d.c." whose amplitude varies in magnitude and polarity (thus tracing the directional pattern) as the array is rotated while looking at a single target. In Dr. Horton's example of a $20 \times 20 \times 20$ array, this would use only 420 elements instead of 8000. The multiplicative array, however, while giving the same directional pattern, would not give the same directional discrimination against multiple targets or noise.

The author (*in reply*): I agree with Professor Tucker that a system which delivered an output proportional to the product of the outputs of a single uncompensated planar array and of a linear array perpendicular to this planar array and compensated for endfire operation, would have the same directivity pattern as the cubic array in question. It would not have the same discrimination against isotropic noise. The discrimination of the actual cubic array is computed by taking the integral of the product of the two responses over all bearings; the discrimination of the system having a signal multiplying component is computed by taking the product of the integrals of the two responses, each taken over all bearings. The two are not the same, as Professor Tucker indicates.